А.И. Орлов

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ДАННЫХ

Учебник

Москва

Ай Пи Ар Медиа

2022

УДК 004.8
ББК 32.81
О-66

Автор:

Орлов А.И. — д-р экон. наук, д-р техн. наук, канд. физ.-мат. наук, проф. кафедры экономики и организации производства (ИБМ-2) Московского государственного технического университета имени Н. Э. Баумана

Орлов, Александр Иванович.

О-66 Искусственный интеллект: статистические методы анализа данных : учебник / А.И. Орлов. — Москва : Ай Пи Ар Медиа, 2022. — 843 с. — Текст : электронный.

ISBN 978-5-4497-1470-1

Учебник посвящен современным методам анализа статистических данных. В первой части рассмотрены основы выборочных исследований и основные задачи описания данных, оценивания и проверки гипотез. Статистические методы анализа числовых данных, многомерный статистический анализ и статистические методы анализа динамики обсуждаются во второй части с непараметрической точки зрения. Основные понятия теории статистического моделирования раскрываются в третьей части на примерах моделей экономики и управления (управления качеством, логистики, взаимовлияния факторов), экспертных исследований, медицины, социологии, демографии, истории, электротехники. Теоретическим инструментам, истории и перспективам развития статистических методов посвящена четвертая часть. Изложение соответствует рекомендациям Российской академии статистических методов.

Подготовлен с учетом требований Федерального государственного образовательного стандарта высшего образования.

Для студентов и преподавателей вузов, слушателей институтов повышения квалификации, структур второго образования и программ MBA («Мастер делового администрирования»), инженеров различных специальностей, менеджеров, экономистов, социологов, научных и практических работников, связанных с анализом данных.

Учебное электронное издание

ISBN 978-5-4497-1470-1

Технический редактор, компьютерная верстка *Ю.В. Семенова*
Обложка *С.С. Сизиумовой*

Подписано к использованию 02.11.2021. Объем данных 12,5 Мб.

Издание представлено в электронно-библиотечных системах
IPR BOOKS (www.iprbookshop.ru),
Библиокомплектатор (www.bibliocomplectator.ru)

Бесплатный звонок по России: **8-800-555-22-35**
Тел.: 8 (8452) 24-77-97, 8 (8452) 24-77-96

Отдел продаж и внедрения ЭБС:
доб. 206, 213, 144, 145
E-mail: sales@iprmedia.ru

Отдел комплектования ЭБС:
доб. 224, 227, 208
E-mail: mail@iprbookshop.ru

По вопросам приобретения издания обращаться:
доб. 208, 201, 222, 224
E-mail: izdat@iprmedia.ru, author@iprmedia.ru

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

В «Национальной стратегии развития искусственного интеллекта на период до 2030 года»[1] принято следующее определение: «... искусственный интеллект — комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе с методами машинного обучения), процессы и сервисы по обработке данных и поиску решений». В этом определении прямо не говорится про научную основу «комплекса технологических решений». По нашему мнению, в социально-экономической области в качестве такой основы можно использовать организационно-экономическое моделирование, включая высокие статистические технологии, в том числе нечисловую статистику, теорию и практику экспертных оценок, статистические методы анализа данных.

Автор занимается проблемами искусственного интеллекта около полвека (первые статьи напечатаны в 1972 г.). Настоящая книга посвящена важной составляющей искусственного интеллекта — статистическим методам анализа данных.

Статистические методы анализа данных активно применяются в технических исследованиях, экономике, теории и практике управления (менеджменте), социологии, медицине, геологии, истории и т.д. С результатами наблюдений, измерений, испытаний, опытов, с их анализом имеют дело специалисты во всех отраслях практической деятельности, почти во всех областях теоретических исследований. Настоящий учебник позволяет овладеть современными статистическими методами на уровне, достаточном для использования этих методов в научной и практической деятельности.

Содержание книги. Учебник состоит из введения и четырех частей, разбитых на 16 глав. Во введении обсуждается внутренняя структура статистических методов анализа данных — развитой области научно-практической деятельности.

Часть 1 посвящена основным постановкам задач анализа данных — методам выборочных исследований, описания данных, оценивания и проверки гипотез. В главе 1 обсуждаются проблемы организации выборочных исследо-

[1] URL: https://www.garant.ru/products/ipo/prime/doc/72738946/.

ваний на примере двух конкретных маркетинговых опросов. Разработаны модели случайных выборок, в том числе гипергеометрическая и биномиальная, методы доверительного оценивания доли и проверки однородности двух биномиальных выборок.

В главе 2 обсуждаются модели порождения данных, методы их описания с помощью таблиц и диаграммы, выборочных характеристик и эмпирической функции распределения, непараметрических оценок плотности (в пространствах произвольной природы). Показано, что распределение результатов наблюдений (испытаний, испытаний, анализов, опытов), как правило, отличается от нормального. Как следствие, в учебнике большое внимание уделено непараметрическим методам анализа статистических данных.

Глава 3 посвящена методам оценивания параметров и характеристик. В частности, разработаны и изучены одношаговые оценки, предназначенные для замены устаревших оценок максимального правдоподобия. Исследована асимптотика решений экстремальных статистических задач и устойчивость (робастность) статистических процедур. Оценивание для сгруппированных данных построено на основе современных вариантов формулы Эйлера — Маклорена и поправок Шеппарда.

Для проверки гипотез в главе 4 разработан метод моментов, реализованный на примере гипотезы согласия с гамма-распределением. Продемонстрирована крайняя неустойчивость параметрических методов отбраковки выбросов, приводящая к выводу о невозможности их научно обоснованного использования. Построена предельная теория непараметрических критериев, опирающаяся на метод приближения ступенчатыми функциями. Разработан метод проверки гипотез по совокупности малых выборок, предназначенный для применения в асимптотике растущей размерности, когда число неизвестных параметров растет вместе с объемом данных. Обсуждается проблема множественных проверок статистических гипотез, актуальная при разработке высоких статистических технологий анализа данных.

В части 2 рассматриваются конкретные статистические методы анализа данных различных типов. В главе 5 элементы выборки — это числа. Разобраны методы точечного и доверительного непараметрического оценивания основных характеристик распределения — математического ожидания, медианы, дисперсии, среднего квадратического отклонения, коэффициента вариации. Подробно рассмотрены методы проверки однородности характеристик двух независимых выборок, обоснована необходимость использования непараметрического критерия Крамера-Уэлча вместо статистики Стьюдента, опирающейся на нереалистические предположения нормальности результатов измерений и совпадения дисперсий элементов двух выборок. Изучены свойства двухвыборочного критерия Вилкоксона, обосновано использование состоятельных критериев проверки однородности независимых выборок. Разработаны методы проверки однородности связанных выборок, в том числе на основе критериев проверки гипотезы симметрии. Обсуждается взаимосвязь реальных и номинальных уровней значимости в задачах проверки статистических гипотез.

Глава 6 посвящена основным постановкам многомерного статистического анализа. Рассмотрены линейный (Пирсона) и непараметрические (Спирмена, Кендалла) коэффициенты парной корреляции. Подробно обсуждается задача восстановления линейной зависимости между двумя переменными на основе непараметрического метода наименьших квадратов. Рассмотрены основы линейного регрессионного анализа, статистических методов классификации и методов снижения размерности. В конце главы 6 разобран индекс инфляции и его применения, в частности, при анализе уровня жизни и доходности банковских вкладов.

Статистические методы анализа динамики обсуждаются в главе 7, в том числе методы анализа и прогнозирования временных рядов и системы эконометрических уравнений. В учебник включены оригинальные подходы к оцениванию длины периода и периодической составляющей сигналов.

Статистические методы анализа нечисловых и интервальных данных не обсуждаются в настоящем учебнике. Они рассмотрены в иных изданиях, в частности, в выпущенных издательством «Экзамен» наших учебниках «Эконометрика» (2002, 2003, 2004), «Прикладная статистика» (2006, наиболее полное изложение) и «Теория принятия решений» (2006), а также в выпущенном издательством МГТУ им. Н.Э. Баумана учебнике «Организационно-экономическое моделирование. Ч.1. Нечисловая статистика».

Наибольшая по объему часть 3, включающая 6 глав, посвящена вероятностно-статистическому моделированию в различных областях применения. В главе 8 рассмотрены основные понятия теории статистического моделирования, затем обсуждаются демографические модели, статистические модели движения товарных потоков в процессе работы склада (в другой терминологии — модели логистики). Большое внимание уделено статистическому моделированию исторических процессов, позволившему существенно уточнить хронологию древнего мира и средневековья. Завершается глава вероятностно-статистическим моделированием помех, создаваемых электровозами, с целью сокращения расходов на защиту проводных линий связи.

В главе 9 подробно описан подход к моделированию взаимовлияний факторов методом ЖОК (название составлено из первых букв фамилий исследователей: Жихарев — Орлов — Кольцов). На основе этого метода разработана система моделей налогообложения и проанализированы макроэкономические балансовые соотношения. Рассмотрена эконометрическая база метода — моделирование и анализ многомерных временных рядов.

Статистические модели управления качеством — предмет обсуждения в главе 10. От основ статистического контроля качества переходим к асимптотической теории одноступенчатых планов, а затем — к практическому применению статистического контроля. Рассмотрен весь комплекс статистических методов управления качеством, в том числе методы обнаружения разладки с помощью контрольных карт, весьма актуальные не только для организации производства, но и в менеджменте.

В главе 11 речь идет о статистическом моделировании в экспертных исследованиях. Приведены примеры процедур экспертных оценок, выделены основные стадии экспертного опроса. Для построения математической теории экспертных технологий важна общенаучная теория измерений. В качестве примера ее применения получены правила выбора вида средних величин в зависимости от типов шкал, в которых измерены ответы экспертов. Обсуждается использование методов средних арифметических и медиан баллов в сочетании с процедурами согласования кластеризованных ранжировок. Кратко рассмотрены математические методы анализа экспертных оценок, в частности, расстояние Кемени и медиана Кемени в пространствах бинарных отношений.

Рассказ о статистических моделях в медицине (глава 12) начинается с обсуждения нового компьютерно-статистического мышление врача, основанного на методах «доказательной медицины». Рассмотрено применение медико-статистических технологий в научных медицинских исследованиях. Особое внимание уделено проблемам внедрения высоких статистических технологий.

Завершаем часть 3 обсуждением статистических методов в социологии (глава 13). Проанализировано развитие статистического инструментария отечественных социологов за последние 30 лет. Подробнее обсуждаются некоторые математические методы в социологии — перспективы применения люсианов, асимптотика квантования и выбор числа градаций в социологических анкетах. Рассмотрен ряд практических применений статистических методов в социологии — в социометрических исследованиях, рассматриваемых как эффективный инструмент менеджера, в выборочных исследованиях научных организаций, в изучении способных к математике школьников.

Заключительная часть 4 посвящена инструментам, истории и перспективам развития статистических методов. В главе 14 кратко рассмотрены такие теоретические инструменты статистических методов, постоянно используемые в предыдущих главах учебника, как законы больших чисел, центральные предельные теоремы, теоремы о наследовании сходимости, метод линеаризации, принцип инвариантности. В конце главы рассмотрены проблемы устойчивости статистических выводов и принцип уравнивания погрешностей.

О развитии статистических методов кратко рассказываем в главе 15. Обсуждаются основные этапы становления статистических методов (от Книги Чисел в Библии до наших дней). Статистические методы в России рассмотрены на примерах исследований А.Н. Колмогорова, Б.В. Гнеденко, Н.В. Смирнова, Л.Н. Большева, В.В. Налимова. Завершается глава рассказом о дискуссии 1980-х гг., посвященной предмету и содержанию прикладной статистики.

В завершающей книгу главе 16 речь идет о современных статистических методах и перспективах их развития. Выделены «точки роста» рассматриваемой научно-практической дисциплины. Введено понятие «высокие статистические технологии» и обоснована необходимость из развития. Обсуждается использование информационных технологий при анализе статистических данных, рассмотрены современные компьютерно-статистические методы. Дано представление о методологии статистических методов. Сформулированы основные нерешенные проблемы статистических методов.

В конце каждой главы приведен список литературных источников, контрольные вопросы и задачи, а также темы докладов, рефератов, исследовательских работ. Нумерация таблиц, рисунков, формул, теорем, примеров дана как по главам, так и по параграфам.

Автор настоящего учебника более 50 лет постоянно занимается статистическими методами. Как практик и как теоретик. В учебник включены теоретические и практические результаты, как достаточно давние (1970-х гг.), так и полученные в последние годы (вплоть до 2021 г.). Их происхождение и авторство заинтересованные читатели проследят по литературным ссылкам, которые пригодятся и для углубленного изучения материала. В конце учебника помещена краткая информация о деятельности автора как научного работника и преподавателя.

Общее количество статей и книг по статистическим методам давно превысило 10^6, из них актуальными к настоящему времени являются, по нашей оценке, не менее 10^5. Конкретный специалист может овладеть несколькими тысячами из них. Следовательно, ни один исследователь не может претендовать на знакомство более чем с 2–3 % актуальных публикаций, и в любом учебнике содержится лишь небольшая часть знаний, накопленных в области разработки и применения статистических методов. Однако автор надеется, что наиболее важные подходы, идеи, результаты и алгоритмы расчетов включены в учебник. Эта надежда основана на собственном почти полувековом опыте теоретической и практической работы в области статистических методов, на совокупном опыте членов научных сообществ, скрупулезном анализе положения в нашей научно-прикладной дисциплине при создании Всесоюзной статистической ассоциации, Российской ассоциации статистических методов и Российской академии статистических методов.

В отличие от учебной литературы по математическим дисциплинам, в настоящей книге практически отсутствуют доказательства. Однако в нескольких случаях мы сочли целесообразным их привести. При первом чтении доказательства теорем можно пропустить.

О роли литературных ссылок в учебнике необходимо сказать достаточно подробно. Прежде всего, эта книга представляет собой замкнутый текст, не требующий для своего понимания ничего, кроме знания стандартных учебных курсов по высшей математике (включая теорию вероятностей и математическую статистику), которые преподаются студентам большинства вузов на первом и втором годах обучения. Зачем же нужны ссылки? Доказательства всех приведенных в учебнике теорем приведены в ранее опубликованных статьях и монографиях. Дотошный читатель, в частности, при подготовке рефератов и при желании глубже проникнуть в материал учебника, может обратиться к приведенным в каждой главе спискам цитированной литературы. Каждая глава учебника — это введение в большую область статистических методов. Приведенные литературные ссылки помогут читателям выйти на передний край теоретических и прикладных работ, познакомиться с доказательствами теорем, включенных в учебник. За многие десятилетия накопились большие книжные богатства, и их надо активно использовать.

Включенные в учебник материалы прошли многолетнюю и всестороннюю проверку. Кроме МГТУ им. Н.Э. Баумана, они использовались при преподавании во многих других отечественных и зарубежных образовательных структурах.

Издательством «Экзамен» в 2002, 2003 и 2004 гг. был выпущен учебник «Эконометрика» А.И. Орлова (три издания), в 2006 г. — его же учебники «Прикладная статистика» и «Теория принятия решений». Это говорит об актуальности тематики настоящего учебника, поскольку под эконометрикой понимают применение статистических методов в экономике и управлении (менеджменте), многие статистические методы входят в прикладную статистику и активно используются при разработке и принятии управленческих решений.

Учебник включен в серию книг «Искусственный интеллект», поскольку в нем рассматриваются современные методы анализа данных, соответствующие последним научным достижениям отечественной вероятностно-статистической школы. Отметим, что субъективные экспертные данные нет оснований противопоставлять объективным результатам измерений (наблюдений, испытаний, анализов, опытов, обследований), поскольку для их описания и анализа используются одни и те же вероятностно-статистические методы и модели.

Для кого написан учебник? Он предназначен для студентов (бакалавров, магистрантов) и аспирантов различных специальностей, прежде всего технических, управленческих и экономических, слушателей институтов повышения квалификации, структур послевузовского (в том числе второго) образования, в частности, программ MBA («Мастер делового администрирования»), преподавателей вузов, сотрудников научно-исследовательских организаций и подразделений.

Учебник будет полезен инженерам, менеджерам, экономистам, социологам, биологам, медикам, психологам, историкам, другим специалистам, самостоятельно повышающим свой научный уровень. Короче говоря, всем научным и практическим работникам, имеющим отношение к анализу данных.

Учебник может быть использован при изучении дисциплин, полностью или частично посвященных методам анализа результатов наблюдений (измерений, испытаний, опытов). Типовые названия таких дисциплин — «Организационно-экономическое моделирование», «Статистика», «Статистические методы», «Прикладная статистика», «Эконометрика», «Анализ данных», «Многомерный статистический анализ», «Общая теория статистики», «Планирование эксперимента», «Биометрика», «Теория принятия решений», «Управленческие решения», «Экономико-математическое моделирование», «Математические методы прогнозирования», «Прогнозирование и технико-экономическое планирование», «Хемометрия», «Математические методы в экономике», «Маркетинговые исследования», «Математические методы оценки», «Математические методы в социологии», «Математические методы в геологии» и т.п.

Математикам — специалистам по теории вероятностей и математической статистике — эта книга также может быть интересна и полезна, поскольку в ней описан современный взгляд на статистические методы и прикладную математическую статистику, основные подходы и результаты в этой области, открывающие большой простор для дальнейших математических исследований.

Благодарности. Книга написана в традициях отечественной вероятностно-статистической школы, начало ее современному этапу развития положил академик АН СССР А.Н. Колмогоров, а в области математической статистики — член-корреспондент АН СССР Н.В. Смирнов. Автор искренне благодарен своим учителям — академику АН УССР Б.Г. Гнеденко, члену-корреспонденту АН СССР Л.Н. Большеву, проф. В.В. Налимову.

Содержание учебника в своих основных чертах соответствует коллективному мнению отечественных специалистов. В 1990 гг. была создана Всесоюзная статистическая ассоциация (ВСА), состоящая из четырех секций. Руководитель секции статистических методов А.И. Орлов был избран вице-президентом ВСА. В XXI в. развитие прикладной статистики продолжается в рамках Российской ассоциации статистических методов и Российской академии статистических методов. Автор искренне благодарен своим многочисленным коллегам, с которыми посчастливилось вместе работать в рамках наших профессиональных объединений.

По ряду причин исторического характера основное место публикаций научных работ по статистическим методам в нашей стране — раздел «Математические методы исследования» журнала «Заводская лаборатория. Диагностика материалов», в котором напечатано более 90 научных статей автора, большинство из которых нашло отражение в настоящем учебнике. Многие работы опубликованы в Политематическом сетевом электронном научном журнале Кубанского государственного аграрного университета (Научном журнале КубГАУ), а также в журналах «Контроллинг» и «Инновации в менеджменте».

Хотелось бы выразить признательность всему коллективу кафедры «Экономика и организация производства» и в целом факультета «Инженерный бизнес и менеджмент» Московского государственного технического университета им. Н.Э. Баумана, декану и членам Ученого совета, поддержавшим инициативу о введении статистических методов в учебный процесс. Автор искренне признателен заведующему кафедрой «Экономика и организация производства» проф. С.Г. Фалько за постоянную поддержку проектов по разработке и внедрению эконометрических и статистических курсов, декану проф. И.Н. Омельченко за помощь в издании книг и совместные научные исследования.

Автор благодарен своим многочисленным коллегам, слушателям и студентам, прежде всего различных образовательных структур Московского государственного технического университета им. Н.Э. Баумана, Московского физи-

ко-технического института, Российской экономической академии им. Г.В. Плеханова и Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации (программа «Топ-Менеджер»), за полезные обсуждения.

С текущей научной информацией по статистическим методам анализа данных можно познакомиться на сайте «Высокие статистические технологии» http://orlovs.pp.ru и его форуме. Большой объем информации по рассматриваемым в учебнике вопросам содержит выходящий с 2000 г. электронный еженедельник «Эконометрика» (электронная газета кафедры «Экономика и организация производства» научно-учебного комплекса «Инженерный бизнес и менеджмент» МГТУ им. Н.Э. Баумана (http://subscribe.ru/catalog/science.humanity.econometrika). Автор искренне благодарен своему сыну А.А. Орлову, разработчику и администратору сайта, редактору еженедельника за многолетний энтузиазм.

Условия для написания книги создала моя любимая жена Л.А. Орлова. Спасибо!

Автор искренне благодарен сотрудникам издательства Ай Пи Ар Медиа Юлии Вадимовне Ермоловой, Юлии Валентиновне Семеновой и Анастасии Дмитриевне Талмаевой за большую работу по подготовке рукописи учебника к публикации.

Включенный в учебник материал дает представление о теории и практике статистических методов анализа данных, соответствующее общепринятому в мире. Изложение доведено до современного уровня научных исследований в этой области. Конечно, возможны различные точки зрения по тем или иным частным вопросам. Автор будет благодарен читателям, если они сообщат свои вопросы и замечания по адресу издательства или непосредственно автору по электронной почте E-mail: prof-orlov@mail.ru.

ВВЕДЕНИЕ. СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ДАННЫХ КАК ОБЛАСТЬ НАУЧНО-ПРАКТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ

Статистические методы анализа данных применяются практически во всех областях деятельности человека. Их используют всегда, когда необходимо получить и обосновать какие-либо суждения о группе (объектов или субъектов) с некоторой внутренней неоднородностью.

Целесообразно выделить три вида научной и прикладной деятельности в области статистических методов анализа данных (по степени специфичности методов, сопряженной с погруженностью в конкретные проблемы):

а) разработка и исследование методов общего назначения, без учета специфики области применения;

б) разработка и исследование статистических моделей реальных явлений и процессов в соответствии с потребностями той или иной области деятельности;

в) применение статистических методов и моделей для статистического анализа конкретных данных.

Кратко рассмотрим три только что выделенных вида научной и прикладной деятельности. По мере движения от а) к в) сужается широта области применения конкретного статистического метода, но при этом повышается его значение для анализа конкретной ситуации. Если работам вида а) соответствуют научные результаты, значимость которых оценивается по общенаучным критериям, то для работ вида в) основное — успешное решение конкретных задач той или иной области применения (техники и технологии, экономики, социологии, медицины и др.). Работы вида б) занимают промежуточное положение, поскольку, с одной стороны, теоретическое изучение свойств статистических методов и моделей, предназначенных для определенной области применения, может быть весьма сложным и математизированным (см., например, монографию [1]), с другой — результаты представляют не всеобщий интерес, а лишь для некоторой группы специалистов. Можно сказать, что работы вида б) нацелены на решение типовых задач конкретной области применения.

Прикладная статистика. Статистические методы анализа данных, относящиеся к группе а), обычно называют методами прикладной статистики. Таким образом, прикладная статистика — это наука о том, как обрабатывать данные произвольной природы, без учета их специфики [2].

Математическая основа прикладной статистики и статистических методов анализа данных в целом — это математическая наука, известная под названием «теория вероятностей и математическая статистика». Следует подчеркнуть, что прикладная статистика — другая область знаний, чем математическая статистика. Это очень четко проявляется в процессе обучения. Курс математической статистики состоит в основном из доказательств теорем, в то время как в курсах статистических методов основное — методология анализа данных и алгоритмы расчетов, а теоремы приводятся для обоснования этих алгоритмов, доказательства же, как правило, опускаются (их можно найти в научной литературе). Так построен и настоящий учебник.

Прикладная статистика — одна из статистических наук, она не относится к математике. Внутренняя структура статистики как науки была выявлена и обоснована при создании в 1990 г. Всесоюзной статистической ассоциации (см., например, статью [3]). Прикладная статистика — методическая дисциплина, являющаяся центром, идейным ядром статистики. Внутри прикладной статистики выделяют задачи описания данных, оценивания и проверки гипотез, рассмотренные в части 1 настоящего учебника.

Описание вида данных и, при необходимости, механизма их порождения — начало любого статистического исследования. Отметим, что для описания данных применяют как детерминированные, так и вероятностные методы. С помощью детерминированных методов можно проанализировать только те данные, которые имеются в распоряжении исследователя. Например, с их помощью получены таблицы, рассчитанные органами официальной государственной статистики на основе представленных предприятиями и организациями статистических отчетов. Перенести полученные результаты на более широкую совокупность, использовать их для предсказания и управления можно лишь на основе вероятностно-статистического моделирования. Поэтому в математическую статистику часто включают лишь методы, опирающиеся на теорию вероятностей, оставляя детерминированные методы экономической учебной дисциплине «Общая теория статистики».

Мы не считаем возможным противопоставлять детерминированные и вероятностно-статистические методы. Мы рассматриваем их как последовательные этапы статистического анализа. На первом этапе необходимо проанализировать имеющие данные, представить их в удобном для восприятия виде с помощью таблиц и диаграмм. Затем статистические данные целесообразно проанализировать на основе тех или иных вероятностно-статистических моделей. Отметим, что возможность более глубокого проникновения в суть реального явления или процесса обеспечивается разработкой адекватной математической модели.

В простейшей ситуации статистические данные — это значения некоторого признака, свойственного изучаемым объектам. Значения могут быть количественными или представлять собой указание на категорию, к которой можно отнести объект. Во втором случае говорят о качественном признаке.

При измерении по нескольким количественным или качественным признакам в качестве статистических данных об объекте получаем вектор. Его можно рассматривать как новый вид данных. В таком случае выборка состоит из набора векторов. Есть часть координат — числа, а часть — качественные (категоризованные) данные, то говорим о векторе разнотипных данных.

Одним элементом выборки, т.е. одним измерением, может быть и функция в целом. Например, описывающая динамику показателя, т.е. его изменение во времени, — электрокардиограмма больного или амплитуда биений вала двигателя. Или временной ряд, описывающий динамику показателей определенной фирмы. Тогда выборка состоит из набора функций.

Элементами выборки могут быть и иные математические объекты. Например, бинарные отношения. Так, при опросах экспертов часто используют упорядочения (ранжировки) объектов экспертизы — образцов продукции, инвестиционных проектов, вариантов управленческих решений. В зависимости от регламента экспертного исследования элементами выборки могут быть различные виды бинарных отношений (упорядочения, разбиения, толерантности), множества, нечеткие множества и т.д.

Итак, математическая природа элементов выборки в различных задачах прикладной статистики может быть самой разной. Однако можно выделить два класса статистических данных — числовые и нечисловые. Соответственно прикладная статистика разбивается на две части — числовую статистику и нечисловую статистику.

Числовые статистические данные — это числа, вектора, функции. Их можно складывать, умножать на коэффициенты. Поэтому в числовой статистике большое значение имеют разнообразные суммы. Математический аппарат анализа сумм случайных элементов выборки — это (классические) законы больших чисел и центральные предельные теоремы (см. главу 14).

Нечисловые статистические данные — это категоризованные данные, вектора разнотипных признаков, бинарные отношения, множества, нечеткие множества и др. Их нельзя складывать и умножать на коэффициенты. Поэтому не имеет смысла говорить о суммах нечисловых статистических данных. Они являются элементами нечисловых математических пространств (множеств). Математический аппарат анализа нечисловых статистических данных основан на использовании расстояний между элементами (а также мер близости, показателей различия) в таких пространствах. С помощью расстояний определяются эмпирические и теоретические средние, доказываются законы больших чисел, строятся непараметрические оценки плотности распределения вероятностей, решаются задачи диагностики и кластерного анализа, и т.д. (см. [2]).

В прикладных исследованиях используют статистические данные различных видов. Это связано, в частности, со способами их получения. Например, если испытания некоторых технических устройств продолжаются до определенного момента времени, то получаем т.н. *цензурированные* данные, состоящие из набора чисел — продолжительности работы ряда устройств до отказа, и информации о том, что остальные устройства продолжали работать в момент окончания испытания. Цензурированные данные часто используются при оценке и контроле надежности технических устройств.

Сведем информацию об основных областях прикладной статистики в табл. 1. Отметим, что модели порождения цензурированных данных входят в состав каждой из рассматриваемых областей.

Таблица 1

Области прикладной статистики

№ п/п	Вид статистических данных	Область прикладной статистики
1	Числа	Статистика (случайных) величин
2	Конечномерные вектора	Многомерный статистический анализ
3	Функции	Статистика случайных процессов и временных рядов
4	Объекты нечисловой природы	Нечисловая статистика (статистика нечисловых данных, статистика объектов нечисловой природы)

В части 2 настоящего учебника рассматриваем статистические методы анализа данных первых трех типов. Это ограничение вызвано тем отмеченным выше обстоятельством, что математический аппарат для анализа данных нечисловой природы — существенно иной, чем для данных в виде чисел, векторов и функций. Статистика нечисловых данных рассмотрена в [2, 4].

Вероятностно-статистическое моделирование. При применении статистических методов в конкретных областях знаний и отраслях народного хозяйства получаем научно-практические дисциплины типа «статистические методы в промышленности», «статистические методы в медицине» и др. С этой точки зрения эконометрика — это «статистические методы в экономике» [4]. Эти дисциплины группы б) обычно опираются на вероятностно-статистические модели, построенные в соответствии с особенностями области применения.

Основная (как по идейному наполнению, так и по объему) часть 3 настоящего учебника посвящена статистическим методам и вероятностно-статистическому моделированию в различных областях деятельности — в технико-экономических исследованиях (логистике, управлении качеством, электротехнике), экономике и управлении (налогообложении, экспертных оценках, маркетинге), демографии, истории, медицине, социологии.

Весьма поучительно сопоставить вероятностно-статистические модели, применяемые в различных областях, обнаружить их близость и вместе с тем констатировать некоторые различия. Так, видна близость постановок задач и применяемых для их решения статистических методов в таких областях, как научные медицинские исследования, конкретные социологические исследования и маркетинговые исследования, или, короче, в медицине, социологии и маркетинге. Они часто объединяются вместе под названием «выборочные исследования».

Отличие выборочных исследований от экспертных проявляется, прежде всего, в числе обследованных объектов или субъектов — в выборочных исследованиях речь обычно идет о сотнях, а в экспертных — о десятках. Зато технологии экспертных исследований гораздо изощреннее. Еще более выражена специфика в демографических или логистических моделях, при обработке нарративной (текстовой, летописной) информации или при изучении взаимовлияния факторов. Ряд иных полезных моделей рассмотрен нами в [5–7].

При отборе вероятностно-статистических моделей для включения в учебник автор во многом исходил из имеющегося у него опыта решения конкретных прикладных задач, а также старался не повторять уже известный в литературе материал. Поэтому в учебнике не рассматриваются вопросы надежности и безопасности технических устройств и технологий, теории массового обслуживания, не описываются изощренные системы эконометрических уравнений, поскольку они подробно рассмотрены, например, в ставших классическими монографиях [8–10].

Статистический анализ конкретных данных. Применение статистических методов и моделей для статистического анализа конкретных данных тесно привязано к проблемам соответствующей области. Результаты третьего из выделенных видов научной и прикладной деятельности находятся на стыке дисциплин. Их можно рассматривать как примеры практического применения статистических методов, как мы и делаем в настоящем учебнике. Но не меньше оснований относить их к соответствующей области деятельности человека.

Например, результаты опроса потребителей растворимого кофе (см. главу 1 данной книги) естественно отнести к маркетингу (что мы и делаем, читая лекции по маркетинговым исследованиям). Исследование динамики роста цен с помощью индексов инфляции, рассчитанных по независимо собранной информации (раздел 6.6), представляет интерес прежде всего с точки зрения экономики и управления народным хозяйством (как на макроуровне, так и на уровне отдельных организаций).

Примеры практического применения статистических методов включены практически во все главы учебника. Вполне естественно, что при отборе примеров предпочтение отдавалось тем исследованиям, в которых автор принимал непосредственное участие. Однако надо подчеркнуть, что описание примеров практического применения статистических методов адаптировано для учебных нужд. Заказчики прикладных исследований получают отчеты, в которых проблемы соответствующих областей деятельности рассмотрены гораздо более подробно. Примером такого отчета является монография [5], посвященная подходам к проблеме вероятностно-статистического моделирования процессов налогообложения.

О высоких статистических технологиях. Термин «высокие технологии» популярен в современной научно-технической литературе. Он используется для обозначения наиболее передовых технологий, опирающихся на последние достижения научно-технического прогресса. Есть такие технологии и среди технологий статистического анализа данных — как в любой интенсивно развивающейся научно-практической области. Они подробно обсуждаются в настоящем учебнике.

Обсудим этот пока не вполне привычный термин (он был введен в статье [11], опубликованной в 2003 г.). Каждое из трех слов несет свою смысловую нагрузку.

«Высокие», как и в других областях, означает, что статистическая технология опирается на современные достижения статистической теории и практики. Это означает, во-первых, что математическая основа технологии получена сравнительно недавно в рамках соответствующей научной дисциплины, во-вторых, что алгоритмы расчетов разработаны и обоснованы в соответствии с нею (а не являются т.н. эвристическими).

Термин «статистические» привычен, он подробно разъясняется на протяжении всего изложения в настоящем учебнике. С нашей точки зрения статистические данные — это результаты измерений, наблюдений, испытаний, анализов, опытов, а «статистические технологии» — это технологии анализа статистических данных.

Наконец, сравнительно редко используемый применительно к статистике термин «технологии». Статистический анализ данных, как правило, включает в себя целый ряд процедур и алгоритмов, выполняемых последовательно, параллельно или по более сложной схеме. В частности, можно выделить следующие этапы:

- планирование статистического исследования;
- организация сбора необходимых статистических данных по оптимальной или рациональной программе (планирование выборки, создание организационной структуры и подбор команды статистиков, подготовка кадров, которые будут заниматься сбором данных, а также контролеров данных и т.п.);
- непосредственный сбор данных и их фиксация на тех или иных носителях (с контролем качества сбора и отбраковкой ошибочных данных по соображениям предметной области);
- первичное описание данных (расчет различных выборочных характеристик, функций распределения, непараметрических оценок плотности, построение гистограмм, корреляционных полей, различных таблиц и диаграмм и т.д.);
- оценивание тех или иных числовых или нечисловых характеристик и параметров распределений (например, непараметрическое интервальное оценивание коэффициента вариации или восстановление зависимости между откликом и факторами, т.е. оценивание функции);
- проверка статистических гипотез (иногда их цепочек — после проверки предыдущей гипотезы принимается решение о проверке той или иной последующей гипотезы);
- более углубленное изучение, т.е. применение различных алгоритмов многомерного статистического анализа, алгоритмов диагностики и построения классификации, статистики нечисловых и интервальных данных, анализа временных рядов и др.;
- проверка устойчивости полученных оценок и выводов относительно допустимых отклонений исходных данных и предпосылок используемых вероятностно-статистических моделей, в частности, изучение свойств оценок методом размножения выборок;
- применение полученных статистических результатов в прикладных целях (например, для диагностики конкретных материалов, построения прогнозов, выбора инвестиционного проекта из предложенных вариантов, нахождения

оптимальных режима осуществления технологического процесса, подведения итогов испытаний образцов технических устройств и др.);

- составление итоговых отчетов, в частности, предназначенных для тех, кто не является специалистами в статистических методах анализа данных, в том числе для руководства — «лиц, принимающих решения».

Возможны и иные структуризации статистических технологий. Важно подчеркнуть, что квалифицированное и результативное применение статистических методов — это отнюдь не проверка одной отдельно взятой статистической гипотезы или оценка параметров одного заданного распределения из фиксированного семейства. Подобного рода операции — только отдельные кирпичики, из которых складывается статистическая технология.

Процедура статистического анализа данных — это информационный технологический процесс, другими словами, та или иная информационная технология. Статистическая информация подвергается разнообразным операциям (последовательно, параллельно или по более сложным схемам). В настоящее время об автоматизации всего процесса статистического анализа данных говорить было бы несерьезно, поскольку имеется слишком много нерешенных проблем, вызывающих дискуссии среди статистиков.

Программное обеспечение статистических методов. В настоящее время статистическая обработка данных проводится, как правило, с помощью соответствующих программных продуктов. Мы не сочли целесообразным включать в учебник ссылки на те или иные пакеты программ по нескольким причинам.

Во-первых, популяции программных продуктов быстро обновляются. Пакеты программ, разработанные 10–15 лет назад, безнадежно устарели. Новые версии, как правило, весьма отличаются от предшественников десятилетней давности. В то же время лучшие книги 1940–1960-х годов по статистическим методам остаются актуальными и сейчас. Например, монографии [12–14].

Во-вторых, каждый программный продукт обладает определенными достоинствами и недостатками. Как показывает наш опыт [15], при сравнении нескольких пакетов программ крайне трудно сделать обоснованный вывод о том, какой из них следует предпочесть.

Необходимо отметить, что между математической и прикладной статистикой имеется и с течением времени углубляется разрыв. Он проявляется, в частности, в том, что большинство методов, включенных в статистические пакеты программ (например, в заслуженные *Statgraphics* и *SPSS* или в более новую систему *Statistica*), даже не упоминается в учебниках по математической статистике. В результате разрыва специалист по математической статистике оказывается зачастую беспомощным при обработке реальных данных, а пакеты программ применяют (что еще хуже — и разрабатывают) лица, не имеющие необходимой теоретической подготовки. Естественно, что они допускают разнообразные ошибки. Типовые ошибки при применении критериев согласия Колмогорова и омега-квадрат давно проанализированы в литературе (например, в статье 1985 г. [16] и учебнике [2]). Об удручающих результатах анализа государственных стандартов по статистическим методам управления качеством рассказано в главе 10.

По оценкам экспертов, распространенные статистические пакеты программ обычно соответствуют уровню научных исследований 1960–1970-х годов. В них нет большинства статистических методов, включенных в современные учебники [2, 4]. Впрочем, как показывает практика преподавания, студенты и слушатели легко реализуют новые статистические методы с помощью подручных вычислительных средств.

О перспективах развития статистических методов. Теория статистических методов нацелена на решение реальных задач. Поэтому в ней постоянно возникают новые постановки математических задач анализа статистических данных, развиваются и обосновываются новые методы. Обоснование часто проводится математическими средствами, т.е. путем доказательства теорем. Для удобства читателей сводка по теоретическим инструментам статистических методов представлена в главе 14. Большую роль играет методологическая составляющая — как именно ставить задачи, какие предположения принять с целью дальнейшего математического изучения. Велика роль современных информационных технологий, в частности, компьютерного эксперимента. Перспективам развития статистических методов посвящена заключительная глава 16 настоящего учебника.

Отметим, что актуальной является задача анализа истории статистических методов с целью выявления тенденций развития и применения их для прогнозирования. Попытка наметить пути такого анализа предпринята в главе 15.

Ситуация с внедрением современных статистических методов на предприятиях и в организациях различных отраслей народного хозяйства внушает оптимизм. На отечественных предприятиях продолжают развиваться структуры, нуждающиеся в статистических методах, — подразделения качества, надежности, управления персоналом, центральные заводские лаборатории и другие. Толчок к развитию в последние годы получили службы контроллинга, маркетинга и сбыта, логистики, сертификации, прогнозирования и планирования, инноваций и инвестиций, управления рисками, которым также полезны различные статистические методы, в частности, методы экспертных оценок. Включенные в учебник методы необходимы органам государственного и муниципального управления, организациям силовых ведомств, транспорта и связи, медицины, образования, агропромышленного комплекса, научным и практическим работникам всех областей деятельности [17, 18].

У статистических методов анализа данных — большое будущее!

Литература

1. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

2. *Орлов, А.И.* Прикладная статистика : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 671 с.

3. *Орлов, А.И.* О перестройке статистической науки и ее применений / А.И. Орлов // Вестник статистики. — 1990. — № 1. — С. 65–71.

4. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 3-е изд., испр. и доп. — Москва : Экзамен, 2004. — 576 с.

5. *Орлов, А.И.* Математическое моделирование процессов налогообложения (подходы к проблеме) / А.И. Орлов., В.Г. Кольцов, Н.Ю. Иванова. — Москва : Изд-во ЦЭО Минобразования РФ, 1997. — 232 с.

6. *Орлов, А.И.* Менеджмент в техносфере : учебное пособие / А.И. Орлов, В.Н. Федосеев. — Москва : Академия, 2003. — 384 с.

7. *Орлов, А.И.* Теория принятия решений : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 576 с.

8. *Гнеденко, Б.В.* Математические методы в теории надежности / Б.В. Гнеденко, Ю.К. Беляев, А.Д. Соловьев. — Москва : Наука, 1965. — 524 с.

9. *Гнеденко, Б.В.* Введение в теорию массового обслуживания / Б.В. Гнеденко, И.Н. Коваленко. — М.: Наука, 1966. — 301 с.

10. *Нейлор, Т.* Машинные имитационные эксперименты с моделями экономических систем / Т. Нейлор. — Москва : Мир, 1975. — 500 с.

11. *Орлов, А.И.* Высокие статистические технологии / А.И. Орлов // Заводская лаборатория. — 2003. — Т. 69. — № 11. — С. 55–60.

12. *Крамер, Г.* Математические методы статистики / Г. Крамер. — Москва : Мир, 1948 (1-е изд.), 1975 (2-е изд.). — 648 с.

13. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.).

14. *Смирнов, Н.В.* Курс теории вероятностей и математической статистики для технических приложений / Н.В. Смирнов, И.В. Дунин-Барковский. — 3-е изд., стереотипное. — Москва : Наука, 1969. — 512 с.

15. *Орлов, А.И.* Математическое обеспечение сертификации: сравнительный анализ диалоговых систем по статистическому контролю / А.И. Орлов // Заводская лаборатория. — 1996. — Т. 62. — № 7. — С.46–49.

16. *Орлов, А.И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат / А.И. Орлов // Заводская лаборатория. — 1985. — Т. 51. — № 1. — С.60–62.

17. *Орлов, А.И.* Системная нечеткая интервальная математика : монография / А.И. Орлов, Е.В. Луценко. — Краснодар : Изд-во КубГАУ, 2014. — 600 с.

18. Лойко В.И. Высокие статистические технологии и системно-когнитивное моделирование в экологии : монография / В.И. Лойко, Е.В. Луценко, А.И. Орлов. — Краснодар : КубГАУ, 2019. — 258 с.

19. *Орлов, А.И.* Организационно-экономическое моделирование : учебник : в 3 ч. / А.И. Орлов. — Ч. 3. Статистические методы анализа данных. — Москва : Изд-во МГТУ им. Н.Э. Баумана, 2012. — 624 с.

20. *Орлов, А.И.* Прикладная статистика — состояние и перспективы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. — 2016. — № 119. — С. 44–74.

21. *Орлов, А.И.* Вероятностно-статистические модели данных — основа методов прикладной статистики / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2020. — Т. 86. — № 7. — С. 5–6.

22. *Орлов, А.И.* Точки роста статистических методов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 103. — С. 136–162.

23. *Орлов, А.И.* Смена парадигм в прикладной статистике / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2021. — Т. 87. — № 7. — С. 6–7.

24. *Орлов, А.И.* Параметрические и непараметрические статистические методы / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2018. — Т. 84. — № 7. — С. 5–6.

25. *Орлов, А.И.* Структура непараметрической статистики (обобщающая статья) / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2015. — Т. 81. — № 7. — С. 62–72.

26. *Орлов, А.И.* Статистика нечисловых данных — центральная часть современной прикладной статистики / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 156. — С. 111–142.

27. *Орлов, А.И.* Статистика нечисловых данных за сорок лет (обзор) / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2019. — Т. 85. — № 11. — С. 69–84.

28. *Орлов, А.И.* Компьютерно-статистические методы: состояние и перспективы / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 103. — С. 163–195.

29. *Орлов, А.И.* Метод статистических испытаний — инструмент исследователя / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2016. — Т. 82. — № 7. — С. 5.

30. *Орлов, А.И.* Предельные теоремы и метод Монте-Карло / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2016. — Т. 82. — № 7. — С. 67–72.

31. *Орлов, А.И.* Значение информационно-коммуникационных технологий для математических методов исследования / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2017. — Т. 83. — № 7. — С. 5–6.

32. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в цифровой экономике (на примере управления качеством) / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 169. — С. 216–242.

33. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в организации производства в эпоху цифровой экономики / А.И. Орлов // Инновации в менеджменте. — 2021. — № 2 (28). — С. 36–45.

ЧАСТЬ 1. ОСНОВНЫЕ ПОСТАНОВКИ ЗАДАЧ АНАЛИЗА ДАННЫХ

ГЛАВА 1. ВЫБОРОЧНЫЕ ИССЛЕДОВАНИЯ

Термин «выборочные исследования применяют, когда невозможно изучить все единицы представляющей интерес совокупности. Приходится знакомиться с частью совокупности — с выборкой, а затем с помощью статистических методов и моделей переносить выводы с выборки на всю совокупность. Выборочные исследования — способ получения статистических данных и важная часть прикладной статистики.

1.1. Организация выборочных исследований

В качестве примера рассмотрим выборочные исследования предпочтений потребителей, которые часто проводят специалисты по маркетингу (изучению рынка).

Оценивание функции спроса. Функция спроса часто встречается в учебниках по экономической теории, но при этом обычно не рассказывается, как она получена. Между тем оценить ее по эмпирическим данным не так уж трудно. Например, можно выяснять ожидаемый спрос с помощью следующего простого приема — спрашиваем потенциальных потребителей: «Какую максимальную цену Вы заплатили бы за такой-то товар?» Пусть для определенности выборка состояла из 20 опрошенных. Они назвали следующие максимально допустимые для них цены:

40, 25, 30, 50, 35, 20, 50, 32, 15, 40,
20, 40, 45, 30, 50, 25, 35, 20, 35, 40.

Сначала названные опрошенными величины упорядочим в порядке возрастания. Результаты представлены в табл. 1. В первом столбце — номера различных численных значений (в порядке возрастания), названных потребителями. Во втором столбце приведены сами значения цены, названные ими. В третьем столбце указано, сколько раз названо то или иное значение.

Таблица 1

Эмпирическая оценка функции спроса и ее использование

№ п/п (i)	Цена p_i	Повторы N_i	Спрос $D(p_i)$	Прибыль $(p$-10)D(p)	Прибыль $(p$-15)$D(p)$	Прибыль $(p$-25)$D(p)$
1	15	1	20	100	0	–
2	20	3	19	190	95	–
3	25	2	16	240	160	0
4	30	2	14	280	210	70
5	32	1	12	264	204	84
6	35	3	11	275	220	110
7	40	4	8	240	200	120
8	45	1	4	140	120	80
9	50	3	3	120	105	75

Таким образом, 20 потребителей назвали 9 конкретных значений цены (максимально допустимых, или приемлемых для них значений), каждое из значений, как видно из третьего столбца, названо от 1 до 4 раз. Теперь легко построить выборочную функцию спроса в зависимости от цены. Она будет представлена в четвертом столбце, который заполним снизу вверх. Спрос как функция от цены p обозначен $D(p)$ (от *demand* (англ.) — спрос). Если мы будем предлагать товар по цене свыше 50 руб., то его не купит никто из опрошенных. При цене 50 руб. появляются 3 покупателя. Записываем 3 в четвертый столбец в девятую строку. А если цену понизить до 45? Тогда товар купят четверо — тот единственный, для кого максимально возможная цена — 45, и те трое, кто был согласен на более высокую цену — 50 руб. Таким образом, легко заполнить столбец 4, действуя по правилу: значение в клетке четвертого столбца равно сумме значений в находящейся слева клетке третьего столбца и в лежащей снизу клетке четвертого столбца. Например, за 30 руб. купят товар 14 человек, а за 20 руб. — 19.

Зависимость спроса от цены — это зависимость четвертого столбца от второго. Табл. 1 дает нам девять точек такой зависимости. Зависимость можно представить на рисунке, в координатах «спрос — цена». Если абсцисса — это спрос, а ордината — цена, то девять точек на кривой спроса, перечисленные в порядке возрастания абсциссы, имеют вид:

$$(3; 50), (4; 45), (8; 40), (11; 35), (12; 32),$$
$$(14; 30), (16; 25), (19; 20), (20; 15).$$

Эти девять точек можно использовать для построения кривой спроса каким-либо графическим (сделайте чертеж!) или расчетным способом, например, методом наименьших квадратов (см. ниже главу 6). Кривая спроса, как и следует ожидать согласно учебникам экономической теории, убывает, имея направления от левого верхнего угла чертежа к правому. Однако заметны отклонения от гладкого вида функции, связанные, в частности, с естественным пристрастием потребителей к круглым числам. Заметьте, все опрошенные, кроме одного, назвали числа, кратные 5 руб.

Расчет оптимальной цены. Данные табл. 1 могут быть использованы для выбора цены продавцом-монополистом. Или организацией, действующей на рынке монополистической конкуренции. Пусть расходы на изготовление или оптовую покупку единицы товара равны 10 руб. По какой цене ее продавать на том рынке, функцию спроса для которого мы только что нашли? Для ответа на этот вопрос вычислим суммарную прибыль, т.е. произведение прибыли на одной единице товара ($p - 10$) на число проданных (точнее, запрошенных) экземпляров $D(p)$. Результаты приведены в пятом столбце табл. 1. Видно, что максимальная прибыль, равная 280 руб., достигается при цене 30 руб. за единицу товара. При этом из 20 потенциальных покупателей окажутся в состоянии заплатить за книгу 14, т.е. 70 %.

Если же удельные издержки производства, приходящиеся на одну единицу товара (или оптовая цена), повысятся до 15 руб., то данные столбца 6 табл. 1 показывают, что максимальная прибыль, равная 220 руб. (она, разумеется, меньше, чем в предыдущем случае), достигается при более высокой цене — 35 руб. Эта цена доступна 11 потенциальным покупателям, т.е. 55 % от всех возможных покупателей. При дальнейшем повышении издержек, скажем, до 25 руб., как вытекает из данных столбца 7 табл. 1, максимальная прибыль, равная 120 руб., достигается при цене 40 руб. за единицу товара, что доступно 8 лицам, т.е. 40 % покупателей. Отметьте, что при повышении оптовой цены на 10 руб. оказалось выгодным увеличить розничную лишь на 5, поскольку более резкое повышение привело бы к такому сокращению спроса, которое перекрыло бы эффект от повышения удельной прибыли (т.е. прибыли, приходящейся на одну проданную единицу товара).

Представляет интерес анализ оптимального объема выпуска при различных значениях удельных издержек (табл. 2).

В табл. 2 звездочками указаны максимальные значения прибыли при том или ином значении издержек, не включенном в табл. 1. Для легкости обозрения результаты об оптимальных объемах выпуска и соответствующих ценах из табл. 1 и 2 приведены в табл. 3.

Таблица 2

Прибыль при различных значениях издержек

№ (i)	Цена p_i	Спрос $D(p_i)$	Прибыль (р-5)$D(p)$	Прибыль (p-20)$D(p)$	Прибыль (p-30)$D(p)$	Прибыль (p-35)$D(p)$	Прибыль (p-40)$D(p)$
1	15	20	200	–	–	–	–
2	20	19	285	0	–	–	–
3	25	16	320	80	–	–	–
4	30	14	350*	140	0	–	–
5	32	12	324	144	24	–	–
6	35	11	330	165*	55	0	–
7	40	8	280	160	80*	40	0
8	45	4	160	100	60	40	20
9	50	3	135	90	60	45*	30*

Таблица 3

Зависимость оптимального выпуска и цены от издержек

Издержки	5	10	15	20	25	30	35	40
Оптимальный выпуск	14	14	11	11	8	8	3	3
Цена	30	30	35	35	40	40	50	50

Как видно из табл. 3, с ростом издержек оптимальный выпуск падает, а цена растет. При этом изменение издержек на 5 единиц может вызывать, а может и не вызывать повышения цены. В этом проявляется микроструктура функции спроса — небольшое повышение цены может привести к тому, что значительные группы покупателей откажутся от покупок, и прибыль упадет.

Этот эффект напоминает известное в экономической теории разделение налогового бремени между производителем и потребителем. Неверно говорить, что производитель перекладывает издержки или, конкретно, налоги, на потребителя, повышая цену на их величину, поскольку при этом сокращается спрос (и выпуск), а потому и прибыль производителя.

Дальнейшее ясно — если оптовая цена будет повышаться, то и дающая максимальную прибыль розничная цена также будет повышаться, и все меньшая доля покупателей сможет приобрести товар. Крайняя точка — оптовая цена, равная 45 руб. Тогда только трое (15 %) купят товар за 50 руб., а прибыль продавца составит только 15 руб. Наглядно видно, что повышение издержек производства приводит к ориентации производителя на наиболее богатые слои населения. Но и повышение цен (до оптимального для монополиста-производителя уровня) не приводит к повышению прибыли, напротив, она снижается, и при этом большинство потенциальных потребителей не в состоянии купить товар.

Отметим, что рыночные структуры не в состоянии обеспечить всех желающих — это просто не выгодно. Так, из 20 опрошенных лишь 14, т.е. 70 %, могут рассчитывать на покупку, даже при минимальных издержках и ценах. Если общество желает чем-либо обеспечить всех граждан, оно должно раздавать это благо бесплатно, как это делается, например, с учебниками в вузах[2].

Для изучения предпочтений потребителей часто используют более изощренные методы. Рассмотрим некоторые из них.

Маркетинговые опросы потребителей. Потенциального покупателя интересует не только цена, но и качество товара, красота упаковки (например, для подарочных наборов конфет) и многое другое. Хочешь узнать, чего желает потребитель — спроси его. Эта простая мысль объясняет популярность маркетинговых опросов.

Бесспорно, что основная цель производственной и торговой деятельности — удовлетворение потребностей людей. Как получить представление об этих потребностях? Очевидно, необходимо опросить потребителей. В американском учебнике по рекламному делу [1] подробно рассматриваются различные методы опроса потребителей и обработки результатов с помощью методов эконометрики. Расскажем о результатах опроса потребителей растворимого кофе. Исследование проведено Институтом высоких статистических технологий и эконометрики по заказу АОЗТ «Д-2» в Москве.

Сбор данных. Один из важнейших разделов прикладной статистики — сбор данных. Обсудим постановку задачи в случае опроса потребителей растворимого кофе. Заказчика интересуют предпочтения как продавцов кофе (розничных и мелкооптовых), так и непосредственно потребителей. В результате совместного обсуждения было признано целесообразным использовать для опроса и тех, и других одну и ту же анкету из 14 основных и 4 социально-демографических вопросов с добавлением двух вопросов специально для продавцов. Анкета была разработана совместно представителями заказчика и исполнителя и утверждена заказчиком. В табл. 4 приведен несколько сокращенный вариант этой анкеты.

---

[2] Описанный здесь метод оценивания спроса был разработан в Институте высоких статистических технологий и эконометрики в [7].

Таблица 4

Анкета для потребителей растворимого кофе (в сокращении)

Дорогой потребитель растворимого кофе,

Институт высоких статистических технологий и эконометрики просит Вас ответить на несколько простых вопросов о том, какой кофе Вы любите. Ваши ответы позволят составить объективное представление о вкусах российских любителей кофе и будут способствовать повышению качества этого товара на российском рынке.

1. Часто ли Вы пьете растворимый кофе: иногда, каждый день 1 чашку, 2–3 чашки, больше, чем 3 чашки.

(Здесь и далее подчеркните нужное.)

2. Что Вы цените в кофе: вкус, аромат, крепость, цвет, отсутствие вредных для здоровья веществ, что-либо еще (сообщите нам, что именно) ______________________________.

3. Как часто покупаете кофе: по мере надобности или по возможности?

4. Какую марку растворимого кофе Вы обычно покупаете? ____________

5. Какой объем упаковки Вы предпочитаете: в пакетиках, маленькая банка, средняя банка, большая банка, обязательно стеклянная банка, все равно.

6. Где покупаете растворимый кофе: в ларьках, в продуктовых магазинах, в специализированных отделах и магазинах, все равно, где купить, где-либо еще (опишите, пожалуйста) _______________.

7. Были ли случаи, когда купленный Вами кофе оказывался низкого качества? Да, нет.

8. Согласны ли Вы, что за высокое и гарантированное качество продукта можно и заплатить несколько дороже? Да, нет.

9. На сколько дороже Вы готовы платить за экологически безопасный кофе? ____

10. Считаете ли Вы нужным, чтобы вредные для здоровья вещества, в частности, ионы тяжелых металлов, не проникали из материала упаковки в растворимый кофе? Да, нет.

Мы планируем сравнить потребительские предпочтения различных категорий жителей нашей страны. Поэтому просим ответить еще на несколько вопросов.

11. Пол: женский, мужской.

12. Возраст: до 20, 20–30, 30–50, более 50.

13. Род занятий: учащийся, работающий, пенсионер, инженер, врач, преподаватель, служащий, менеджер, предприниматель, научный работник, рабочий, др. (пожалуйста, расшифруйте).

14. Вся Ваша семья любит растворимый кофе или же Вы — единственный любитель этого восхитительного напитка современного человека? Вся семья, я один (одна).

Спасибо за Ваше содействие работе по повышению качества продуктов на российском рынке!

Выбор метода опроса. Широко применяются процедуры опроса, когда респонденты (так социологи и маркетологи называют тех, от кого получают информацию, т.е. опрашиваемых) самостоятельно заполняют анкеты (розданные им или полученные по почте), а также личные и телефонные интервью. Из этих процедур нами было выбрано личное интервью по следующим причинам.

Возврат почтовых анкет сравнительно невелик (в данном случае можно было ожидать не более 5–10 %), оттянут по времени и искажает структуру совокупности потребителей (наиболее динамичные люди вряд ли найдут время для ответа на подобную анкету).

Самостоятельное заполнение анкеты, как показали специально проведенные эксперименты, не позволяет получить полные ответы на поставленные вопросы. Респондент утомляется или отвлекается, отказывается отвечать на часть вопросов, иногда не понимает их или отвечает не по существу. Некоторые категории респондентов, например, продавцы в киосках, отказываются заполнять анкеты, но готовы устно ответить на вопросы.

Телефонный опрос искажает совокупность потребителей, поскольку наиболее активных индивидуумов трудно застать дома и уговорить ответить на вопросы анкеты. Репрезентативность нарушается также и потому, что на один номер телефона может приходиться различное количество продавцов и потребителей растворимого кофе, а некоторые из них не имеют телефонов вообще. Анкета достаточно длинная, и разговор по домашнему и служебному телефону респондента может быть прекращен досрочно по его инициативе. Иногородних продавцов и потребителей растворимого кофе, приехавших в Москву, по телефону опросить практически невозможно.

Метод личного интервью лишен перечисленных недостатков. Соответствующим образом подготовленный интервьюер, получив согласие на интервью, удерживает внимание собеседника на анкете, добивается получения ответов на все ее вопросы, контролируя при этом соответствие ответов реальной позиции респондента. Ясно, что успех интервьюирования зависит от личных качеств и подготовки интервьюера. Однако расходы на получение одной анкеты при использовании этого метода больше, чем для других рассмотренных методов.

Формулировки вопросов. В маркетинговых и социологических опросах используют три типа вопросов — закрытые, открытые и полузакрытые, они же полуоткрытые. При ответе на закрытые вопросы респондент может выбирать лишь из сформулированных составителями анкеты вариантов ответа. В качестве ответа на открытые вопросы респондента просят изложить свое мнение в свободной форме. Полузакрытые, они же полуоткрытые вопросы занимают промежуточное положение — кроме перечисленных в анкете вариантов, респондент может добавить свои соображения.

В социологических публикациях, посвященных выборочным исследованиям, продолжается дискуссия по поводу «мягких» и «жестких» форм сбора данных, т.е. фактически о том, какого типа вопросы более целесообразно использовать — открытые или закрытые (см., например, статью известного социолога В.А. Ядова [2]).

Преимущество открытых вопросов состоит в том, что респондент может свободно высказать свое мнение так, как сочтет нужным. Их недостаток — в сложности сопоставления мнений различных респондентов. Для такого сопоставления и получения сводных характеристик организаторы опроса вынуждены сами шифровать ответы на открытые вопросы, применяя разработанную ими схему шифровки.

Преимущество закрытых вопросов в том и состоит, что такую шифровку проводит сам респондент. Однако при этом организаторы опроса уподобляются древнегреческому мифическому персонажу Прокрусту. Как известно, Прокруст приглашал путников заночевать у него. Укладывал их на кровать. Если путник был маленького роста, он вытягивал его ноги так, чтобы они доставали до конца кровати. Если же путник оказывался высоким и ноги его торчали — он обрубал их так, чтобы достигнуть стандарта: «рост» путника должен равняться длине кровати. Так и организаторы опроса, применяя закрытые вопросы, заставляют респондента «вытягивать» или «обрубать» свое мнение, чтобы выразить его с помощью приведенных в формулировке вопроса возможных ответов.

Ясно, что для обработки данных по группам и сравнения групп между собой нужны формализованные данные, и фактически речь может идти лишь о том, кто — респондент или маркетолог (социолог, психолог и др.) — будет шифровать ответы. В проекте «Потребители растворимого кофе» практически для всех вопросов варианты ответов можно перечислить заранее, т.е. можно широко использовать закрытые вопросы. В отличие от опросов с вопросами типа: «Одобряете ли Вы идущие в России реформы?», в которых естественно просить респондента расшифровать, что он понимает под «реформами» (открытый вопрос). Поэтому в используемой в описываемом проекте анкете использовались в основном закрытые и полузакрытые вопросы. Как показали результаты обработки, этот подход оказался правильным — лишь в небольшом числе анкет оказались вписаны свои варианты ответов. Вместе с тем демонстрировалось уважение к мнению респондента, не выдвигалось требование обязательного выбора из заданного множества ответов — респондент мог добавить свое, но редко пользовался этой возможностью (не более чем в 5 % случаев).

В последнем вопросе анкеты респонденту предлагалось стать постоянным участником опросов о качестве товаров народного потребления. Ряд респондентов откликнулся на это предложение, в результате стало возможным развертывание постоянной сети «экспертов по качеству», подобной аналогичным в США и других странах.

1.2. Модели случайных выборок

Статистические методы выборочных исследований основаны на вероятностных моделях, описывающих получение ответов опрашиваемых на вопросы анкет. В случае ответов типа «да» — «нет» наиболее распространенными являются две вероятностные модели — биномиальная и гипергеометрическая.

В биномиальной модели предполагается, что ответы n опрашиваемых можно рассматривать как совокупность n независимых одинаково распределенных случайных величин $X_1, X_2,...,X_n$, где $X_i = 1$, если i-й респондент сказал «да», и $X_i = 0$, если его ответ — «нет». Тогда число X ответов «да» в выборке равно:

$$X = X_1 + X_2 + ... + X_n\ . \qquad (1)$$

Из формулы (1) и Центральной Предельной Теоремы теории вероятностей (см. главу 14) вытекает, что при увеличении объема выборки n распределение X сближается с нормальным распределением. Известно, что распределение X имеет вид:

$$P(X = k) = C_n^k\, p^k\,(1—p)^{n-k}, \qquad (2)$$

где C_n^k — число сочетаний из n элементов по k, а p — доля ответов «да» в генеральной совокупности, т.е. $p = P(X_i = 1)$. Формула (2) задает биномиальное распределение, часто используемое при вероятностном моделировании реальных явлений и процессов.

Гипергеометрическое распределение соответствует иной схеме — случайному отбору респондентов в выборку. Пусть среди N лиц, составляющих генеральную совокупность, имеется D лиц, чье мнение — «да». Случайность отбора респондентов в выборку означает, что каждое лицо имеет одинаковые шансы быть отобранным. Мало того, ни одна пара потенциальных респондентов не должна иметь при отборе в выборку преимущества перед любой другой парой. То же самое — для троек, четверок и т.д. Это условие выполнено тогда и только тогда, когда каждое из C_N^n сочетаний по n лиц из N имеет одинаковые шансы быть отобранным в качестве выборки. Вероятность того, что будет отобрано заранее заданное сочетание, равна, очевидно, $1/C_N^n$.

Пусть Y —число сказавших «да» лиц в случайной выборке, организованной таким образом. Известно, что тогда $P(Y = k)$ — гипергеометрическое распределение, т.е.:

$$P(Y = k) = \frac{C_D^k C_{N-D}^{n-k}}{C_N^n}. \qquad (3)$$

Отбор случайной выборки, согласно описанным правилам, организуют при проведении различных лотерей. Например, отбирают 6 номеров из 49. Тогда генеральная совокупность состоит из 49 единиц (номеров), а выборка — из 6. В этом случае отбирают номера, а не респондентов, но вероятностная модель — та же. Удобно говорить, что генеральная совокупность и выборка состоят из единиц. В одном случае единицы — это люди (лица, потенциальные респонденты), в другом — номера. В главе 10 рассматриваются единицы продукции — детали или изделия.

Замечательный математический факт состоит в том, что биномиальная и гипергеометрическая модели *весьма близки* (с практической точки зрения совпадают), когда объем генеральной совокупности (партии) по крайней мере в 10 раз превышает объем выборки. Другими словами, можно принять, что:

$$P(X = k) = P(Y = k), \qquad (4)$$

если объем выборки мал по сравнению с объемом партии. При этом в качестве p в формуле (4) берут D/N.

Близость результатов, получаемых с помощью биномиальной и гипергеометрической моделей, весьма важна не только с практической, но и с методологической точки зрения. Дело в том, что эти модели исходят из принципиально различных методологических предпосылок. В биномиальной модели случайность *присуща каждому респонденту*. Он с какой-то вероятностью отвечает «да», а с какой-то — «нет» (сумма этих вероятностей, очевидно, равна 1). В то же время в гипергеометрической модели ответ респондента полностью определен, а случайность проявляется лишь *в отборе*, вносится социологом или маркетологом при составлении выборки.

В науках о человеке противоречие между рассматриваемыми моделями выборки четко выражено. В среде специалистов, изучающих человека (маркетологов, социологов, психологов, политологов и др.) давно идет дискуссия о роли случайности в поведении человека. А именно, о том, есть ли случайность в поведении отдельно взятого человека или же случайность проявляется лишь в отборе выборки из генеральной совокупности.

Биномиальная модель предполагает, что поведение человека, в частности, выбор им определенного варианта при ответе на вопрос, определяется с участием случайных причин. Например, человек может случайно сказать «да», случайно — «нет». Некоторые философы отрицают случайность, присущую поведению человека согласно биномиальной модели. Они верят в причинность и считают поведение конкретного человека практически полностью детерминированным (его взглядами, психофизиологическими особенностями, прежним опытом и др.). Поэтому они принимают гипергеометрическую модель и считают, что случайность отличия ответов в выборке от ответов во всей генеральной совокупности определяется всецело случайностью, вносимой при отборе единиц наблюдения в выборку.

Сформулированные выше математические результаты (соотношение (4)) показывают, что позиция в этой давней дискуссии практически не влияет на алгоритмы обработки данных. Следовательно, во многих случаях нет необходимости принимать чью-либо сторону в этом споре, поскольку обе модели дают близкие численные результаты.

Отличия проявляются лишь при обсуждении вопроса о том, какую выборку считать *представительной*. В терминах контроля качества (глава 10) — является ли таковой выборка, составленная из 20 изделий, лежащих сверху в первом вскрытом ящике? В биномиальной модели вполне допустим ответ «да», в гипергеометрической — только «нет».

Биномиальная модель легче для теоретического изучения, поэтому будем ее рассматривать в дальнейшем. Однако при реальном опросе лучше формировать выборку, исходя из гипергеометрической модели. Это делают, выбирая респондентов из списка избирателей (для включения в выборку) с помощью датчиков псевдослучайных чисел на ЭВМ или с помощью таблиц псевдослучайных чисел. Алгоритмы формирования выборки встраивают во все современные программные продукты, предназначенные для поддержки проведения маркетинговых или социологических опросов, организации статистического контроля качества и др.

Обоснование объема выборки и проведение опроса. Вернемся к анализу результатов опроса потребителей растворимого кофе, о котором шла речь в предыдущем разделе. Как уже говорилось, модели выборочных исследований часто опираются на предположение о том, что реальную выборку можно описывать как «случайную выборку из конечной совокупности». Типа той, когда из списков избирателей с помощью датчика случайных чисел отбирается необходимое число номеров для формирования жюри присяжных заседателей. В рассматриваемом исследовании нельзя обеспечить формирование подобной выборки — не существует реестра потребителей растворимого кофе. Однако в этом и нет необходимости. Поскольку гипергеометрическое распределение хорошо приближается биномиальным, если объем выборки по крайней мере в 10 раз меньше объема всей совокупности (в рассматриваемом случае это так), то правомерно использование биномиальной модели, согласно которой мнение респондента (ответы на все вопросы анкеты) рассматривается как случайный вектор, а все такие вектора независимы между собой. Другими словами, можно использовать модель простой случайной выборки.

1.3. Доверительное оценивание доли

Зачем проводятся выборочные исследования? Чтобы получить необходимую информацию о генеральной совокупности. Для этого необходимо перенести выводы с выборки на генеральную совокупность. Как и с какой точностью можно это сделать?

Рассмотрим эту проблему для простейшего случая одного вопроса с двумя возможными ответами — «да» и «нет».

Напомним, что биномиальная модель выборки как раз и применяется для описания ответов на закрытые вопросы, имеющие две подсказки, например, «да» и «нет». Конечно, пары подсказок могут быть иными. Например, «согласен» и «не согласен». Или при опросе потребителей кондитерских товаров первая подсказка может иметь такой вид: «Больше люблю «Марс», чем «Сникерс». А вторая тогда такова: «Больше люблю «Сникерс», чем «Марс».

Пусть объем выборки равен n. Тогда ответы опрашиваемых можно представить как $X_1, X_2, \ldots, X_n$, где $X_i = 1$, если i-й респондент выбрал первую подсказку, и $X_i = 0$, если i-й респондент выбрал вторую подсказку, $i = 1, 2, \ldots, n$. В вероятностной модели предполагается, что случайные величины $X_1, X_2, \ldots, X_n$ независимы и одинаково распределены. Поскольку эти случайные величины принимают два значения, то ситуация описывается одним параметром p — долей выбирающих первую подсказку во всей генеральной совокупности. Тогда:

$$P(X_i = 1) = p,\ P(X_i = 0) = 1 - p,\ i = 1, 2, \ldots, n.$$

Пусть $m = X_1 + X_2 + \ldots + X_n$. Оценкой вероятности p является частота $p^* = m/n$. При этом математическое ожидание $M(p^*)$ и дисперсия $D(p^*)$ имеют вид:

$$M(p^*) = p,\ D(p^*) = p(1-p)/n.$$

По Закону Больших Чисел (ЗБЧ) теории вероятностей (в данном случае — по теореме Бернулли) частота p^* сходится (т.е. безгранично приближается) к вероятности p при росте объема выборки (см. главу 14). Это означает, что оценивание проводится тем точнее, чем больше объем выборки. Точность оценивания можно указать. Займемся этим.

По теореме Муавра — Лапласа теории вероятностей:

$$\lim_{n\to\infty} P\{\frac{m - np}{\sqrt{np(1-p)}} \le x\} = \Phi(x),$$

где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy,$$

где π= 3,1415925… — отношение длины окружности к ее диаметру, е = 2,718281828… — основание натуральных логарифмов. График плотности стандартного нормального распределения:

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}},$$

был очень точно изображен на германской денежной банкноте в 10 немецких марок (до введения евро). Банкнота была посвящена великому немецкому математику Карлу Гауссу (1777–1855), среди основных работ которого есть относящиеся к нормальному распределению. Эта подробность демонстрирует, что в Германии (и тем более в англосаксонских странах) гораздо шире распространено знакомство с основами теории вероятностей и математической статистики, чем в нашей стране.

В настоящее время нет необходимости вычислять функцию стандартного нормального распределения и ее плотность по приведенным выше формулам, поскольку давно составлены подробные таблицы (см., например, [3]), а распространенные программные продукты содержат алгоритмы нахождения этих функций.

С помощью теоремы Муавра — Лапласа могут быть построены доверительные интервалы для неизвестной статистику вероятности. Сначала заметим, что из этой теоремы непосредственно следует, что:

$$\lim_{n\to\infty}P\left\{-x\le\frac{m-np}{\sqrt{np(1-p)}}\le x\right\}=\Phi(x)-\Phi(-x).$$

Поскольку функция стандартного нормального распределения симметрична относительно 0, т.е. $\Phi(x)+\Phi(-x)=1$, то справедливо полезное равенство $\Phi(x)-\Phi(-x)=2\Phi(x)-1$.

Зададим характеристику надежности переноса выводов с выборки на генеральную совокупность — доверительную вероятность γ, близкую к 1. Пусть функция $U(\gamma)$ удовлетворяет условию:

$$\Phi(U(\gamma))-\Phi(-U(\gamma))=\gamma,$$

т.е.

$$U(\gamma)=\Phi^{-1}\left(\frac{1+\gamma}{2}\right).$$

Из последнего предельного соотношения следует, что:

$$\lim_{n\to\infty} P\left\{ p^* - U(\gamma)\frac{\sqrt{p(1-p)}}{\sqrt{n}} \le p \le p^* + U(\gamma)\frac{\sqrt{p(1-p)}}{\sqrt{n}} \right\} = \gamma.$$

К сожалению, это соотношение нельзя непосредственно использовать для доверительного оценивания, поскольку верхняя и нижняя границы зависят от неизвестной вероятности. Однако с помощью метода наследования сходимости (см. главу 14 или [4, п. 2.4]) можно доказать, что:

$$\lim_{n\to\infty} P\{ p^* - U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}} \le p \le p^* + U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}} \} = \gamma.$$

Следовательно, нижняя доверительная граница имеет вид:

$$p_{нижн} = p^* - U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}},$$

в то время как верхняя доверительная граница такова:

$$p_{верх} = p^* + U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}}.$$

Наиболее распространенным (в прикладных исследованиях) значением доверительной вероятности является $\gamma = 0{,}95$. Иногда употребляют термин «95 % доверительный интервал». Тогда $U(\gamma) = 1{,}96$.

Пример 1. Пусть n=500, m=200. Тогда p^* = 0,40. Найдем доверительный интервал для $\gamma = 0{,}95$:

$$p_{нижн} = 0{,}40 - 1{,}96\frac{\sqrt{0{,}4\times 0{,}6}}{\sqrt{500}} = 0{,}40 - 0{,}043 = 0{,}357, \quad p_{верх} = 0{,}40 + 0{,}043 = 0{,}443.$$

Таким образом, хотя в достаточно большой выборке 40 % респондентов говорят «да», можно утверждать лишь, что во всей генеральной совокупности таких от 35,7 до 44,3 % — крайние значения отличаются на 8,6 %.

Замечание. С достаточной для практики точностью можно заменить 1,96 на 2.

Удобные для использования в практической работе специалиста по выборочным исследованиям, маркетолога и социолога таблицы точности оценивания разработаны во ВЦИОМ (Всероссийском центре по изучению общественного мнения). Приведем здесь несколько модифицированный вариант одной из них.

Таблица 5

Допустимая величина ошибки выборки (в процентах)

Объем группы / Доля p*	1000	750	600	400	200	100
Около 10 % или 90 %	2	3	3	4	5	7
Около 20 % или 80 %	3	4	4	5	7	9
Около 30 % или 70 %	4	4	4	6	9	10
Около 40 % или 60 %	4	4	5	6	8	11
Около 50 %	4	4	5	6	8	11

В условиях рассмотренного выше примера надо взять вторую снизу строку. Объема выборки 500 нет в таблице, но есть объемы 400 и 600, которым соответствуют ошибки в 6 и 5 % соответственно. Следовательно, в условиях примера целесообразно оценить ошибку как ((5+6)/2) % = 5,5 %. Эта величина несколько больше, чем рассчитанная выше (4,3 %). С чем связано это различие? Дело в том, что таблица ВЦИОМ связана не с доверительной вероятностью $\gamma = 0{,}95$, а с доверительной вероятностью $\gamma = 0{,}99$, которой соответствует множитель $U(\gamma) = 2{,}58$. Расчет ошибки по приведенным выше формулам дает 5,65 %, что практически совпадает со значением, найденным по табл. 5.

Необходимый объем выборки. В биномиальной модели выборки оценивание характеристик происходит тем точнее, чем объем выборки больше. Часто спрашивают: «Какой объем выборки нужен?» Разработан ряд методов определения необходимого объема выборки. Они основаны на разных подходах. Либо на задании необходимой точности оценивания параметров. Либо на явной формулировке альтернативных гипотез, между которыми необходимо сделать выбор. Либо на учете погрешностей измерений (методы статистики интервальных данных). Ни один из этих подходов нельзя применить в рассматриваемом случае.

Минимальный из обычно используемых объемов выборки n в маркетинговых или социологических исследованиях — 100, максимальный — до 5 000 (обычно в исследованиях, охватывающих ряд регионов страны, т.е. фактически разбивающихся на ряд отдельных исследований — как в ряде исследований ВЦИОМ). По данным Института социологии Российской академии наук [5], среднее число анкет в социологическом исследовании не превышает 700. Поскольку стоимость исследования растет, по крайней мере, как линейная функция объема выборки, а точность повышается как квадратный корень из этого объема, то верхняя граница объема выборки определяется обычно из экономических соображений. Объемы пилотных исследований (т.е. проводящихся впервые, предварительно или как первые в сериях подобных) обычно ниже, чем объемы исследований по обкатанной программе.

Нижняя граница определяется тем, что в минимальной по численности анализируемой подгруппе должно быть несколько десятков человек (не менее 30), поскольку по ответам попавших в эту подгруппу необходимо сделать обоснованные заключения, например, о предпочтениях соответствующей подгруппы в совокупности всех потребителей растворимого кофе. Учитывая деление опрашиваемых на продавцов и покупателей, на мужчин и женщин, на четыре градации по возрасту и восемь — по роду занятий, наличие 5–6 подсказок во многих вопросах, приходим к выводу о том, что в рассматриваемом проекте объем выборки должен быть не менее 400–500. Вместе с тем существенное превышение этого объема было признано нецелесообразным, поскольку исследование являлось пилотным.

Поэтому в проекте «Потребители растворимого кофе» объем выборки был выбран равным 500. Анализ полученных результатов позволяет утверждать, что в соответствии с целями исследования выборку следует считать репрезентативной.

1.4. Два прикладных выборочных исследования

Продолжим обсуждение выборочного исследования потребителей растворимого кофе.

Организация опроса. Интервьюерами работали молодые люди — студенты первого курса экономико-математического факультета Московского государственного института электроники и математики (технического университета) и лицея № 1 140, проходившие обучение по экономике, всего 40 человек, имеющих специальную подготовку по изучению рынка и проведению маркетинговых опросов потребителей и продавцов (в объеме 8 часов). Опрос продавцов проводился на рынках г. Москвы, действующих в Лужниках, у Киевского вокзала и в других местах. Опрос покупателей проводился на рынках, в магазинах, на улицах около киосков и ларьков, а также в домашней и служебной обстановке.

Большое внимание уделялось качеству заполнения анкет. Интервьюеры были разбиты на шесть бригад, бригадиры персонально отвечали за качество заполнения анкет. Второй уровень контроля осуществляла специально созданная «группа организации опроса», третий происходил при вводе информации в базу данных. Каждая анкета заверена подписями интервьюера и бригадира, на ней указано место и время интервьюирования. Поэтому необходимо признать высокую достоверность собранных анкет.

Обработка данных. В соответствии с целью исследования основной метод первичной обработки данных — построение частотных таблиц для ответов на отдельные вопросы. Кроме того, проводилось сравнение различных групп потребителей и продавцов, выделенных по социально-демографическим данным, с помощью критериев проверки однородности выборок (см. ниже). При более углубленном анализе применялись различные методы статистики объектов нечисловой природы (более 90 % маркетинговых и социологических данных имеют нечисловую природу [6]). Использовались средства графического представления данных.

Итоги опроса. Итак, по заданию одной из торговых фирм были изучены предпочтения покупателей и мелкооптовых продавцов растворимого кофе. Совместно с представителями заказчика был составлен опросный лист (анкета типа социологической) из 16 основных вопросов и 4 дополнительных, посвященных социально-демографической информации. Опрос проводился в форме интервью с 500 покупателями и продавцами кофе. Места опроса — рынки, лотки, киоски, продуктовые и специализированные магазины. Другими словами, были охвачены все виды мест продаж кофе. Интервью проводили более 40 специально подготовленных (примерно по 8-часовой программе) студентов, разбитых на 7 бригад. После тщательной проверки бригадирами и группой обработки информация была введена в специально созданную базу данных. Затем проводилась разнообразная статистическая обработка, строились таблицы и диаграммы, проверялись статистические гипотезы и т.д. Заключительный этап — осмысление и интерпретация данных, подготовка итогового отчета и предложений для заказчиков.

Технология организации и проведения маркетинговых опросов лишь незначительно отличается от технологии социологических опросов, многократно описанной в литературе. Так, мы предпочли использовать полуоткрытые вопросы, в которых для опрашиваемого дан перечень подсказок, а при желании он может высказать свое мнение в свободной форме. Не уложившихся в подсказки оказалось около 5 %, их мнения были внесены в базу данных и анализировались дополнительно. Для повышения надежности опроса о наиболее важных с точки зрения маркетинга моментах спрашивалось в нескольких вопросах. Были вопросы — ловушки, с помощью которых контролировалась «осмысленность» заполнения анкеты. Например, в вопросе: «Что Вы цените в кофе: вкус, аромат, крепость, наличие пенки...» ловушкой является включение «крепости» — ясно, что крепость зависит не от кофе самого по себе, а от его количества в чашке. В ловушку никто из 500 не попался — никто не отметил «крепость». Этот факт свидетельствует о надежности выводов проведенного опроса. Мы считали нецелесообразным задавать вопрос об уровне доходов (поскольку в большинстве случаев отвечают «средний», что невозможно связать с определенной величиной). Вместо такого вопроса мы спрашивали: «Как часто Вы покупаете кофе: по мере надобности или по возможности?». Поскольку кофе не является дефицитным товаром, первый ответ свидетельствовал о наличии достаточных денежных средств, второй — об их ограниченности (потребитель не всегда имел возможность позволить себе купить банку растворимого кофе).

Стоимость подобных исследований — 5–10 долларов США на одного обследованного. При этом трудоемкость (и стоимость) начальной стадии — подготовки анкеты и интервьюеров, пробный опрос и др. — 30 % от стоимости исследования. Стоимость непосредственно опроса — тоже 30 %, ввод информации в компьютер и проведение расчетов, построение таблиц и графиков — 20 %, интерпретация результатов, подготовка итогового отчета и предложений для заказчиков — 20 %. Таким образом, стоимость собственно опроса в два с лишним раза меньше стоимости остальных стадий исследования. И в выполнении работы участвуют различные специалисты. На первой стадии — в основном нужны высококвалифицированные аналитики. На второй — многочисленные интервьюеры, в роли которых могут выступать студенты и школьники, прошедшие конкретный курс обучения в 8–10 часов. На третьей — работа с компьютером (надо уметь строить и обсчитывать электронные таблицы или базы данных, использовать статистические пакеты, составлять и печатать таблицы и диаграммы и т.п.). На четвертой — опять в основном нужны высококвалифицированные аналитики.

Приведем некоторые из полученных результатов:

а) в отличие от западных потребителей, отечественные не отдавали предпочтения стеклянным банкам по сравнению с жестяными. Поскольку жестяные банки дешевле стеклянных, то можно было порекомендовать (в 1994 г., когда проходил опрос) с целью снижения расходов закупку кофе в жестяных банках;

б) отечественные потребители готовы платить на 10–20 % больше за экологически безопасный кофе более высокого качества, имеющий сертификат Минздрава и символ экологической безопасности на упаковке;

в) средний объем потребления растворимого кофе одной семьей — 850 г в месяц;

г) потребители растворимого кофе могут быть разделены на классы (кластеры в терминологии главы 6). Есть «продвинутые» потребители, обращающие большое внимание на качество и экологическую безопасность, марку и страну производства, терпимо относящиеся к изменению цены. Эти «тонкие ценители» — в основном женщины от 30 до 50 лет, служащие, менеджеры, научные работники, преподаватели, врачи (т.е. лица с высшим образованием), пьющие кофе как дома, так и на работе, причем «кофейный ритуал» зачастую входит в процедуру деловых переговоров или совещаний. Противоположный по потребительскому поведению класс состоит из мужчин двух крайних возрастных групп — школьников и пенсионеров. Для них важна только цена, что очевидным образом объясняется недостатком денег.

Результаты были использованы заказчиком в рекламной кампании. В частности, обращалось внимание на сертификат Минздрава и на экологическую безопасность упаковки.

Оценивание функции спроса и моделирование рынка. Выпускник программы «Топ-менеджер» А.А. Пивень оценил функцию спроса на продукцию своего предприятия. Расчет и установление оптимальной цены на изделие с точки зрения максимизации прибыли был произведен по описанному выше методу. В табл. 6 приведена функция ожидаемого спроса в зависимости от цены. Как подсчитал А.А. Пивень, уровень издержек на производство 1 изделия составляет 42 824,7 руб. (1 350 у.е.). Для удобства все расчеты будем производить в условных единицах.

Таблица 6

Функция ожидаемого спроса в зависимости от цены

№ п/п	Цена, у.е.	Объем продаж в год, шт.	Издержки на объем производства	Выручка, у.е.	Прибыль, у.е.
1	1 400	1 600	2 160 000	2 240 000	80 000
2	1 500	1 500	2 025 000	2 250 000	225 000
3	1 600	1 200	1 620 000	1 920 000	300 000
4	1 700	1 000	1 350 000	1 700 000	350 000
5	1 800	720	972 000	1 246 000	324 000
6	1 900	500	675 000	950 000	275 000
7	2 000	320	432 000	640 000	208 000
8	2 100	170	229 500	357 000	127 500
9	2 200	110	148 500	242 000	93 500

Как видно из приведенных расчетов, оптимальная цена на подъемник должна находиться в диапазоне 1 600–1 700 у.е.

На основе многомерной регрессионной зависимости (см. главу 6) была построена математическая модель рынка. Она довольно точно отражает реальное положение дел. При исходной цене 1 650 у.е. продажи ориентировочно должны составить 1 010 шт. На рис. 1 приведена кривая спроса.

Эти расчеты были сделаны при допущении, что издержки не меняются в течение длительного промежутка времени. Однако, в реальных условиях постоянный рост стоимости энергоресурсов и непрекращающаяся инфляция издержек (рост затрат на сырье, материалы, комплектующие изделия, рабочую силу) приводит к увеличению издержек. Поэтому А.А. Пивень проанализировал оптимальный объем выпуска при их различных значениях. Данные приведены в табл. 7.

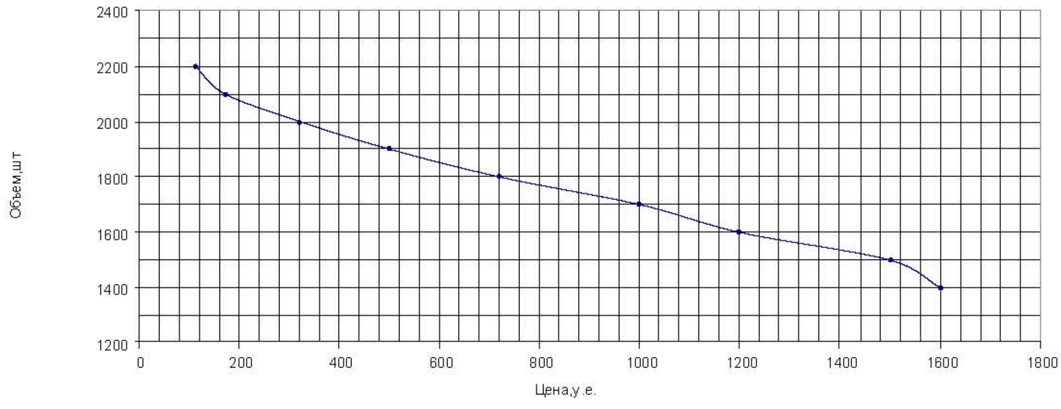


Рис. 1. Кривая спроса на изделие

Таблица 7

Прибыль в зависимости от цены и издержек

№ п/п	Цена, у.е.	Объем продаж, шт.	Прибыль (тыс. у.е.) при издержках на единицу продукции, у.е.							
			1350	**1400**	**1450**	**1500**	**1550**	**1600**	**1650**	**1700**
1	1 400	1 600	80	0	–	–	–	–	–	–
2	1 500	1 500	225	150	75	0	–	–	–	–
3	1 600	1 200	300	240	180	120	60	0	–	–
4	1 700	1 000	350	300	250	200	150	100	50	0
5	1 800	720	324	288	252	216	180	144	108	72
6	1 900	500	275	250	225	200	175	150	125	100
7	2 000	320	208	192	176	160	144	128	112	96
8	2 100	170	127,5	119	110,5	102	93,5	85	76,5	68
9	2 200	110	93,5	88	82,5	77	71,5	66	60,5	55

Для удобства рассмотренные результаты оптимальных объемов производства при соответствующих ценах приведены в табл. 8.

Таблица 8

Оптимальные выпуск и цена в зависимости от издержек

Издержки	1 350	1 400	1 450	1 500	1 550	1 600	1 650	1 700
Оптимальный выпуск	1 000	1 000	720	720	720	500	500	500
Цена	1 700	1 700	1 800	1 800	1 800	1 900	1 900	1 900

Как видно из табл. 8, увеличение издержек ведет к снижению оптимального выпуска при росте цены. Хотя изменение издержек на 50 у.е. может не сразу привести к изменению цены. Необоснованная цена может «переключить» большую группу потребителей на другое, аналогичное изделие, имеющее сходный по уровню набор технических характеристик, но более низкую рыночную цену.

По данным функции спроса (табл. 7) проведем расчет эластичности спроса по цене. Под ценовой эластичностью спроса понимается степень реагирования рыночного спроса на изменение цен. В классическом понимании эластичность спроса по цене показывает, насколько изменится объем спроса при изменении цены на 1 %. Спрос квалифицируется как эластичный, если понижение цены вызывает такой рост оборота, при котором увеличение объема продаж с лихвой компенсирует более низкие цены. Если же понижение цены, приводя к некоторому увеличению объема продаж, тем не менее, не ведет к увеличению оборота или даже уменьшает его, то такой спрос называется неэластичным. Коэффициент ценовой эластичности спроса определяется по формуле:

$$K_{ЦЭС} = \frac{(Q_1 - Q_2)/(Q_1 + Q_2)}{(P_1 - P_2)/(P_1 + P_2)},$$

где Q_1, Q_2 — значения объема продаж; P_1, P_2 — значения цены изделия.

В рассматриваемом случае $K_{ЦЭС}$ будет различен на протяжении всей функции спроса (рис.1). Однако, произведем расчет на той части кривой (в том диапазоне), где присутствует расчетная цена подъемника, а именно: Q_1 = 1200 шт. ; Q_2 = 720 шт.; P_1 = 1 600 у.е.; P_2 = 1 800 у.е. В этом случае:

$$K_{ЦЭС} = \frac{(1200 - 720)/(1200 + 720)}{(1600 - 1800)/(1600 + 1800)} = -4{,}25.$$

Коэффициент $K_{ЦЭС}$ имеет отрицательный знак и абсолютную величину, значительно превышающую 1. Это говорит о сильной обратной зависимости объемов продаж от цены. Спрос на подъемник эластичен. Валовая выручка увеличивается при снижении цены и уменьшается при ее повышении. Компании необходимо быть готовой к тому, что покупатели очень чутко реагируют на всякое повышение цены на изделие значительным снижением объемов закупок. Как отмечает А.А. Пивень, снижение же эластичности спроса на изделие возможно только при общем росте благосостояния населения страны и в частности, значительного роста доходной части бюджетов промышленных предприятий.

1.5. Проверка однородности двух биномиальных выборок

Проверка однородности — одна из базовых проблем, решаемых статистическими методами. Она часто обсуждается в литературе, а методы проверки однородности применяются при решении многих практических задач. Например, как сравнить две группы — мужчин и женщин, молодых и пожилых, и т.п.? В маркетинге это важно для сегментации рынка. Если две группы не отличаются по ответам, значит, их можно объединить в один сегмент и проводить по отношению к ним одну и ту же маркетинговую политику, в частности, осуществлять одни и те же рекламные воздействия. Если же две группы различаются, то и относиться к ним надо по-разному. Это — представители двух разных сегментов рынка, требующих разного подхода при борьбе за их завоевание.

Обсуждаемая далее постановка задачи в статистических терминах такова. Рассматривается вопрос с двумя возможными ответами, например, «да» и «нет». В первой группе из n_1 опрошенных m_1 человек сказали «да», а во второй группе из n_2 опрошенных m_2 сказали «да». В вероятностной модели предполагается, что m_1 и m_2 — биномиальные случайные величины $B(n_1, p_1)$ и $B(n_2, p_2)$ соответственно. Запись $B(n, p)$ означает, что случайная величина m имеет биномиальное распределение с параметрами n — объем выборки и p — вероятность определенного ответа (скажем, ответа «да»). Такая случайная величина может быть представлена в виде суммы $m = X_1 + X_2 + ... + X_n$, где случайные величины $X_1, X_2, ..., X_n$ независимы, одинаково распределены, принимают два значения 1 и 0, причем $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$, $i = 1,2, ..., n$.

Однородность двух групп означает, что соответствующие им вероятности равны, неоднородность — что эти вероятности отличаются. В терминах прикладной математической статистики задача ставится так: необходимо проверить гипотезу однородности:

$$H_0: p_1 = p_2$$

при альтернативной гипотезе о наличии эффекта:

$$H_1: p_1 \neq p_2.$$

(Иногда представляют интерес односторонние альтернативные гипотезы $H_1': p_1 > p_2$ и $H_1'': p_1 < p_2$.)

Оценкой вероятности p_1 является частота $p_1^*=m_1/n_1$, а оценкой вероятности p_2 является частота $p_2^*=m_2/n_2$. Даже при совпадении вероятностей p_1 и p_2 частоты, как правило, различаются. Как говорят, «по чисто случайным причинам». Рассмотрим случайную величину $p_1^* - p_2^*$. Тогда:

$$M(p_1^* - p_2^*) = p_1 - p_2,\ D(p_1^* - p_2^*) = p_1 (1 - p_1)/n_1 + p_2(1 - p_2)/n_2.$$

Из теоремы Муавра — Лапласа и теоремы о наследовании сходимости (глава 14 и [4, п. 2.4]) следует, что:

$$\lim_{n_1\to\infty, n_2\to\infty} P\left\{\frac{p_1^* - p_2^* - M(p_1^* - p_2^*)}{\sqrt{D(p_1^* - p_2^*)}} \le x\right\} = \Phi(x),$$

где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Для практического применения этого соотношения следует заменить неизвестную статистику дисперсию разности частот на оценку этой дисперсии:

$$D^*(p_1^* - p_2^*) = p^*_1 (1 - p^*_1)/n_1 + p^*_2 (1 - p^*_2)/n_2.$$

(Могут использоваться и другие оценки рассматриваемой дисперсии, например, по объединенной выборке.) С помощью указанной выше математической техники можно показать, что при такой замене предельное распределение не меняется:

При справедливости гипотезы однородности (т.е. при отсутствии эффекта) $M(p_1^* - p_2^*) = 0$. Поэтому правило принятия решения при проверке однородности двух выборок выглядит так:

$$\lim_{n_1\to\infty, n_2\to\infty} P\{\frac{p_1^* - p_2^* - M(p_1^* - p_2^*)}{\sqrt{D^*(p_1^* - p_2^*)}} \le x\} = \Phi(x).$$

1. Вычислить статистику:

$$Q = \frac{p_1^* - p_2^*}{\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}}.$$

2. Сравнить значение модуля статистика $|Q|$ с граничным значением K. Если $|Q| \le K$, то принять гипотезу однородности H_0 . Если же $|Q| > K$, то заявить об отсутствии однородности и принять альтернативную гипотезу H_1.

Граничное значение K определяется выбором уровня значимости статистического критерия проверки однородности. Из приведенных выше предельных соотношений следует, что при справедливости гипотезы однородности H_0 для уровня значимости $\alpha = P(|Q| > K)$ имеем (при $n_1 \to \infty, n_2 \to \infty$):

$$\alpha \to \Phi(K) - \Phi(-K) = 2\Phi(K) - 1.$$

Следовательно, граничное значение в зависимости от уровня значимости целесообразно выбирать из условия:

$$K = K(\alpha) = \Phi^{-1}\left(\frac{1+\alpha}{2}\right).$$

Здесь $\Phi^{-1}(\cdot)$ — функция, обратная к функции стандартного нормального распределения. В социально-экономических исследованиях наиболее распространен 5 % уровень значимости, т.е. $\alpha = 0{,}05$. Для него $K = 1{,}96$.

Пример 2. Пусть в первой группе из 500 опрошенных мужчин ответили «да» 200, а во второй группе из 700 опрошенных женщин сказали «да» 350. Есть ли разница по доле отвечающих «да» между генеральными совокупностями, представленными этими двумя группами?

Для установления взаимопонимания с маркетологом уберем из формулировки примера относящийся к теории статистики термин «генеральная совокупность». Получим следующую постановку.

Пусть из 500 опрошенных мужчин ответили «да, я люблю “Пепси-колу”» 200, а из 700 опрошенных женщин 350 сказали «да, я люблю “Пепси-колу”». Есть ли разница между мужчинами и женщинами по доле отвечающих «да» на вопрос о любви к «Пепси-коле»?

В рассматриваемом примере нужные для расчетов величины таковы: $n_1 = 500, p_1^* = 200/500 = 0{,}4; n_2 = 700, p_2^* = 350/700 = 0{,}5$. Вычислим статистику:

$$Q = \frac{0{,}4 - 0{,}5}{\sqrt{\frac{0{,}4\cdot 0{,}6}{500} + \frac{0{,}5\cdot 0{,}5}{700}}} = \frac{-0{,}1}{\sqrt{\frac{0{,}24}{500} + \frac{0{,}25}{700}}} = \frac{-0{,}1}{\sqrt{0{,}00048 + 0{,}0003571}} =$$

$$= \frac{-0{,}1}{\sqrt{0{,}0008371}} = \frac{-0{,}1}{0{,}029} = -3{,}45.$$

Поскольку $|Q| = 3{,}45 > 1{,}96$, то необходимо отклонить нулевую гипотезу и принять альтернативную. Таким образом, мужчины и женщины отличаются по рассматриваемому признаку — любви к «Пепси-коле».

Необходимо отметить, что результат проверки гипотезы однородности зависит не только от частот, но и от объемов выборок. Предположим, что частоты (доли) зафиксированы, а объемы выборок растут. Тогда числитель статистики Q не меняется, а знаменатель уменьшается, значит, вся дробь возрастает. Поскольку знаменатель стремится к 0, то дробь возрастает до бесконечности и рано или поздно превзойдет любую границу. Есть только одно исключение — когда в числителе стоит 0. Следовательно, при строгом подходе к формулировкам вывод статистика должен выглядеть так: «различие обнаружено» или «различие не обнаружено». Во втором случае различие, возможно, было бы обнаружено при увеличении объемов выборок.

Как и для доверительного оценивания вероятности, во ВЦИОМ разработаны две полезные таблицы, позволяющие оценить вызванные чисто случайными причинами допустимые расхождения между частотами в группах. Эти таблицы рассчитаны при выполнении нулевой гипотезы однородности и соответствуют ситуациям, когда частоты близки к 50 % (табл. 9) или к 20 % (табл. 10). Если наблюдаемые частоты — от 30 % до 70 %, то рекомендуется пользоваться первой из этих таблиц, если от 10 % до 30 % или от 70 % до 90 % — то второй. Если наблюдаемые частоты меньше 10 % или больше 90 %, то теорема Муавра-Лапласа и основанные на ней асимптотические формулы дают не очень хорошие приближения, целесообразно применять иные, более продвинутые математические средства, в частности, приближения с помощью распределения Пуассона.

Таблица 9

Допустимые расхождения (в %) между частотами в двух группах, когда наблюдаются частоты от 30 % до 70 %

Объемы групп	750	600	400	200	100
750	6	7	7	10	12
600	7	8	8	11	13
400	7	8	10	11	14
200	10	11	11	13	16
100	12	13	14	16	18

Таблица 10

Допустимые расхождения (в %) между частотами в двух группах, когда наблюдаются частоты от 10 до 30 % или от 70 до 90 %

Объемы групп	750	600	400	200	100
750	5	5	6	8	10
600	5	6	7	8	10
400	6	7	8	9	11
200	8	8	9	10	12
100	10	10	11	12	14

В условиях разобранного выше примера табл. 9 дает допустимое расхождение 7 %. Действительно, объем первой группы 500 отсутствует в таблице, но строки, соответствующие объемам 400 и 600, совпадают для первых двух столбцов слева. Эти столбцы соответствуют объемам второй группы 750 и 600, между которыми расположен объем 700, данный в примере. Он ближе к 750, поэтому берем величину расхождения, стоящую на пересечении первого столбца и второй (и третьей) строк, т.е. 7 %. Поскольку реальное расхождение (10 %) больше, чем 7 %, то делаем вывод о наличии значимого различия между группами. Естественно, этот вывод совпадает с полученным ранее расчетным путем.

Как и в случае табл. 5, значения в таблицах 9 и 10 несколько больше, чем рассчитанные по приведенным выше формулам. Дело в том, что таблицы ВЦИОМ связаны не с уровнем значимости $\alpha = 0{,}05$, а с уровнем значимости $\alpha = 0{,}01$, которому соответствует граничное значение 2,58.

Допустимое расхождение $\Delta = \Delta(\alpha)$ между частотами нетрудно получить расчетным путем. Для этого достаточно воспользоваться формулой для статистики Q и определить, при каком максимальном расхождении частот все еще делается вывод о том, что верна гипотеза однородности. Следовательно, допустимое расхождение $\Delta = \Delta(\alpha)$ находится из уравнения:

$$K(\alpha) = \frac{\Delta(\alpha)}{\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}}.$$

Таким образом,

$$\Delta(\alpha) = K(\alpha)\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}.$$

Для данных примера 2 $\Delta = \Delta(\alpha)$ = 1,96 · 0,029 = 0,057, или 5,7 %, для уровня значимости 0,05.

Для других уровней значимости надо использовать другие коэффициенты $K(\alpha)$. Так, K(0,01) = 2,58 для уровня значимости 1 % и K(0,10) = 1,64 для уровня значимости 10 %. Для данных примера $\Delta = \Delta(\alpha)$ = 2,58 · 0,029 = 0,7482 ≈ ≈ 0,075, или 7,5 %, для уровня значимости 0,01. Если округлить до ближайшего целого числа процентов, то получим 7 %, как при использовании табл. 9 выше.

Анализ таблиц 9 и 10 показывает, что для обнаружения эффекта (констатации различия генеральных совокупностей) частоты должны отличаться не менее чем на 6 %. А при некоторых объемах выборок — более чем на 10 %, например, при объемах выборок 100 и 100 — на 19 %. Если же частоты отличаются на 5 % или менее, можно сразу сказать, что статистический анализ приведет к выводу о том, что различие не обнаружено (для выборок объемов не более 750).

В связи со сказанным возникает вопрос: каково типовое отличие частот в двух выборках из одной и той же совокупности? Разность частот в этом случае имеет нулевое математическое ожидание и дисперсию:

$$p(1-p)(\frac{1}{n_1}+\frac{1}{n_2})=\frac{p(1-p)(n_1+n_2)}{n_1 n_2}.$$

Величина $p(1\text{-}p)$ достигает максимума при p=1/2, и этот максимум равен 1/4. Если p=1/2, а объемы двух выборок совпадают и равны 500, то дисперсия разности частот равна:

$$\frac{0{,}25\times 1000}{500\times 500}=\frac{250}{250\times 1000}=\frac{1}{1000}.$$

Следовательно, среднее квадратического отклонение σ равно 0,032, или 3,2 %. Поскольку для стандартной нормальной случайной величины в 50 % случаев ее значение не превосходит по модулю 0,67 (а в 50 % случаев — больше 0,67), то типовой разброс равен $0{,}67\sigma$, а в рассматриваемом случае — 2,1 %.

Приведенные соображения дают возможность построить метод контроля правильности (корректности) проведения повторных опросов. Если частоты излишне устойчивы, значения при повторных опросах слишком близки — это подозрительно! Возможно, нарушены правила проведения опросов, выборки не являются случайными, ответы фальсифицированы, и т.д.

Литература

1. *Сэндидж, Ч.* Реклама: теория и практика / Ч. Сэндидж, В. Фрайбургер, К. Ротцолл : перевод с английского. — Москва : Прогресс, 1989. — 630 с.

2. *Ядов, В.А.* Стратегии и методы качественного анализа данных / В.А. Ядов // Социология: методология, методы, математические модели. — 1991. — № 1. — С. 14–31.

3. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1983. — 416 с.

4. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

5. Опыт применения ЭВМ в социологических исследованиях. — Москва : Институт социологических исследований АН СССР ; Советская социологическая ассоциация, 1977. — 158 с.

6. *Орлов, А.И.* Общий взгляд на статистику объектов нечисловой природы / А.И. Орлов // Анализ нечисловой информации в социологических исследованиях. — Москва : Наука, 1985. — С. 58–92.

7. *Орлов, А.И.* Метод ценообразования на основе оценивания функции спроса / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 158. — С. 250–267.

8. *Агаларов, З.С.* Эконометрика : учебник / З.С. Агаларов, А.И. Орлов. — Москва : Дашков и К°, 2021. — 380 с.

Контрольные вопросы и задачи

1. Почему выборочные исследования необходимы для решения многих практических задач?

2. Рассчитайте коэффициент ценовой эластичности спроса по данным табл. 1 при цене $p = 35$ и при цене $p = 40$.

3. Какова роль теоремы Муавра-Лапласа в теории выборочных исследований?

В задачах 4–8 выберите наиболее подходящий вариант ответа.

4. При какой цене максимальна прибыль в условиях пункта 1.1, если издержки (оптовая цена товара) равны 12?

а) 30;

б) 32;

в) 35.

5. При исследовании предпочтений потребителей открытые вопросы:

а) труднее для опрашиваемых, но легче для обработки;

б) легче для опрашиваемых, но труднее для обработки.

6. Пусть из 657 опрошенных 289 сказали «да». Доверительный интервал для доли отвечающих «да» в генеральной совокупности, соответствующий доверительной вероятности 0,95, таков:

а) [0,245; 0,398];

б) [0,435; 0,445];

в) [0,405; 0,556];

г) [0,402; 0,478];

д) [0,247; 0,633].

7. Из 513 юношей 193 любят «Сникерс», а из 748 девушек — 327. Значение статистического критерия Q для проверки гипотезы о равенстве вероятностей равно:

а) 3,38;

б) – 2,176;

в) 0,25;

г) 12,56;

д) – 0,173.

8. В условиях задачи 7 гипотеза об одинаковой привлекательности «Сникерса» для юношей и девушек (на уровне значимости 0,05):

а) принимается;

б) отклоняется.

9. Как понятие допустимого расхождения между частотами можно использовать при планировании выборочных исследований?

Темы докладов, рефератов, исследовательских работ

1. Проведите выборочное исследование с целью построения оценки функции ожидаемого спроса на выбранный Вами товар (услугу).

2. Найдите адекватное приближение функции спроса в табл. 1 с помощью метода наименьших квадратов.

3. Постройте экономико-математическую модель оптимизации цены при заданных функциях спроса (в зависимости о цены) и издержек (в зависимости от выпуска).

4. Сопоставьте теорию квотной выборки с теорией простой случайной выборки.

5. Рассмотрите статистическую теорию доверительного оценивания и проверки гипотез о равенстве вероятностей в случае нескольких возможных ответов (соответственно, с использованием мультиномиального распределения вместо биномиального).

6. В каких случаях может быть использована теория малых выборок (теорема Пуассона) для доверительного оценивания вероятности определенного ответа?

ГЛАВА 2. ОПИСАНИЕ ДАННЫХ

Выделяют три основные области статистических методов обработки результатов наблюдений — описание данных, оценивание (характеристик и параметров распределений, регрессионных зависимостей и др.) и проверка статистических гипотез.

Описание данных — предварительный этап статистической обработки. Используемые при описании данных величины применяются при дальнейших этапах статистического анализа — оценивании и проверке гипотез, а также при решении иных задач, возникающих при применении вероятностно-статистических методов принятия решений, например, при статистическом контроле качества продукции и статистическом регулировании технологических процессов.

2.1. Модели порождения данных

Статистические данные — это результаты наблюдений (измерений, испытаний, опытов, анализов). Функции результатов наблюдений, используемые, в частности, для оценки параметров распределений и (или) для проверки статистических гипотез, называют «статистиками». (Для математиков надо добавить, что речь идет об измеримых функциях.) Если в вероятностной модели результаты наблюдений рассматриваются как случайные величины (или случайные элементы), то статистики, как функции случайных величин (элементов), сами являются случайными величинами (элементами). Статистики, являющиеся выборочными аналогами характеристик случайных величин (математического ожидания, медианы, дисперсии, моментов и др.) и используемые для оценивания этих характеристик, называют статистическими характеристиками.

Виды выборок. Основополагающее понятие в вероятностно-статистических методах принятия решений — выборка. Как уже говорилось, выборка — это:

1) набор наблюдаемых значений;

2) множество объектов, отобранных из изучаемой совокупности.

Например, единицы продукции, отобранные из контролируемой партии или потока продукции для контроля и принятия решений. Наблюдаемые значения обозначим x_1, x_2, ..., x_n, где n — объем выборки, т.е. число наблюдаемых значений, составляющих выборку. О втором виде выборок уже шла речь в главе 1 при рассмотрении гипергеометрического распределения, когда под выборкой понимался набор единиц продукции, отобранных из партии. Там же обсуждалась вероятностная модель случайной выборки.

В вероятностной модели выборки первого типа наблюдаемые значения обычно рассматривают как реализацию независимых одинаково распределенных случайных величин $X_1(\omega), X_2(\omega),...,X_n(\omega), \omega \in \Omega$. При этом считают, что полученные при наблюдениях конкретные значения x_1, x_2, ..., x_n соответствуют определенному элементарному событию $\omega = \omega_0$, т.е.:

$$x_1 = X_1(\omega_0), x_2 = X_2(\omega_0),...,x_n = X_n(\omega_0), \omega_0 \in \Omega.$$

При повторных наблюдениях будут получены иные наблюдаемые значения, соответствующие другому элементарному событию $\omega = \omega_1$. Цель обработки статистических данных состоит в том, чтобы по результатам наблюдений, соответствующим элементарному событию $\omega = \omega_0$, сделать выводы о вероятностной мере P и результатах наблюдений при различных возможных $\omega = \omega_1$.

Для выборок второго вида отбор объектов может проводиться в несколько этапов. Например, для входного контроля сигарет могут сначала отбираться коробки, в отобранных коробках — блоки, в выбранных блоках — пачки, а в пачках — сигареты. Четыре ступени отбора. Ясно, что выборка будет обладать иными свойствами, чем простая случайная выборка из совокупности сигарет.

Частоты. Из приведенного выше определения математической статистики следует, что описание статистических данных дается с помощью частот. Частота — это отношение числа X наблюдаемых единиц, которые принимают заданное значение или лежат в заданном интервале, к общему числу наблюдений n, т.е. частота — это X/n. (В более старой литературе иногда X/n называется относительной частотой, а под частотой имеется в виду X. В старой терминологии можно сказать, что относительная частота — это отношение частоты к общему числу наблюдений.)

Число X имеет биномиальное распределение, задаваемое вероятностью p того, что случайная величина, с помощью которой моделируются результаты наблюдений, принимает заданное значение или лежит в заданном интервале, и общим числом наблюдений n. Из закона больших чисел (теорема Бернулли) следует, что:

$$\frac{X}{n} \to p$$

при $n \to \infty$ (сходимость по вероятности), т.е. частота сходится к вероятности. Теорема Муавра — Лапласа позволяет уточнить скорость сходимости в этом предельном соотношении (см. главу 1).

Детерминированный и модельно-вероятностный подходы. В статистических методах есть два подхода к исходным данным — детерминированный и модельно-вероятностный. В первом из них данные рассматриваются сами по себе, без попыток связать их с какой-либо более общей ситуацией. Например, при анализе данных о производственной деятельности конкретного предприятия за конкретный период времени подсчитывается процент брака по конкретным технологическим процессам, число работников на различных должностях, объем реализованной продукции по месяцам. К этой же категории данных относятся различные виды отчетности — бухгалтерская, налоговая, статистическая (для органов Росстата РФ). Преимуществом детерминированного подхода является отсутствие каких-либо дополнительных предположений о данных. Недостаток состоит в невозможности обоснованного переноса выводов с конкретной ситуации на другие, ей аналогичные. Например, на другие периоды времени или на другие предприятия. При детерминированном подходе невозможно также оценить погрешность рассчитанных характеристик.

Чтобы выйти за пределы конкретной ситуации, необходимо использовать модельно-вероятностный подход, согласно которому основой алгоритмов расчетов является вероятностная модель порождения данных. При этом конкретные данные рассматриваются как реализации случайных величин, векторов, более общо — элементов, т.е. как значения задающих их функций, определенных на вероятностном пространстве, в конкретной точке (элементарном событии ω).

Наиболее распространенная вероятностная модель порождения данных — это модель случайной выборки. Согласно этой модели данные $x_1, x_2, \ldots, x_n$ рассматриваются как реализация независимых одинаково распределенных случайных элементов (величин, векторов, множеств и других объектов нечисловой природы) $X_1 = X_1(\omega), X_2 = X_2(\omega), \ldots, X_n = X_n(\omega)$, т.е. $x_1 = X_1(\omega_0), x_2 = X_2(\omega_0), \ldots, x_n = X_n(\omega_0)$ при некотором ω_0 из пространства элементарных событий Ω. Модель выборки обычно используется для описания результатов независимых наблюдений, измерений, анализов. опытов.

В некоторых случаях используют более специальные модели порождения данных. Например, при проведении испытаний на надежность используют план испытаний, согласно которому испытания прекращаются через время T. Это значит, что фиксируются только моменты отказа изделий, которые произошли до момента T. Пусть $x_1, x_2, \ldots, x_n$ — наработки на отказ n изделий. Статистику доступны только значения $y_1, y_2, \ldots, y_n$, где $y_j = x_j$ при $x_j < T$ и $y_j = T$ при $x_j \geq T$. Такая выборка, в которой часть описывающих реальное явление случайных величин заменена граничным значением, называется цензурированной. Иногда используются и более сложные модели порождения данных. Например, если аппаратурой не фиксируются значения, меньшие некоторого порога, то выборка не только цензурирована, но и состоит из случайного числа элементов. Бывают и процедуры, когда минимальный и максимальный элементы выборки отбрасываются, а остальные предоставляются статистику, и т.д.

Параметрические и непараметрические модели случайной выборки. Рассмотрим ситуацию, когда элементы выборки — числа. Модель описывается функцией распределения элементов выборки. Можно ли что-либо сказать об этой функции?

В учебных курсах по теории вероятностей и математической статистике часто рассматривают различные параметрические семейства распределений числовых случайных величин. А именно, изучают семейства нормальных распределений, логарифмически нормальных, экспоненциальных, гамма-распределений, распределений Вейбулла — Гнеденко и др. Все они зависят от одного, двух или трех параметров. Поэтому для полного описания распределения достаточно знать (или оценить по выборке) одно, два или три числа. Очень удобно. Поэтому широко развита и представлена в литературе параметрическая теория математической статистики, в которой предполагается, что распределения результатов наблюдений принадлежат тем или иным параметрическим семействам.

К сожалению, параметрические семейства существуют лишь в головах авторов учебников по теории вероятностей и математической статистике. В реальной жизни их нет. Поэтому прикладная статистика использует в основном непараметрические методы, в которых распределения результатов наблюдений могут иметь произвольный вид.

В настоящем разделе на примере нормального распределения подробно обоснуем невозможность практического использования параметрических семейств для описания распределений конкретных данных. В главе 4.2 разобраны параметрические методы отбраковки резко выделяющихся наблюдений и продемонстрирована невозможность практического использования ряда методов параметрической статистики, ошибочность выводов, к которым они приводят. В главе 5.1 рассмотрены непараметрические методы доверительного оценивания основных характеристик числовых случайных величин — математического ожидания, медианы, дисперсии, среднего квадратического отклонения, коэффициента вариации.

К настоящему времени непараметрические методы полностью покрывают область задач, которые ранее решались с помощью параметрической статистики. Поэтому можно порекомендовать использовать только непараметрическую статистику. Однако в литературе много внимания уделяется параметрическим методам, поэтому игнорировать в настоящем учебнике параметрическую статистику было признано нецелесообразным.

Часто ли распределение результатов наблюдений является нормальным? Статистические модели применяют при изучении и оптимизации процессов маркетинга, управления предприятием и регионом, точности и стабильности технологических процессов. Их используют в задачах надежности, обеспечения безопасности, в том числе экологической, для описания функционирования технических устройств и объектов, при разработке организационных схем и т.д. При этом зачастую используют те или иные параметрические семейства распределений вероятностей. Наиболее популярно нормальное распределение. Используют также логарифмически нормальное распределение, экспоненциальное распределение, гамма-распределение, распределение Вейбулла — Гнеденко и т.д.

Очевидно, всегда необходимо проверять соответствие моделей реальности. Возникают два вопроса. Отличаются ли реальные распределения от используемых в модели? Насколько это отличие влияет на выводы?

На примере нормального распределения покажем, что реальные распределения практически всегда отличаются от включенных в классические параметрические семейства. Имеющиеся отклонения от заданных семейств делают неверными выводы, основанные на использовании этих семейств. Например, выводы об отбраковке резко отличающихся наблюдений (выбросов).

Есть ли основания априори предполагать нормальность результатов измерений?

Иногда утверждают, что в случае, когда погрешность измерения (или иная случайная величина) определяется в результате совокупного действия многих малых факторов, то в силу Центральной Предельной Теоремы (ЦПТ) теории вероятностей эта величина хорошо приближается (по распределению) нормальной случайной величиной. Такое утверждение справедливо, если малые факторы действуют аддитивно и независимо друг от друга. Если же они действуют мультипликативно, то в силу той же ЦПТ аппроксимировать надо логарифмически нормальным распределением. В прикладных задачах обосновать аддитивность, а не мультипликативность действия малых факторов обычно не удается. Если же зависимость имеет общий характер, не приводится к аддитивному или мультипликативному виду, а также нет оснований принимать модели, дающие экспоненциальное, Вейбулла — Гнеденко, гамма или иные распределения, то о распределении итоговой случайной величины практически ничего не известно, кроме внутриматематических свойств типа регулярности.

Экспериментальное изучение распределений погрешностей. При обработке конкретных данных иногда считают, что погрешности измерений имеют нормальное распределение. На предположении нормальности построены классические модели регрессионного, дисперсионного, факторного анализов, метрологические модели, которые еще продолжают встречаться как в отечественной нормативно-технической документации, так и в международных стандартах. На то же предположение опираются модели расчетов максимально достигаемых уровней тех или иных характеристик, применяемые при проектировании систем обеспечения безопасности функционирования экономических структур, технических устройств и объектов. Однако теоретических оснований для такого предположения нет. Необходимо экспериментально изучать распределения погрешностей.

Что же показывают результаты экспериментов? Сводка, данная в монографии [1], позволяет утверждать, что в большинстве случаев распределение погрешностей измерений отличается от нормального. Так, в Машинно-электротехническом институте (г. Варна, Болгария) было исследовано распределение погрешностей градуировки шкал аналоговых электроизмерительных приборов. Изучались приборы, изготовленные в Чехословакии, СССР и Болгарии. Закон распределения погрешностей оказался одним и тем же. Он имеет плотность:

$$f(x) = 0{,}534 \exp(1 - |x|^7).$$

Были проанализированы данные о параметрах 219 фактических распределений погрешностей, исследованных разными авторами, при измерении как электрических, так и иных физических величин самыми разнообразными (электрическими) приборами. В результате этого исследования оказалось, что 111 распределений, т.е. примерно 50 %, принадлежат классу распределений с плотностью:

$$f(x;\alpha,b.\sigma) = \frac{\alpha}{2\lambda\sigma\Gamma(1/\alpha)} \exp\left(-\left|\frac{x-b}{\lambda\sigma}\right|^{\alpha}\right),$$

где α — параметр степени (формы); b — параметр сдвига; σ — параметр масштаба; $\Gamma(\beta)$ — гамма-функция от аргумента β;

$$\lambda = \sqrt{\frac{\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}$$

(см. [1, с. 56]); 63 распределения, т.е. 30 %, имеют плотности с плоской вершиной и пологими длинными спадами и не могут быть описаны как нормальные или, например, экспоненциальные. Оставшиеся 45 распределений оказались двухмодальными.

В книге известного метролога профессора П.В. Новицкого [2] приведены результаты исследования законов распределения для различного рода погрешностей измерения. Он изучил распределения погрешностей электромеханических приборов на кернах, электронных приборов для измерения температур и усилий, цифровых приборов с ручным уравновешиванием. Объем выборок экспериментальных данных для каждого экземпляра составлял 100–400 отсчетов. Оказалось, что 46 из 47 распределений значимо отличались от нормального. Исследована форма распределения погрешностей у 25 экземпляров цифровых вольтметров Щ-1411 в 10 точках диапазона. Результаты аналогичны. Дальнейшие сведения содержатся в монографии [1].

В лаборатории прикладной математики Тартуского государственного университета проанализировано более 2 500 выборок из архива реальных статистических данных. В 92 % гипотезу нормальности пришлось отвергнуть.

Приведенные описания экспериментальных данных показывают, что погрешности измерений в большинстве случаев имеют распределения, отличные от нормальных. Это означает, в частности, что большинство применений критерия Стьюдента, классического регрессионного анализа и других статистических методов, основанных на нормальной теории, строго говоря, не является обоснованным. Поскольку неверна лежащая в их основе аксиома нормальности распределений соответствующих случайных величин.

Очевидно, для оправдания или обоснованного изменения существующей практики анализа статистических данных требуется изучить свойства процедур анализа данных при «незаконном» применении. Изучение процедур отбраковки показало, что они крайне неустойчивы к отклонениям от нормальности, а потому применять их для обработки реальных данных нецелесообразно (см. раздел 4.2); поэтому нельзя утверждать, что произвольно взятая процедура устойчива к отклонениям от нормальности.

Иногда предлагают перед применением, например, критерия Стьюдента однородности двух выборок проверять нормальность. Хотя для этого имеется много критериев, но проверка нормальности — более сложная и трудоемкая статистическая процедура, чем проверка однородности (как с помощью статистик типа Стьюдента, так и с помощью непараметрических критериев). Для достаточно надежного установления нормальности требуется весьма большое число наблюдений. Так, чтобы гарантировать, что функция распределения результатов наблюдений отличается от некоторой нормальной не более, чем на 0,01 (при любом значении аргумента), требуется порядка 2 500 наблюдений. В большинстве экономических, технических, медико-биологических и других прикладных исследований число наблюдений существенно меньше. Особенно это справедливо для данных, используемых при изучении проблем, связанных с обеспечением безопасности функционирования экономических структур и технических объектов.

ЦПТ и нормальность. Иногда пытаются использовать ЦПТ для приближения распределения погрешности к нормальному, включая в технологическую схему измерительного прибора специальные сумматоры. Оценим полезность этой меры. Пусть Z_1 , Z_2 , ..., Z_k — независимые одинаково распределенные случайные величины с функцией распределения $H = H(x)$ такие, что:

$$M(Z_1) = 0,\ \ D(Z_1) = 1,\ \ M\,|\,Z_1\,|^3 = \rho < +\infty.$$

Рассмотрим

$$w = \frac{Z_1 + Z_2 + \ldots + Z_k}{\sqrt{k}}.$$

Показателем обеспечиваемой сумматором близости к нормальности является:

$$C = \sup_H \sup_x |\,P(w < x) - \Phi(x)\,|\,.$$

Тогда

$$0{,}3989\frac{\rho}{\sqrt{k}} \le C \le 0{,}7975\frac{\rho}{\sqrt{k}}.$$

Правое неравенство в последнем соотношении вытекает из оценок константы в неравенстве Берри — Эссеена, полученном в книге [3, с. 172], а левое — из примера в монографии [4, с. 140–141]. Для нормального закона ρ =1,6, для равномерного ρ = 1,3, для двухточечного ρ = 1 (это — нижняя граница для ρ). Следовательно, для обеспечения расстояния (в метрике Колмогорова) до нормального распределения не более 0,01 для «неудачных» распределений необходимо не менее k_0 слагаемых, где:

$$0{,}4\sqrt{k_0} < 0{,}01,\ \ k_0 > 1600.$$

В обычно используемых сумматорах слагаемых значительно меньше.

Сужая класс возможных распределений H, можно получить, как показано в монографии [5], более быструю сходимость, но теория здесь еще не позволяет сформулировать практических рекомендаций. Кроме того, не ясно, обеспечивает ли близость распределения к нормальному (в определенной метрике) также и близость распределений статистик. Речь идет о сравнении распределения статистики, построенной по случайным величинам, полученным суммированием, с распределением статистики, соответствующей нормальным результатам наблюдений. Видимо, для каждой конкретной статистики необходимы специальные теоретические исследования. Именно к такому выводу приходит автор монографии [5]. В задачах отбраковки выбросов ответ: «Не обеспечивает» (см. ниже).

Отметим, что результат любого реального измерения записывается с помощью конечного числа десятичных знаков, обычно небольшого (2–5), так что любые реальные данные целесообразно моделировать лишь с помощью дискретных случайных величин, принимающих сравнительно небольшое число значений. Нормальное распределение — лишь аппроксимация реального распределения. Так, например, данные конкретного исследования, приведенные в работе [6], принимают значения от 1,0 до 2,2, т.е. всего 13 возможных значений. Из принципа Дирихле следует, что в какой-то точке построенная по данным работы [6] функция распределения отличается от ближайшей функции нормального распределения не менее чем на 1/26, т.е. на 0,04. Кроме того, очевидно, что для нормального распределения случайной величины вероятность попасть в дискретное множество десятичных чисел с заданным числом знаков после запятой равна 0.

Из сказанного выше следует, что результаты измерений и вообще статистические данные имеют свойства, приводящие к тому, что моделировать их следует случайными величинами с распределениями, более или менее отличными от нормальных. В большинстве случаев распределения существенно отличаются от нормальных. В других ситуациях нормальные распределения могут, видимо, рассматриваться как некоторая аппроксимация. Но никогда нет полного совпадения. Отсюда вытекает необходимость изучения свойств классических статистических процедур в неклассических вероятностных моделях (подобно тому, как это сделано в главе 5.2 для критерия Стьюдента). А также целесообразность разработки устойчивых (учитывающих наличие отклонений от нормальности) и непараметрических, в том числе свободных от распределения процедур, их широкого внедрения в практику статистической обработки данных.

Опущенные здесь рассмотрения для других параметрических семейств приводят к аналогичным выводам. Итог можно сформулировать так. Распределения реальных данных практически никогда не входят в какое-либо конкретное параметрическое семейство. Реальные распределения всегда отличаются от тех, что включены в параметрические семейства. Отличия могут быть большими или малыми, но они всегда есть.

2.2. Таблицы и диаграммы

Исходные статистические данные могут быть достаточно обширными. В качестве примера приведем результаты экспертного опроса, проведенного Институтом высоких статистических технологий и эконометрики (табл. 1). В первом столбце приведены номера экспертов, в остальных четырех — четыре прогнозных значения, полученных от каждого эксперта. Отметим, что эксперт № 28 не ответил на вопрос об инфляции. В таблицах реальных данных приходится сталкиваться с пропусками.

Таблица 1

Прогнозы экспертов на 8 декабря (сделаны 19 октября)

№ п/п	Курс доллара США, руб.	Инфляция (%) за период прогноза	Цена батона белого хлеба, руб.	Цена 1 л молока, руб.
1	4 185	4,0	800	1 305
2	4 270	2,8	1 028	1 322

№ п/п	Курс доллара США, руб.	Инфляция (%) за период прогноза	Цена батона белого хлеба, руб.	Цена 1 л молока, руб.
3	3 200	17,0	760	755
4	4 000	16,0	950	1 000
5	3 500	16,0	820	800
6	3 800	5,0	1 000	1 000
7	3 500	3,5	500	1 500
8	3 300	62,0	800	780
9	4 100	54,0	900	899
10	3 560	10,0	870	1 050
11	4 000	54,0	1 000	1 000
12	5 200	54,0	1 500	1 500
13	4 000	9,0	830	1 300
14	6 000	54,0	2 000	2 000
15	4 000	40,0	950	1 200
16	3 400	13,0	750	900
17	3 500	15,0	1 000	1 250
18	4 200	2,5	1 000	1 500
19	3 560	200,0	940	1 200
20	4 300	6,0	950	1 570
21	4 000	3,0	1 000	1 100
22	4 500	12,0	950	1 100
23	4 200	11,0	890	1 100
24	3 900	54,0	1 000	1 000
25	5 500	62,0	1 000	1 400
26	5 000	73,0	1 000	1 200
27	5 600	54,0	1 200	2 000
28	3 900	–	1 500	1 400
29	4 200	38,0	950	1 100
30	3 680	38,0	850	1 100
31	4 000	2,0	840	1 100
32	4 600	46,0	1 000	1 100
33	4 560	92,0	1 300	1 400

Описание данных — это первичное сжатие информации с целью сделать ее более обозримой, легкой для восприятия. Самый древний способ — это составление различных таблиц, вторичных по отношению к таблицам исходных данных.

Например, рассмотрим последний столбец табл. 1. Для лучшего восприятия прогнозов экспертов о цене 1 л молока сгруппируем данные по интервалам, как это сделано в табл. 2.

Таблица 2

Прогнозируемая цена молока

№ п/п	Интервал, руб.	Число ответов
1	700–799	2
2	800–899	2
3	900–999	1
4	1 000–1 099	5
5	1 100–1 199	7
6	1 200–1 299	4
7	1 300–1 399	3
8	1 400–1 499	3
9	1 500–1 599	4
10	2 000	2
	Всего:	33

Группировка данных в табл. 2 по 10 интервалам может показаться слишком дробной. Нетрудно объединить градации и получить, например, табл. 3.

Таблица 3

Прогнозируемая цена молока (крупные градации)

№ п/п	Интервал, руб.	Число ответов
1	700–999	5
2	1 000–1 299	16
3	1 300–1 599	10
4	2 000	2
	Всего:	33

Сколько использовать градаций (т.е. строк в таблице)? Общих рекомендаций дать нельзя. Ответ зависит от цели статистического исследования, от структуры конкретных данных.

Табличный материал может быть выражен в виде различных диаграмм, в том числе круговых и столбчатых. Несколько десятков лет назад были популярны гистограммы — столбчатые диаграммы, для которых интервалы группирования имеют одинаковую длину.

В настоящее время гистограммы рассматривают как устаревшие инструменты статистического анализа. Для описания массива данных рекомендуется использовать вариационные ряды, эмпирические функции распределения и — особенно настоятельно — непараметрические оценки плотности.

2.3. Выборочные характеристики распределения

Для описания данных применяют различные статистические характеристики. В качестве выборочной средней величины чаще всего используют выборочное среднее арифметическое, т.е. сумму значений рассматриваемой величины, полученную по результатам испытания выборки, деленную на ее объем:

$$\bar{x} = \frac{1}{n}\sum_{1 \le i \le n} x_i,$$

где n — объем выборки, x_i — результат измерения (испытания) i-го элемента выборки.

Другой вид выборочного среднего — выборочная медиана. Она определяется через порядковые статистики.

Порядковые статистики — это члены вариационного ряда, который получается, если элементы выборки $x_1, x_2, \ldots, x_n$ расположить в порядке неубывания:

$$x(1) \le x(2) \le \ldots \le x(k) \le \ldots \le x(n).$$

Пример 1. Для выборки $x_1 = 1$, $x_2 = 7$, $x_3 = 4$, $x_4 = 2$, $x_5 = 8$, $x_6 = 0$, $x_7 = 5$, $x_8 = 7$ вариационный ряд имеет вид 0, 1, 2, 4, 5, 7, 7, 8, т.е. $x(1) = 0 = x_6$, $x(2) = 1 = x_1$, $x(3) = 2 = x_4$, $x(4) = 4 = x_3$, $x(5) = 5 = x_7$, $x(6) = x(7) = 7 = x_2 = x_8$, $x(8) = 8 = x_5$.

В вариационном ряду элемент $x(k)$ называется k-той порядковой статистикой. Порядковые статистики и функции от них широко используются в вероятностно-статистических методах принятия решений, в эконометрике и в других прикладных областях [7].

Выборочная медиана $\tilde{x}$ — результат наблюдения, занимающий центральное место в вариационном ряду, построенном по выборке с нечетным числом элементов, или полусумма двух результатов наблюдений, занимающих два центральных места в вариационном ряду, построенном по выборке с четным числом элементов. Таким образом, если объем выборки n — нечетное число, $n = 2k+1$, то медиана $\tilde{x} = x(k+1)$, если же n — четное число, $n = 2k$, то медиана $\tilde{x} = [x(k) + x(k+1)]/2$, где $x(k)$ и $x(k+1)$ — порядковые статистики.

В качестве выборочных показателей рассеивания результатов наблюдений чаще всего используют выборочную дисперсию, выборочное среднее квадратическое отклонение и размах выборки.

Согласно [8] выборочная дисперсия s^2 — это сумма квадратов отклонений выборочных результатов наблюдений от их среднего арифметического, деленная на объем выборки:

$$s^2 = \frac{1}{n} \sum_{1 \le i \le n} (x_i - \bar{x})^2.$$

Выборочное среднее квадратическое отклонение s — неотрицательный квадратный корень из дисперсии, т.е. $s = +\sqrt{s^2}$.

В некоторых литературных источниках выборочной дисперсией называют другую величину:

$$s_0^2 = \frac{1}{n-1} \sum_{1 \le i \le n} (x_i - \bar{x})^2.$$

Она отличается от s^2 постоянным множителем:

$$s^2 = \left(1 - \frac{1}{n}\right) s_0^2.$$

Соответственно выборочным средним квадратическим отклонением в этих литературных источниках называют величину $s_0 = +\sqrt{s_0^2}$. Тогда очевидно:

$$s = \sqrt{1 - \frac{1}{n}} s_0.$$

Различие в определениях приводит к различию в алгоритмах расчетов, правилах принятия решений и соответствующих таблицах. Поэтому при использовании тех или иных нормативно-технических и инструктивно-методических материалов, программных продуктов, таблиц необходимо обращать внимание на способ определения выборочных характеристик.

Выбор s_0^2, а не s^2 объясняется тем, что:

$$M(s_0^2) = D(X) = \sigma^2,$$

где X — случайная величина, имеющая такое же распределение, как и результаты наблюдений. В терминах теории статистического оценивания это означает, что s_0^2 — несмещенная оценка дисперсии. В то же время статистика s^2 не является несмещенной оценкой дисперсии результатов наблюдений, поскольку:

$$M(s^2) = \left(1 - \frac{1}{n}\right)\sigma^2.$$

Однако у s^2 есть другое свойство, оправдывающее использование этой статистики в качестве выборочного показателя рассеивания. Для известных результатов наблюдений $x_1, x_2, \ldots, x_n$ рассмотрим случайную величину $У$ с распределением вероятностей:

$$P(Y = x_i) = \frac{1}{n}, \quad i = 1,2,\ldots,n,$$

и $P(У = x) = 0$ для всех прочих x. Это распределение вероятностей называется *эмпирическим*. Тогда функция распределения $У$ — это эмпирическая функция распределения, построенная по результатам наблюдений $x_1, x_2, \ldots, x_n$. Вычислим математическое ожидание и дисперсию случайной величины $У$:

$$M(Y) = \bar{x}, \quad D(Y) = s^2.$$

Второе из этих равенств и является основанием для использования s^2 в качестве выборочного показателя рассеивания.

Отметим, что математические ожидания выборочных средних квадратических отклонений $M(s)$ и $M(s_0)$, вообще говоря, не равняются теоретическому среднему квадратическому отклонению σ. Например, если X имеет нормальное распределение, объем выборки $n = 3$, то:

$$M(s) = 0{,}724\sigma, \; M(s_0) = 0{,}887\sigma.$$

Кроме перечисленных выше статистических характеристик, в качестве выборочного показателя рассеивания используют размах R — разность между n-й и первой порядковыми статистиками в выборке объема n, т.е. разность между наибольшим и наименьшим значениями в выборке: $R = x(n) - x(1)$.

В ряде вероятностно-статистических методов применяют и иные показатели рассеивания. В частности, в методах статистического регулирования процессов используют средний размах — среднее арифметическое размахов, полученных в определенном количестве выборок одинакового объема. Популярно и межквартильное расстояние, т.е. расстояние между выборочными квартилями $x([0,75n])$ и $x([0,25n])$ порядка 0,75 и 0,25 соответственно, где $[0,75n]$ — целая часть числа $0,75n$, а $[0,25n]$ — целая часть числа $0,25n$.

Подведем итоги. Целесообразно рассчитывать и приводить в документации статистического исследования в разделе «Описание данных» выборочные характеристики:

- выборочное среднее арифметическое;
- выборочную дисперсию;
- выборочное среднее квадратическое отклонение;
- выборочный коэффициент вариации $v_n = \frac{s}{\bar{x}}$;
- выборочную медиану;
- минимум (первый член вариационного ряда);
- максимум (последний член вариационного ряда);
- размах;
- моду (наиболее часто встречающееся значение) и амплитуду моды (число элементов выборки, равных моде);
- верхний квартиль $x([0,75n])$;
- нижний квартиль $x([0,25n])$;
- межквартильное расстояние.

2.4. Эмпирическая функция распределения

Чтобы от отдельных событий перейти к одновременному рассмотрению многих событий, используют накопленную частоту. Так называется отношение числа единиц, для которых результаты наблюдения меньше заданного значения, к общему числу наблюдений. Функция, которая выражает зависимость между значениями количественного признака и накопленной частотой, называется эмпирической функцией распределения. Итак, эмпирической функцией распределения $F_n(x)$ называется доля элементов выборки, меньших x. Эмпирическая функция распределения содержит всю информацию о результатах наблюдений. Эмпирическая функция распределения — это функция введенного выше эмпирического распределения.

Чтобы выразить эмпирическую функцию распределения не словами, а формулой, введем функцию $c(x, y)$ двух переменных:

$$c(x,y)=\begin{cases}0, & x\le y,\\ 1, & x>y.\end{cases}$$

Случайные величины, моделирующие результаты наблюдений, обозначим $X_1(\omega), X_2(\omega), ..., X_n(\omega), \omega\in\Omega$. Тогда эмпирическая функция распределения $F_n(x)$ имеет вид:

$$F_n(x)=F_n(x,\omega)=\frac{1}{n}\sum_{1\le i\le n}c(x,X_i(\omega)).$$

Из закона больших чисел следует, что для каждого действительного числа x эмпирическая функция распределения $F_n(x)$ сходится к функции распределения $F(x)$ результатов наблюдений, т.е.:

$$F_n(x)\to F(x) \tag{1}$$

при $n\to\infty$. Советский математик В.И. Гливенко (1897–1940) доказал в 1933 г. более сильное утверждение: сходимость в (1) равномерна по x, т.е.:

$$\sup_x|F_n(x)-F(x)|\to 0 \tag{2}$$

при $n\to\infty$ (сходимость по вероятности).

В (2) использовано обозначение sup (читается как «супремум»). Для функции $g(x)$ под $\sup_x g(x)$ понимают наименьшее из чисел a таких, что $g(x)\le a$ при всех x. Если функция $g(x)$ достигает максимума в точке x_0, то $\sup_x g(x)=g(x_0)$. В таком случае вместо sup пишут max. Хорошо известно, что не все функции достигают максимума.

В том же 1933 г. А.Н. Колмогоров усилил результат В.И. Гливенко для непрерывных функций распределения $F(x)$. Рассмотрим случайную величину:

$$D_n=\sqrt{n}\sup_x|F_n(x)-F_0(x)|$$

и ее функцию распределения:

$$K_n(x) = P\{D_n \le x\}.$$

По теореме А.Н. Колмогорова:

$$\lim_{n\to\infty} K_n(x) = K(x)$$

при каждом x, где $K(x)$ — т.н. функция распределения Колмогорова.

Рассматриваемая работа А.Н. Колмогорова породила одно из основных направлений математической статистики — т.н. непараметрическую статистику. И в настоящее время непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат широко используются. Они были разработаны для проверки согласия с *полностью известным* теоретическим распределением, т.е. предназначены для проверки гипотезы $H_0 : F(x) \equiv F_0(x)$. Основная идея критериев Колмогорова, омега-квадрат и аналогичных им состоит в измерении расстояния между функцией эмпирического распределения и функцией теоретического распределения. Различаются эти критерии видом расстояний в пространстве функций распределения. Аналитические выражения для предельных распределений статистик, расчетные формулы, таблицы распределений и критических значений широко распространены [8], поэтому не будем их приводить.

Непараметрическое оценивание функции распределения. По теореме Гливенко эмпирическая функция распределения $F_n(x)$ является состоятельной оценкой функции распределения $F(x)$. Если $F(x)$ — непрерывная функция, то на основе теоремы Колмогорова доверительные границы для функции распределения $F(x)$ задают в виде:

$$(F(x))_{\text{Н}} = \max\left\{0, F_n(x) - \frac{k(\gamma,n)}{\sqrt{n}}\right\},\ (F(x))_{\text{В}} = \min\left\{1, F_n(x) + \frac{k(\gamma,n)}{\sqrt{n}}\right\},$$

где $k(\gamma,n)$ — квантиль порядка γ распределения статистики Колмогорова при объеме выборки n (напомним, что распределение этой статистики не зависит от $F(x)$).

Как уже отмечалось, правила определения оценок и доверительных границ иногда строятся на основе параметрического семейства распределений $F(x;\theta)$. При обработке реальных данных возникает вопрос — соответствуют ли эти данные принятой вероятностной модели? Статистической гипотезе о том, что результаты наблюдений имеют функцию распределения из семейства $\{F(x;\theta), \theta \in \Theta\}$ при некотором $\theta = \theta_0$? Такие гипотезы называют гипотезами согласия, а критерии их проверки — критериями согласия.

Если истинное значение параметра $\theta = \theta_0$ известно, функция распределения $F(x;\theta_0)$ непрерывна, то для проверки гипотезы согласия часто применяют критерий Колмогорова, основанный на статистике:

$$D_n = \sqrt{n} \sup_x | F_n(x) - F(x, \theta_0 |,$$

где $F_n(x)$ — эмпирическая функция распределения.

Если истинное значение параметра θ_0 неизвестно, например, при проверке гипотезы о нормальности распределения результатов наблюдения (т.е. при проверке принадлежности этого распределения к семейству нормальных распределений), то иногда используют статистику типа Колмогорова:

$$D_n(\theta^*) = \sqrt{n} \sup_x | F_n(x) - F(x, \theta^*) | .$$

Она отличается от статистики Колмогорова D_n тем, что вместо истинного значения параметра θ_0 подставлена его оценка θ^*.

Распределение статистики $D_n(\theta^*)$ сильно отличается от распределения статистики D_n. В качестве примера рассмотрим проверку нормальности, когда $\theta = (m, \sigma^2)$, а $\theta^* = (\bar{x}, s^2)$. Для этого случая квантили распределений статистик D_n и $D_n(\theta^*)$ приведены в табл. 4 (см., например, [9]). Таким образом, квантили отличаются примерно в 1,5 раза.

Таблица 4

Квантили статистик D_n и $D_n(\theta^*)$ при проверке нормальности

p	0,85	0,90	0,95	0,975	0,99
Квантили порядка p для D_n	1,138	1,224	1,358	1,480	1,626
Квантили порядка p для $D_n(\theta^*)$	0,775	0,819	0,895	0,955	1,035

2.5. Непараметрические оценки плотности

Эмпирическая функция распределения — это состоятельная непараметрическая оценка функции распределения числовой случайной величины. А как оценить плотность? Если продифференцировать эмпирическую функцию распределения, то получим бесконечности в точках, соответствующих элементам выборки, и 0 во всех остальных. Ясно, что это не оценка плотности.

Как же действовать? Каждому элементу выборки соответствует в эмпирическом распределении вероятность $1/n$, где n — объем выборки. Целесообразно эту вероятность не помещать в одну точку, а «размазать» вокруг нее, построив «холмик». Если «холмики» налегают друг на друга, то получаем положительную плотность на всей прямой. Чтобы получить состоятельную оценку плотности, необходимо выбирать ширину «холмика» в зависимости от объема выборки. При этом число «холмиков», покрывающих фиксированную точку, должно безгранично расти. Но одновременно доле таких «холмиков» следует убывать, поскольку покрывающие «холмики» должны быть порождены лишь ближайшими членами вариационного ряда.

Реализация описанной идеи привела к различным вариантам непараметрических оценок плотности. Основополагающей является работа Н.В. Смирнова 1951 г. [10]. Вначале рассматривались непараметрические оценки плотности распределения числовых случайных величин и конечномерных случайных векторов. В 1980-х годах удалось сконструировать такие оценки в пространствах произвольной природы [11], а затем и для конкретных видов нечисловых данных [12].

Чтобы определить понятие плотности, в пространстве X должна быть выделена некоторая специальная мера μ, относительно которой будут рассматриваться плотности, соответствующие другим мерам, например, мере ν, задающей распределение вероятностей некоторого случайного элемента ξ. В таком случае $\nu(A) = \mathrm{P}(\xi \in A)$ для любого случайного события A. Плотность $f(x)$, соответствующая мере ν — это такая функция, что:

$$\nu(A) = \int_A f(x)d\mu$$

для любого случайного события A. Для случайных величин и векторов мера μ — это объем множества A, в математических терминах — мера Лебега.

Как могут быть использованы непараметрические оценки плотности распределения вероятностей в пространствах нечисловой природы? Например, для решения задач классификации (диагностики, распознавания образов — см. главу 6.4). Зная плотности распределения классов, можно решать основные задачи диагностики — как задачи выделения кластеров, так и задачи отнесения вновь поступающего объекта к одному из диагностических классов. В задачах кластер-анализа можно находить моды плотности и принимать их за центры кластеров или за начальные точки итерационных методов типа k-средних или динамических сгущений. В задачах собственно диагностики (дискриминации, распознавания образов с учителем) можно принимать решения о диагностике объектов на основе отношения плотностей, соответствующих классам. При неизвестных плотностях представляется естественным использовать их состоятельные оценки.

Методы оценивания плотности вероятности в пространствах общего вида предложены и первоначально изучены в работе [11]. В частности, в задачах диагностики и кластер-анализа целесообразно использовать непараметрические ядерные оценки плотности типа Парзена — Розенблатта (этот вид оценок и его название впервые были введены в статье [11]). Они имеют вид:

$$f_n(x) = \frac{1}{\eta_n(h_n, x)} \sum_{1 \le i \le n} K\left(\frac{d(x_i, x)}{h_n}\right),$$

где K: $R^1_+ \to R^1$ — так называемая ядерная функция, $x_1, x_2, \ldots, x_n \in X$ — выборка, по которой оценивается плотность, $d(x_i, x)$ — показатель различия (метрика, расстояние, мера близости) между элементом выборки x_i и точкой x, в которой оценивается плотность, последовательность h_n показателей размытости такова, что $h_n \to 0$ и $nh_n \to \infty$ при $n \to \infty$, а $\eta_n(h_n, x)$ — нормирующий множитель, обеспечивающий выполнение условия нормировки (интеграл по всему пространству от непараметрической оценки плотности $f_n(x)$ по мере μ должен равняться 1). Ранее американские исследователи Парзен и Розенблатт использовали подобные статистики в случае $X = R^1$ с $d(x_i, x) = |x_i - x|$.

В качестве ядерной функции распределения можно взять функцию, равную 1/2 на отрезке [–1; 1] и 0 вне этого отрезка. Или плотность стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Показателем различия $d(x_i, x)$ может служить обычное евклидово расстояние. Показатель размытости $h_n = n^{-\alpha}$ при любом α из интервала (0; 1) удовлетворяет введенным выше условиям.

Введенные описанным выше образом ядерные оценки плотности — частный случай так называемых линейных оценок, также впервые предложенных в работе [11]. В теоретическом плане они выделяются тем, что удается получать результаты такого же типа, что в классическом одномерном случае, но с помощью совсем иного математического аппарата.

Свойства непараметрических ядерных оценок плотности. Рассмотрим выборку со значениями в некотором пространстве, в котором заданы показатель различия d и мера μ. Одна из основных идей рассматриваемого подхода состоит в том, чтобы согласовать их между собой. А именно, на их основе построим новый показатель различия d_1, так называемый «естественный», в терминах которого проще формулируются свойства непараметрической оценки плотности. Для этого рассмотрим шары $L_t(x) = \{y \in X : d(y,x) \le t\}$ радиуса $t \ge 0$ и их меры $F_x(t) = \mu(L_t(x))$. Предположим, что $F_x(t)$ как функция t при фиксированной точке x непрерывна и строго возрастает. Введем функцию $d_1(x,y) = F_x(d(x,y))$. Это — монотонное преобразование показателя различия или расстояния, а потому $d_1(x,y)$ — также показатель различия (даже если d — *метрика,* для d_1 неравенство треугольника может быть не выполнено). Другими словами, $d_1(x,y)$, как и $d(x,y)$, можно рассматривать как показатель различия (меру близости) между x и y.

Для вновь введенного показателя различия $d_1(x,y)$ введем соответствующие шары $L_{1t}(x) = \{y \in X : d_1(y,x) \le t\}$. Поскольку обратная функция $F^{-1}{}_x(t)$ определена однозначно, то:

$$L_{1t}(x) = \{y \in X : d_1(y,x) \le F_x^{-1}(t)\} = L_T(x)\ ,$$

где $T = F^{-1}{}_x(t)$. Следовательно, справедлива цепочка равенств $F^1{}_x(t) = \mu(L_{1t}(x)) = \mu(L_T(x)) = F_x(F^{-1}{}_x(t)) = t$ (для всех тех значений параметра t, для которых определены все участвующие в записи математические объекты).

Переход от d к d_1 напоминает классическое преобразование, использованное Н.В. Смирновым при изучении непараметрических критериев согласия и однородности, а именно, преобразование $\eta = F(\xi)$, переводящее случайную величину ξ с непрерывной функцией распределения $F(x)$ в случайную величину η, равномерно распределенную на отрезке [0; 1]. Оба рассматриваемых преобразования существенно упрощают дальнейшие рассмотрения. Преобразование $d_1 = F_x(d)$ зависит от точки x, что не влияет на дальнейшие рассуждения, поскольку ограничиваемся изучением сходимости в отдельно взятой точке.

Функцию $d_1(x,y)$, для которой мера шара радиуса t равна t, называем в соответствии с работой [11] «естественным показателем различия» или «естественной метрикой». В случае конечномерного пространства R^k и евклидовой метрики d имеем $d_1(x,y) = c_k\, d^k(x,y)$, где c_k — объем шара единичного радиуса в R^k.

Поскольку можно записать, что:

$$K\left(\frac{d(x_i,x)}{h_n}\right)=K_1\left(\frac{d_1(x_i,x)}{h_n}\right),$$

где:

$$K_1(u)=K\left(\frac{F_x^{-1}(uh_n)}{h_n}\right),$$

то переход от одного показателя различия к другому, т.е. от d к d_1, соответствует переходу от одной ядерной функции к другой, т.е. от K к K_1. Выгода от такого перехода заключается в том, что утверждения о поведении непараметрических оценок плотности приобретают более простую формулировку.

Теорема 1. Пусть d — естественная метрика, плотность f непрерывна в точке x и ограничена на всем пространстве X, причем $f(x)>0$, ядерная функция $K(u)$ удовлетворяет простым условиям регулярности:

$$\int_0^{+\infty}K(u)du=1,\ \int_0^{+\infty}(|K(u)|+K^2(u))du<\infty\cdot$$

Тогда $\eta_n(h_n,x)=nh_n$, оценка $f_n(x)$ является состоятельной, т.е. $f_n(x)\to f(x)$ по вероятности при $n\to\infty$ и, кроме того,

$$\lim_{n\to\infty}(nh_nDf_n(x))=f(x)\int_0^{+\infty}K^2(u)du.$$

Теорема 1 доказывается методами, развитыми в работе [11]. Однако остается открытым вопрос о скорости сходимости ядерных оценок, в частности, о поведении величины $\alpha_n=M(f_n(x)-f(x))^2$ — среднего квадрата ошибки, и об оптимальном выборе показателей размытости h_n. Для того, чтобы продвинуться в решении этого вопроса, введем новые понятия. Для случайного элемента $X(\omega)$ со значениями в X рассмотрим т.н. круговое распределение $G(x,t)=$ $=P\{d(X(\omega),x)\leq t\}$ и круговую плотность $g(x,t)=G'_t(x,t)$.

Теорема 2. Пусть ядерная функция $K(u)$ непрерывна и финитна, т.е. существует число E такое, что $K(u)=0$ *при* $u>E$. Пусть круговая плотность является достаточно гладкой, т.е. допускает разложение:

$$g(x,t)=f(x)+tg'_t(x,0)+\frac{t^2}{2}g''_{tt}(x,0)+\frac{t^3}{3!}g'''_{ttt}(x,0)+\ldots+\frac{t^k}{k!}g^{(k)}_{t^{(k)}}(x,0)+o(h_n^k)$$

при некотором k, причем остаточный член равномерно ограничен на $[0,hE]$. Пусть:

$$\int_0^E u^i K(u)du = 0, i = 1,2,\ldots,k-1.$$

Тогда:

$$\alpha_n = [Mf_n(x) - f(x)]^2 + Df_n(x) =$$
$$= h_n^{2k}\left(\int_0^E u^k K(u)du\right)^2 (g^k_{t^{(k)}}(x,0))^2 + \frac{f(x)}{nh_n}\int_0^E K^2(u)du + o\left(h_n^{2k} + \frac{1}{nh_n}\right).$$

Доказательство теоремы 2 проводится с помощью разработанной в нечисловой статистике математической техники, образцы которой представлены, в частности, в работе [11]. Если коэффициенты при основных членах в правой части последней формулы не равны 0, то величина α_n достигает минимума, равного $\alpha_n = O\left(n^{-1+\frac{1}{2k+1}}\right)$, при $h_n = n^{-\frac{1}{2k+1}}$. Эти выводы совпадают с классическими результатами, полученными ранее рядом авторов для весьма частного случая прямой $X = R^1$ (см., например, монографию [13, с.316]). Заметим, что для уменьшения смещения оценки приходится применять знакопеременные ядра $K(u)$.

Непараметрические оценки плотности лучше, чем гистограммы, отражают эмпирическое распределение. Однако для расчета таких оценок необходимо использование вычислительной техники. Программная реализация описания статистических данных с помощью непараметрических оценок плотности включена в ряд программных продуктов по статистическим методам. Примерами являются диалоговая система анализа данных ДИСАН и пакет программ анализа данных ППАНД [14], разработанные под научным руководством автора настоящего учебника.

Литература

1. *Новицкий, П.В.* Оценка погрешностей результатов измерений / П.В. Новицкий, И.А. Зограф. — Ленинград : Энергоатомиздат, 1985. — 248 с.

2. *Новицкий, П.В.* Основы информационной теории измерительных устройств / П.В. Новицкий. — Ленинград : Энергия, 1968. — 248 с.

3. *Боровков, А.А.* Теория вероятностей / А.А. Боровков. — Москва : Наука, 1976. — 352 с.

4. *Петров, В.В.* Суммы независимых случайных величин / В.В. Петров. — Москва : Наука, 1972. — 416 с.

5. *Золотарев, В.М.* Современная теория суммирования независимых случайных величин / В.М. Золотарев. — Москва : Наука, 1986. — 416 с.

6. *Егорова, Л.А.* О применении непараметрического Х-критерия Ван-дер-Вардена при статистической обработке результатов наблюдений / Л.А. Егорова, Ю.С. Харитонов, Л.В. Соколовская // Заводская лаборатория. — 1976. — Т. 42. — № 11. — С. 1237–1237.

7. *Орлов, А.И.* Эконометрика : учебник / А.И. Орлов. — 2-е изд. — Москва : Экзамен, 2003. — 576 с.

8. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.).

9. *Орлов, А.И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат / А.И. Орлов // Заводская лаборатория. — 1985. — Т. 51. — № 1. — С. 60–62.

10. *Смирнов, Н.В.* О приближении плотностей распределения случайных величин / Н.В. Смирнов // Ученые записки МГПИ им. В.П. Потемкина. — 1951. — Т. XVI. — Вып. 3. — С. 69–96.

11. *Орлов, А.И.* Непараметрические оценки плотности в топологических пространствах / А.И. Орлов // Прикладная статистика. Ученые записки по статистике. — Т. 45. — Москва : Наука, 1983. — С. 12–40.

12. *Орлов, А.И.* Ядерные оценки плотности в пространствах произвольной природы / А.И. Орлов // Статистические методы оценивания и проверки гипотез : Межвузовский сборник научных трудов. — Пермь : Изд-во Пермского госуниверситета, 1996. С. 68–75.

13. *Ибрагимов, И.А.* Асимптотическая теория оценивания / И.А. Ибрагимов, Р.З. Хасьминский. — Москва : Наука, 1979. — 528 с.

14. Пакет программ анализа данных «ППАНД» : учебное пособие / А.И. Орлов, И.Л. Легостаева [и др.]. — Москва : Сотрудничающий центр ВОЗ по профессиональной гигиене, 1990. — 93 с.

15. *Орлов, А.И.* Распределения реальных статистических данных не являются нормальными / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 117. — С. 71–90.

16. *Орлов, А.И.* Непараметрическое оценивание характеристик распределений вероятностей / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 112. — С. 1–20.

17. *Орлов, А.И.* Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат и ошибки при их применении / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 97. — С. 32–45.

18. *Орлов, А.И.* Оценки плотности распределения вероятностей в пространствах произвольной природы / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 99. — С. 15–32.

19. *Орлов, А.И.* Предельные теоремы для ядерных оценок плотности в пространствах произвольной природы / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 108. — С. 316–333.

20. *Орлов, А.И.* Непараметрические ядерные оценки плотности вероятности в дискретных пространствах / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 122. — С. 833–855.

21. *Орлов, А.И.* Асимптотика оценок плотности распределения вероятностей / А.И. Орлов // Научный журнал КубГАУ. — 2017. — № 131. — С. 845–873.

Контрольные вопросы и задачи

1. Часто ли результаты измерений имеют нормальное распределение?

2. По выборке фактических данных о величине годового дохода (в тыс. долл.), взятых на конец 1970-х гг. (США), постройте вариационный ряд, гистограмму (группируя данные по шести равным интервалам); определить выборочные среднее арифметическое, медиану и моду:

2,0; 13,4; 2,2; 6,7; 11,1; 10,0; 2,6; 12,9; 10,5; 9,2; 11,1;
14,0; 26,0; 17,5; 7,2; 18,7; 9,9; 7,6; 11,7;11,3; 6,5.

3. Дано распределение по градациям (табл. 5) почасовой зарплаты 303 рабочих, занятых в промышленности (f_i — число рабочих, имеющих почасовую зарплату x_i). Постройте эмпирическую функцию распределения, найти выборочные медиану, моду и среднее арифметическое.

Таблица 5

Распределение рабочих по ставкам почасовой оплаты

x_i	2,5	2,6	2,7	2,8	2,9	3,0	3,1	3,2	3,3	3,4
f_i	10	25	41	74	58	34	17	14	11	3

4. Как связаны эмпирическое распределение и эмпирическая функция распределения?

5. Можно ли проверять нормальность распределения с помощью критерия согласия Колмогорова?

6. Приведите конкретные примеры непараметрических оценок плотности.

7. Почему описание числовых данных с помощью непараметрических оценок плотности предпочтительнее их описания с помощью гистограмм?

Темы докладов, рефератов, исследовательских работ

1. Сравнение непараметрических и параметрических подходов к анализу статистических данных.

2. Различные виды статистических таблиц.

3. Методы наглядного представления статистических данных.

4. Система выборочных характеристик распределения.

5. Проведите описание данных, приведенных в табл. 1 (раздел 2.2). Постройте таблицы (типа табл. 2 и 3 там же), рассчитайте выборочные характеристики.

6. Аналоги эмпирической функции распределения.

7. Методы изучения асимптотических распределений статистик типа Колмогорова.

8. Ядерные непараметрические оценки плотности и оценки, основанные на разложении плотности в ряд по базисным функциям.

9. Непараметрические оценки плотности в непрерывных и дискретных пространствах.

ГЛАВА 3. ОЦЕНИВАНИЕ

3.1. Методы оценивания параметров

Статистические методы иногда опираются на разнообразные параметрические модели. Термин «параметрический» означает, то вероятностно-статистическая модель полностью описывается конечномерным вектором фиксированной размерности. Причем эта размерность не зависит от объема выборки.

Рассмотрим выборку x_1, x_2,..., x_n из распределения с плотностью $f(x;\theta_0)$, где $f(x;\theta_0)$ — элемент параметрического семейства плотностей распределения вероятностей $\{f(x;\theta), \theta \in \Theta\}$. Здесь Θ — заранее известное k-мерное пространство параметров, являющееся подмножеством евклидова пространства R^k, а конкретное значение параметра θ_0 статистику неизвестно. Обычно в прикладной статистике применяются параметрические семейства с $k = 1, 2, 3$. В статистике нечисловых данных вместо плотности часто рассматриваются вероятности попадания в точки. Напомним, что в параметрических задачах оценивания принимают вероятностную модель, согласно которой результаты наблюдений x_1, x_2, ..., x_n рассматривают как реализации n независимых случайных величин.

Задача оценивания состоит в том, чтобы оценить неизвестное статистику значение параметра θ_0 наилучшим (в каком-либо смысле) образом.

Пример 1. В статистических задачах стандартизации и управления качеством используют семейство гамма-распределений. Плотность гамма-распределения имеет вид:

$$f(x;a,b,c)=\begin{cases}\dfrac{1}{\Gamma(a)}(x-c)^{a-1}b^{-a}\exp\left[-\dfrac{x-c}{b}\right], & x\ge c,\\ 0, & x<c.\end{cases} \quad (1)$$

Плотность вероятности в формуле (1) определяется тремя параметрами a, b, c, где a>2, b>0. При этом a является параметром формы, b — параметром масштаба и c — параметром сдвига. Множитель $1/\Gamma(a)$ является нормировочным, он введен, чтобы:

$$\int_{-\infty}^{+\infty} f(x;a,b,c)dx=1.$$

Здесь $\Gamma(a)$ — одна из используемых в математике специальных функций, так называемая «гамма-функция», по которой названо и распределение, задаваемое формулой (1),

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx.$$

Подробные решения задач оценивания параметров для гамма-распределения содержатся в разработанном нами государственном стандарте ГОСТ 11.011-83 «Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения» (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) [1]. В настоящее время эта публикация используется в качестве методического материала для инженерно-технических работников промышленных предприятий и прикладных научно-исследовательских институтов.

Поскольку гамма-распределение зависит от трех параметров, то имеется $2^3 - 1 = 7$ вариантов постановок задач оценивания. Они описаны в табл. 1.

Таблица 1

Постановки задач оценивания для параметров гамма-распределения

№ п/п	Параметр формы	Параметр масштаба	Параметр сдвига
1	Известен	Оценивается	Известен
2	Оценивается	Известен	Известен
3	Известен	Известен	Оценивается
4	Оценивается	Оценивается	Известен
5	Известен	Оценивается	Оценивается
6	Оценивается	Известен	Оценивается
7	Оценивается	Оценивается	Оценивается

В табл. 2 приведены реальные данные о наработке резцов до предельного состояния, в часах. Упорядоченная выборка (вариационный ряд) объема $n = 50$ взята из государственного стандарта (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) [1]. Проверка согласия данных о наработке резцов с семейством гамма-распределений проведена в главе 4.1. Именно эти данные будут служить исходным материалом для демонстрации тех или иных методов оценивания параметров.

Таблица 2

Наработка резцов до предельного состояния, ч

№ п/п	Наработка, ч	№ п/п	Наработка, ч	№ п/п	Наработка, ч
1	9	18	47,5	35	63
2	17,5	19	48	36	64,5
3	21	20	50	37	65
4	26,5	21	51	38	67,5
5	27,5	22	53,5	39	68,5
6	31	23	55	40	70
7	32,5	24	56	41	72,5
8	34	25	56	42	77,5
9	36	26	56,5	43	81
10	36,5	27	57,5	44	82,5
11	39	28	58	45	90
12	40	29	59	46	96
13	41	30	59	47	101,5
14	42,5	31	60	48	117,5
15	43	32	61	49	127,5
16	45	33	61,5	50	130
17	46	34	62	–	–

Выбор «наилучших» оценок в определенной параметрической модели прикладной статистики — научно-исследовательская работа, растянутая во времени. Выделим два этапа. *Этап асимптотики*: оценки строятся и сравниваются по их свойствам при безграничном росте объема выборки. На этом этапе рассматривают такие характеристики оценок, как состоятельность, асимптотическая эффективность и др. *Этап конечных объемов выборки:* оценки сравниваются, скажем, при n = 10. Ясно, что исследование начинается с этапа асимптотики: чтобы сравнивать оценки, надо сначала их построить и быть уверенными, что они не являются абсурдными (такую уверенность дает доказательство состоятельности).

С какой оценки начинать? Одним из наиболее известных и простых в употреблении методов является метод моментов. Название связано с тем, что этот метод опирается на использование выборочных моментов:

$$M_{nm} = \frac{1}{n}\sum_{i=1}^{n} x_i^m, \quad m = 1, 2, \ldots,$$

где x_1, x_2,..., x_n — выборка, т.е. набор независимых одинаково распределенных случайных величин с числовыми значениями.

В прикладной статистике метод анализа данных называется *методом моментов*, если он использует статистику:

$$Y_n = g(M_{n1}, M_{n2}, ... M_{nq}), \tag{2}$$

где g: $R^q \to R^k$ — некоторая функция (здесь k — число неизвестных числовых параметров). Чаще всего термин «метод моментов» используют, когда речь идет об оценивании параметров. В этом случае обычно предполагают, что плотность вероятности распределения элементов выборки $f(x)$ входит в заранее известное статистику параметрическое семейство $\{f(x;\theta), \theta \in \Theta\}$, т.е. $f(x) = f(x;\theta_0)$ при некотором θ_0. Здесь Θ — заранее заданное k-мерное пространство параметров, являющееся подмножеством евклидова пространства R^k, а конкретное значение параметра θ_0 статистику неизвестно, его и следует оценить. Известно также, что неизвестный параметр определяется с помощью известной статистику функции через начальные моменты элементов выборки:

$$\theta_0 = g(a_1, a_2, ..., a_q), \quad a_m = M(x_i^m), \quad m = 1,2,... \tag{3}$$

В методе моментов в качестве оценки θ_0 используют статистику Y_n вида (2), которая отличается от правой части формулы (2) заменой теоретических моментов заменены выборочными.

Статистики Y_n вида (2) применяются не только для оценивания параметров, но и для непараметрического оценивания характеристик случайной величины, таких, как коэффициент вариации, и для проверки гипотез. Во всех случаях применения статистики Y_n вида (2) говорят о методе моментов.

Распределение вектора Y_n во всех практически важных случаях является асимптотически нормальным. Это утверждение опирается на следующий общий факт.

Пусть случайный вектор $Z_n \in R^q$ асимптотически нормален с математическим ожиданием z_∞ и ковариационной матрицей $||c_{ij}||/n$, а функция h: $R^q \to R^1$ достаточно гладкая. Тогда случайная величина $h(Z_n)$ асимптотически нормальна с математическим ожиданием $h(z_\infty)$ и дисперсией:

$$\sigma^2 = \frac{1}{n}\sum_{r=1}^{q}\sum_{s=1}^{q}\frac{\partial h}{\partial x_r}\frac{\partial h}{\partial x_s}c_{rs}. \tag{4}$$

Этот способ нахождения предельного распределения известен как δ-метод Рао [2], метод линеаризации [3]. Последний термин и будем использовать. Условия регулярности, накладываемые на распределение случайной величины Z_n и функцию h, при которых метод линеаризации обоснован, хорошо известны (см. [4], [2, с.337–339], а также главу 14 настоящего учебника).

Для получения асимптотического распределения статистики Y_n вида (2) можно применить метод линеаризации к асимптотически нормальному вектору выборочных моментов ($M_{n1}, M_{n2}, \ldots, M_{nq}$) и функции g из формулы (2).

В силу многомерной центральной предельной теоремы (см. главу 14) указанная асимптотическая нормальность имеет место, если, например,

$$M|x_i|^{2q+1} < +\infty.$$

Это условие выполнено, в частности, для результатов измерений, распределения которых сосредоточены на ограниченных сверху и снизу интервалах.

При реализации намеченного плана для применения формулы (4) необходимо использовать асимптотические дисперсии и ковариации выборочных моментов, т.е. величины, обозначенные в формуле (4) как c_{rs}. Эти величины имеют вид [2, с. 388]:

$$c_{rr} = \mu_{2r} - \mu_r^2 - 2r\mu_{r-1}\mu_{r+1} + r^2\mu_{r-1}^2\mu_2,$$
$$c_{rs} = \mu_{r+s} - \mu_r\mu_s + rs\mu_2\mu_{r-1}\mu_{s-1} - r\mu_{r-1}\mu_{s+1} - s\mu_{r+1}\mu_{s-1}, \quad r,s = 1,2,\ldots, \quad \mu_0 = 0. \qquad (5)$$

Здесь μ_r — теоретический центральный момент порядка r, т.е.:

$$\mu_r = M(x_i - M(x_i))^r, \quad r = 1,2,\ldots$$

Таким образом, для получения асимптотического распределения случайной величины Y_n вида (2) достаточно знать теоретические центральные моменты результатов наблюдений и вид функции g. Отметим, что асимптотическим смещением оценок в рассматриваемом случае можно пренебречь, поскольку его вклад в средний квадрат ошибки статистической оценки — бесконечно малая величина более высокого порядка по сравнению с асимптотической дисперсией.

Однако моменты неизвестны. Их приходится оценивать. Согласно теоремам о наследовании сходимости для нахождения асимптотического распределения функции от выборочных моментов можно воспользоваться не теоретическими моментами, а их состоятельными оценками. Эти оценки можно получить разными способами. Можно непосредственно применить формулы (5), заменив теоретические моменты выборочными. Можно выразить моменты через параметры рассматриваемого распределения. Можно применять более сложные процедуры, например, на основе непараметрических устойчивых (робастных) оценок моментов типа урезанных средних Пуанкаре и др. (в первой в России книге по общей теории устойчивости [5] проблематика робастных оценок рассмотрена в главе 2).

Для оценивания параметров гамма-распределения воспользуемся известной формулой [6, с. 184–185], согласно которой для случайной величины X, имеющей гамма-распределение с параметрами формы a, масштаба b =1 и сдвига с=0,

$$M(X^m)=\frac{\Gamma(a+m)}{\Gamma(a)}=a(a+1)...(a+m-1),\quad m=1,2,... \tag{6}$$

Следовательно, $M(X) = a$, $M(X^2) = a(a+1)$, $D(X) = M(X^2) - (M(X))^2 = a(a+1) - - a^2 = a$. Найдем третий центральный момент $M(X - M(X))^3$. Справедливо равенство:

$$M(X - M(X))^3 = M(X^3) - 3\,M(X^2)\,M(X) + 3\,M(X)\,(M(X))^2 - (M(X))^3.$$

Из равенства (6) вытекает, что:

$$M(X - M(X))^3 = a(a + 1)(a + 2) - 3\,a\,(a + 1)\,a + 3\,a\,a^2 - a^3 = 2a.$$

Если Y — случайная величина, имеющая гамма-распределение с произвольными параметрами формы a, масштаба b и сдвига c, то $Y = bX + c$. Следовательно, $M(Y) = ab+c$, $D(Y) = ab^2$, $M(Y - M(Y))^3 = 2ab^3$.

Пример 2. Оценивание методом моментов параметров гамма-распределения в случае трех неизвестных параметров (строка 7 табл. 1).

В соответствии с проведенными выше рассуждениями для оценивания трех параметров достаточно использовать три выборочных момента — выборочное среднее арифметическое:

$$\bar{x}=\frac{x_1+x_2+...+x_n}{n},$$

выборочную дисперсию:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

и выборочный третий центральный момент:

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3.$$

Приравнивая теоретические моменты, выраженные через параметры распределения, и выборочные моменты, получаем систему уравнений метода моментов:

$$ab + c = \bar{x}, \quad ab^2 = s^2, \quad 2ab^3 = m_3.$$

Решая эту систему, находим оценки метода моментов. Подставляя второе уравнение в третье, получаем оценку метода моментов для параметра сдвига:

$$2s^2 b = m_3, \quad b^* = \frac{1}{2}\frac{m_3}{s^2}.$$

Подставляя эту оценку во второе уравнение, находим оценку метода моментов для параметра формы:

$$a(b^*)^2 = a\left(\frac{1}{2}\frac{m_3}{s^2}\right)^2 = \frac{a}{4}\frac{m_3^2}{s^4} = s^2, \quad a^* = 4\frac{s^6}{m_3^2}.$$

Наконец, из первого уравнения находим оценку для параметра сдвига:

$$c^* = \bar{x} - a^* b^* = \bar{x} - 4\frac{s^6}{m_3^2}\frac{1}{2}\frac{m_3}{s^2} = \bar{x} - 2\frac{s^4}{m_3}.$$

Для реальных данных [1], приведенных выше в табл. 2, выборочное среднее арифметическое $\bar{x}$ = 57,88, выборочная дисперсия s^2 = 663,00, выборочный третий центральный момент m_3 = 14 927,91. Согласно только что полученным формулам оценки метода моментов таковы:

a^* = 5,23; b^* = 11,26, c^* = – 1,01.

Оценки параметров гамма-распределения, полученные методом моментов, являются функциями от выборочных моментов. В соответствии со сказанным выше они являются асимптотически нормальными случайными величинами. Их распределения аппроксимируются нормальными распределениями, математические ожидания которых равны соответствующим параметрам, а дисперсии находятся с помощью формулы (4) с учетом формул (5) и (6). В табл. 3 приведены оценки метода моментов и их асимптотические дисперсии при различных вариантах сочетания известных и неизвестных параметров гамма-распределения.

Таблица 3

Оценки метода моментов и их асимптотические дисперсии

№ п/п	Описание вероятностной модели			Оцениваемый параметр	Вид оценки	Асимптотическая дисперсия оценки
	a	***b***	***c***			
1	–	–	+	a	$\frac{\left(\bar{x}\right)^2}{s^2}$	$\frac{2a(a+1)}{n}$
2	–	–	+	b	$\frac{s^2}{\bar{x}}$	$\frac{b^2}{n}\left(2+\frac{3}{a}\right)$
3	–	–	–	a	$4\frac{s^6}{m_3^2}$	$\frac{6a}{n}(a^2+6a+5)$
4	–	–	–	b	$\frac{1}{2}\frac{m_3}{s^2}$	$\frac{b^2}{2an}(6a^2+25a+24)$
5	–	–	–	c	$\bar{x}-2\frac{s^4}{m_3}$	$\frac{ab^2}{n}(3a^2+13a+10)$
6	+	–	–	b	$\frac{s}{\sqrt{a}}$	$\frac{b}{2n}(a+3)$
7	+	–	–	c	$\bar{x}-s\sqrt{a}$	$\frac{ab^2}{2n}(a+1)$
8	–	+	–	A	$\frac{s^2}{b^2}$	$\frac{2a}{n}(a+3)$
9	–	+	–	c	$\bar{x}-\frac{s^2}{b}$	$\frac{ab^2}{n}(2a+3)$
10	+	+	–	c	$\bar{x}-ab$	$\frac{ab^2}{n}$

Примечание. При описании вероятностной модели известные статистику параметры отмечены плюсами, оцениваемые — минусами.

Все оценки метода моментов, приведенные в табл. 3, включены в государственный стандарт [1]. Они охватывают все постановки задач оценивания параметров гамма-распределения (см. табл. 1), кроме тех, когда неизвестен только один параметр — a или b. Для этих исключительных случаев в [1] разработаны специальные методы оценивания.

Поскольку асимптотическое распределение оценок метода моментов известно, то не представляет труда формулировка правил проверки статистических гипотез относительно значений параметров распределений, а также построение доверительных границ для параметров. Например, в вероятностной модели, когда все три параметра неизвестны, в соответствии с третьей строкой таблицы 3 нижняя доверительная граница для параметра a, соответствующая доверительной вероятности $\gamma = 0{,}95$, в асимптотике имеет вид:

$$a_H = a^* - 1{,}96\left\{\frac{6a^*}{n}([a^*]^2 + 6a^* + 5)\right\}^{1/2},$$

а верхняя доверительная граница для той же доверительной вероятности такова:

$$a_B = a^* + 1{,}96\left\{\frac{6a^*}{n}([a^*]^2 + 6a^* + 5)\right\}^{1/2},$$

где a^* — оценка метода моментов параметра формы (табл. 3).

Метод моментов является универсальным. Однако получаемые с его помощью оценки лишь в редких случаях обладают оптимальными свойствами. Поэтому в прикладной статистике применяют и другие виды оценок.

В работах, предназначенных для первоначального знакомства с математической статистикой, обычно рассматривают оценки максимального правдоподобия (сокращенно ОМП):

$$\theta_0(n) = \theta_0(n; x_1, x_2, \ldots, x_n) = \underset{\theta\in\Theta}{Arg\min}\prod_{i=1}^{n} f(x_i; \theta). \qquad (7)$$

Таким образом, сначала строится плотность распределения вероятностей, соответствующая выборке. Поскольку элементы выборки независимы, то эта плотность представляется в виде произведения плотностей для отдельных элементов выборки. Совместная плотность рассматривается в точке, соответствующей наблюденным значениям. Это выражение как функция от параметра (при заданных элементах выборки) называется функцией правдоподобия. Затем тем или иным способом ищется значение параметра, при котором значение совместной плотности максимально. Это и есть оценка максимального правдоподобия.

Хорошо известно, что оценки максимального правдоподобия входят в класс наилучших асимптотически нормальных оценок (строгое определение дано ниже). Однако при конечных объемах выборки в ряде задач ОМП недопустимы, т.к. они хуже (дисперсия и средний квадрат ошибки больше), чем другие оценки, в частности, несмещенные [6]. Именно поэтому в ГОСТ 11.010-81 (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) для оценивания параметров отрицательного биномиального распределения используются несмещенные оценки, а не ОМП [7]. Из сказанного следует: априорно предпочитать ОМП другим видам оценок можно — если можно — лишь на этапе изучения асимптотического поведения оценок.

В отдельных случаях ОМП находятся явно, в виде конкретных формул, пригодных для вычисления.

Пример 3. Найдем ОМП для выборки из нормального распределения, каждый элемент которой имеет плотность:

$$f(x;m,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}.$$

Таким образом, надо оценить двумерный параметр (m, σ^2).

Произведение плотностей вероятностей для элементов выборки, т.е. функция правдоподобия, имеет вид:

$$H(m;\sigma^2)=\sigma^{-n}(2\pi)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-m)^2\right\}. \qquad (8)$$

Требуется решить задачу оптимизации:

$$H(m;\sigma^2)\to\max.$$

Как и во многих иных случаях, задача оптимизации проще решается, если прологарифмировать функцию правдоподобия, т.е. перейти к функции:

$$h(m;\sigma^2)=\ln H(m;\sigma^2),$$

называемой логарифмической функцией правдоподобия. Для выборки из нормального распределения:

$$h(m;\sigma^2) = (-n)\ln\sigma + \left(-\frac{n}{2}\right)\ln(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2. \tag{9}$$

Необходимым условием максимума является равенство 0 частных производных от логарифмической функции правдоподобия по параметрам, т.е.:

$$\frac{\partial h(m,\sigma^2)}{\partial m} = 0, \quad \frac{\partial h(m,\sigma^2)}{\partial(\sigma^2)} = 0. \tag{10}$$

Система (10) называется системой уравнений максимального правдоподобия. В общем случае число уравнений равно числу неизвестных параметров, а каждое из уравнений выписывается путем приравнивания 0 частной производной логарифмической функции правдоподобия по тому или иному параметру.

При дифференцировании по *m* первые два слагаемых в правой части формулы (9) обращаются в 0, а последнее слагаемое дает уравнение:

$$\frac{\partial}{\partial m}\sum_{i=1}^{n}(x_i - m)^2 = \sum_{i=1}^{n}2(x_i - m)(-1) = 0, \quad \sum_{i=1}^{n}x_i = nm.$$

Следовательно, оценкой *m** максимального правдоподобия параметра *m* является выборочное среднее арифметическое:

$$m^* = \bar{x}.$$

Для нахождения оценки дисперсии необходимо решить уравнение:

$$\frac{\partial}{\partial(\sigma^2)}h(m;\sigma^2) = \frac{\partial}{\partial(\sigma^2)}(-n)\ln\sqrt{\sigma^2} - \frac{\partial}{\partial(\sigma^2)}\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2 = 0.$$

Легко видеть, что:

$$\frac{\partial}{\partial(\sigma^2)}(-n)\ln\sqrt{\sigma^2} = \frac{(-n)}{2\sigma^2}, \quad -\frac{\partial}{\partial(\sigma^2)}\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2 = \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - m)^2.$$

Следовательно, оценкой $(\sigma^2)^*$ максимального правдоподобия для дисперсии σ^2 с учетом найденной ранее оценки для параметра m является выборочная дисперсия:

$$(\sigma^2)^* = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Итак, система уравнений максимального правдоподобия решена аналитически, ОМП для математического ожидания и дисперсии нормального распределения — это выборочное среднее арифметическое и выборочная дисперсия. Отметим, что последняя оценка является смещенной.

Отметим, что в условиях примера 3 оценки метода максимального правдоподобия совпадают с оценками метода моментов. Причем вид оценок метода моментов очевиден и не требует проведения каких-либо рассуждений.

В большинстве случаев аналитических решений не существует, для нахождения ОМП необходимо применять численные методы. Так обстоит дело, например, с выборками из гамма-распределения или распределения Вейбулла — Гнеденко. Во многих работах каким-либо итерационным методом решают систему уравнений максимального правдоподобия ([8] и др.) или впрямую максимизируют функцию правдоподобия типа (8) (см. [9] и др.).

Однако применение численных методов порождает многочисленные проблемы. Сходимость итерационных методов требует обоснования. В ряде примеров функция правдоподобия имеет много локальных максимумов, а потому естественные итерационные процедуры не сходятся [10]. Для данных ВНИИ железнодорожного транспорта, относящихся к усталостным испытаниям стали, уравнение максимального правдоподобия имеет 11 корней [11]. Какой из одиннадцати использовать в качестве оценки параметра?

Как следствие осознания указанных трудностей, стали появляться работы по доказательству сходимости алгоритмов, предназначенных для нахождения оценок максимального правдоподобия для конкретных вероятностных моделей и конкретных алгоритмов. Примером является статья [12].

Однако теоретическое доказательство сходимости итерационного алгоритма — это еще не все. Возникает вопрос об обоснованном выборе момента прекращения вычислений в связи с достижением требуемой точности. В большинстве случаев он не решен.

Но и это не все. Точность вычислений необходимо увязывать с объемом выборки — чем он больше, тем точнее надо находить оценки параметров, в противном случае нельзя говорить о состоятельности метода оценивания. Более того, при увеличении объема выборки необходимо увеличивать и количество используемых в компьютере разрядов, переходить от одинарной точности расчетов к двойной и далее к тройной, четверной и т.д. — опять-таки ради достижения состоятельности оценок.

Таким образом, при отсутствии явных формул для оценок максимального правдоподобия нахождение ОМП натыкается на ряд проблем вычислительного характера. Специалисты по математической статистике позволяют себе игнорировать все эти проблемы, рассуждая об ОМП в теоретическом плане. Однако специалисты по прикладной статистике не может их игнорировать. Отмеченные проблемы ставят под вопрос целесообразность практического использования ОМП.

Нет необходимости абсолютизировать ОМП. Кроме них, существуют другие виды оценок, обладающих хорошими статистическими свойствами. Примером являются одношаговые оценки (ОШ-оценки), обсуждаемые ниже.

В прикладной статистике разработано много видов оценок. Упомянем квантильные оценки. Они основаны на идее, аналогичной методу моментов, но только вместо выборочных и теоретических моментов приравниваются выборочные и теоретические квантили. Другая группа оценок базируется на идее минимизации расстояния (показателя различия) между эмпирическими данными и элементом параметрического семейства. В простейшем случае минимизируется евклидово расстояние между эмпирическими и теоретическими гистограммами, а точнее, векторами, составленными из высот столбиков гистограмм.

3.2. Одношаговые оценки

Одношаговые оценки имеют столь же хорошие асимптотические свойства, что и оценки максимального правдоподобия, при тех же условиях регулярности, что и ОМП. Грубо говоря, они представляют собой результат первой итерации при решении системы уравнений максимального правдоподобия по методу Ньютона-Ватсона. Одношаговые оценки выписываются в виде явных формул, а потому требуют существенно меньше машинного времени, а также могут применяться при ручном счете (на калькуляторах). Снимаются вопросы о сходимости алгоритмов, о выборе момента прекращения вычислений, о влиянии округлений при вычислениях на окончательный результат. ОШ оценки были использованы при разработке ГОСТ 11.011-83 (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) вместо ОМП.

Как и раньше, рассмотрим выборку x_1, x_2, ..., x_n из распределения с плотностью $f(x;\theta_0)$, где $f(x;\theta_0)$ — элемент параметрического семейства плотностей распределения вероятностей $\{f(x;\theta),\ \theta\in\Theta\}$. Здесь Θ — известное статистику k-мерное пространство параметров, являющееся подмножеством евклидова пространства R^k, а конкретное значение параметра θ_0 неизвестно. Его и будем оценивать.

Обозначим $\theta = (\theta^1, \theta^2,\ldots, \theta^k)$. Рассмотрим вектор-столбец частных производных логарифма плотности вероятности:

$$s(x;\theta) = \left\| \frac{\partial}{\partial\theta^\alpha} \ln f(x;\theta),\ \ \alpha = 1,2,\ldots,k \right\|$$

и матрицу частных производных второго порядка для той же функции:

$$b(x;\theta) = \left\| \frac{\partial^2}{\partial\theta^\alpha \partial\theta^\beta} \ln f(x;\theta),\ \ \alpha,\beta = 1,2,\ldots,k \right\|.$$

Положим:

$$s_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} s(x_i,\theta),\ \ b_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} b(x_i,\theta).$$

Пусть матрица информации Фишера $I(\theta_0) = M[-b_n(\theta_0)]$ положительно определена.

Определение 1 [10, с. 269]. Оценку $\theta(n)$ параметра θ_0 называют наилучшей асимптотически нормальной оценкой (сокращенно НАН-оценкой), если распределение случайного вектора $\sqrt{n}(\theta(n)-\theta_0)$ сходится при $n \to \infty$ к нормальному распределению с математическим ожиданием 0 и ковариационной матрицей $I^{-1}(\theta_0)$.

Определение 1 корректно: $I^{-1}(\theta_0)$ является нижней асимптотической границей для ковариационной матрицы случайного вектора $\sqrt{n}(\theta^*(n)-\theta_0)$, где $\theta^*(n)$ — произвольная оценка; кроме ОМП есть и иные НАН-оценки, например, байесовские оценки (см. [10] и др.). Сказанное об ОМП и байесовских оценках справедливо при некоторых условиях регулярности (см., например, [13]). В ряде случаев несмещенные оценки являются НАН-оценками, более того, они лучше, чем ОМП (их дисперсия меньше), при конечных объемах выборки [6].

Для анализа реальных данных естественно рекомендовать какую-либо из НАН-оценок. Это утверждение всегда верно на этапе асимптотики при изучении конкретной задачи прикладной статистики. Теоретически можно предположить, что при тщательном изучении для конкретных конечных объемов выборки наилучшей окажется какая-либо оценка, не являющаяся НАН-оценкой. Однако такие ситуации специалистам пока не известны.

Пусть $\theta_1(n)$ и I_n^{-1} — некоторые оценки θ_0 и $I^{-1}(\theta_0)$ соответственно.

Определение 2. Одношаговой оценкой (ОШ-оценкой, или ОШО) называется оценка:

$$\theta_2(n) = \theta_1(n) + I_n^{-1} s_n(\theta_1(n)).$$

Теорема 1 [14]. Пусть выполнены следующие условия.

(I) Распределение $\sqrt{n} s_n(\theta_0)$ сходится при $n \to \infty$ к нормальному распределению с математическим ожиданием 0 и ковариационной матрицей $I(\theta_0)$ и, кроме того, существует $Mb_n(\theta_0)b_n^{'}(\theta_0)$.

(II) При некотором $\varepsilon > 0$ и $n \to \infty$:

$$\sup_{\theta: 0<|\theta-\theta_0|<\varepsilon} \frac{|s_n(\theta)-s_n(\theta_0)-b_n(\theta_0)(\theta-\theta_0)|}{|\theta-\theta_0|^2} = O_p. \qquad (1)$$

(III) Для любого $\varepsilon > 0$:

$$\lim_{n\to\infty} P\left\{n^{1/4}(|\theta_1(n)-\theta_0| + \| I_n^{-1} - I^{-1}(\theta_0)\|) > \varepsilon\right\} = 0.$$

Тогда ОШ-оценка является НАН-оценкой.

Доказательство. Рассмотрим тождество

$$\sqrt{n}(\theta_2(n)-\theta_0) = \sqrt{n}(\theta_1(n)-\theta_0) + \sqrt{n} I_n^{-1} s_n(\theta_1(n)). \qquad (2)$$

В силу условия (II) теоремы

$$\sqrt{n} I_n^{-1} s_n(\theta_1(n)) = \sqrt{n} I_n^{-1} s_n(\theta_0) + \sqrt{n} I_n^{-1} b_n(\theta_0)(\theta_1(n)-\theta_0) + \sqrt{n} I_n^{-1} O_p(|\theta_1(n)-\theta_0|^2). \qquad (3)$$

Из условия (I) теоремы следует, что первое слагаемое в правой части равенства (2) сходится при $n \to \infty$ по распределению к многомерному нормальному закону с математическим ожиданием 0 и ковариационной матрицей $I^{-1}(\theta_0)$. Согласно условию (III):

$$\sqrt{n}\,|\,\theta_1(n)-\theta_0\,|^2 \to 0$$

по вероятности. Кроме того, согласно тому же условию, последовательность матриц I_n^{-1} ограничена по вероятности. Поэтому третье слагаемое в правой части формулы (2) сходится к 0 по вероятности. Для завершения доказательства теоремы осталось показать, что:

$$\sqrt{n}(\theta_1(n)-\theta_0)+\sqrt{n}I_n^{-1}b_n(\theta_0)(\theta_1(n)-\theta_0)\to 0 \qquad (4)$$

по вероятности. Левая часть формулы (3) преобразуется к виду:

$$(E+I_n^{-1}b_n(\theta_0))\sqrt{n}(\theta_1(n)-\theta_0), \qquad (5)$$

где E — единичная матрица. Поскольку из условия (I) теоремы следует, что для $b_n(\theta_0)$ справедлива (многомерная) центральная предельная теорема, то:

$$b_n(\theta_0)=-I(\theta_0)+O_p(n^{-1/2}).$$

С учетом условия (III) теоремы заключаем, что:

$$E+I_n^{-1}b_n(\theta_0)=o_p(n^{-1/4}). \qquad (6)$$

Из соотношений (4), (5) и условия (III) теоремы вытекает справедливость формулы (3). Теорема доказана.

Прокомментируем условия теоремы. Условия (I) и (II) обычно предполагаются справедливыми при рассмотрении оценок максимального правдоподобия [10]. Эти условия можно выразить в виде требований, наложенных непосредственно на плотность $f(x;\theta)$ из параметрического семейства, как это сделано, например, в [13]. Условие (III) теоремы, наложенное на исходные оценки, весьма слабое. Обычно используемые оценки $\theta_1(n)$ и I_n^{-1} являются не $n^{-1/4}$-состоятельными, а $\sqrt{n}$-состоятельными, т.е. условие (III) заведомо выполняется.

Какие оценки годятся в качестве начальных? В качестве $\theta_1(n)$ можно использовать оценки метода моментов, как это сделано в ГОСТ 11.011-83 (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) [1], или, например, квантильные. В качестве I_n^{-1} в теоретической работе [10] предлагается использовать простейшую оценку:

$$I_n^{-1} = -b_n^{-1}(\theta_1(n)). \tag{7}$$

Для гамма-распределения с неизвестными параметрами формы, масштаба и сдвига ОШ-оценки применены в [1]. При этом оценка (6) оказалась непрактичной, поскольку с точностью до погрешностей измерений и вычислений $\det(b_n) = 0$ для реальных данных о наработке резцов до предельного состояния, приведенных выше в табл. 2 раздела 3.1. Поскольку $\det(b_n) = 0$, то обратная матрица не существует, вычисления по формуле (6) невозможны. Поэтому в [1] в качестве ОШ-оценки была применена непосредственно первая итерация метода Ньютона — Рафсона решения системы уравнений максимального правдоподобия, т.е. была использована оценка:

$$I_n^{-1} = I^{-1}(\theta_1(n)). \tag{8}$$

В формуле (7) непосредственно используется явный вид зависимости матрицы информации Фишера от неизвестных параметров распределения.

В других случаях выбор тех или иных начальных оценок, в частности, выбор между (6) и (7), может определяться, например, простотой вычислений. Можно использовать также устойчивые аналоги [5] перечисленных выше оценок.

Полезно отметить, что еще в 1925 г., т.е. непосредственно при разработке метода максимального правдоподобия, его создатель Р. Фишер считал, что первая итерация по методу Ньютона — Рафсона дает хорошую оценку вектору неизвестных параметров [10, с. 298]. Он, однако, рассматривал эту оценку как аппроксимацию ОМП. А.А. Боровков воспринимает ОШ-оценки как способ «приближенного вычисления оценок максимального правдоподобия» [15, с. 225] и показывает асимптотическую эквивалентность ОШ-оценок и ОМП (в более сильных предположениях, чем в теореме 1; другими словами, теорема 1 обобщает результаты А.А. Боровкова относительно ОШ-оценок). Мы же полагаем, что ОШ-оценки имеют самостоятельную ценность, причем не меньшую, а в ряде случаев большую, чем ОМП. По нашему мнению, ОМП целесообразно применять (на этапе асимптотики) только тогда, когда они находятся явно. Во всех остальных случаях следует использовать на этом этапе ОШ-оценки (или какие-либо иные, выбранные из дополнительных соображений).

С чем связана популярность оценок максимального правдоподобия? Из всех НАН-оценок они наиболее просто вводятся, ранее других предложены. Поэтому среди математиков сложилась устойчивая традиция рассматривать ОМП в курсах математической статистики. Однако при этом игнорируются вычислительные вопросы, а также отодвигаются в сторону многочисленные иные НАН оценки.

В прикладной статистике — иные приоритеты. На первом месте — ОШ-оценки, все остальные НАН-оценки, в том числе ОМП, рассматриваются в качестве дополнительных возможностей.

Пример 1. Найдем ОШ-оценки для гамма-распределения с плотностью:

$$f(x;a,b,c)=\begin{cases}\dfrac{1}{\Gamma(a)}(x-c)^{a-1}b^{-a}\exp\left[-\dfrac{x-c}{b}\right], & x\geq c,\\ 0, & x<c.\end{cases} \tag{9}$$

Плотность вероятности в формуле (8) определяется тремя параметрами a, b, c, где a>0, b>0. При этом a является параметром формы, b — параметром масштаба и c — параметром сдвига. Здесь $\Gamma(a)$ — одна из используемых в математике специальных функций, так называемая «гамма-функция», по которой названо и распределение, задаваемое формулой (8):

$$\Gamma(a)=\int_0^{+\infty}x^{a-1}e^{-x}dx.$$

Как следует из явного вида плотности (8), логарифмическая функция правдоподобия имеет вид [16, с. 98]:

$$L=\sum_{i=1}^{n}\ln f(x_i;a,b,c)=-n\ln\Gamma(a)-na\ln b+(a-1)\sum_{i=1}^{n}\ln(x_i-c)-\frac{1}{b}\sum_{i=1}^{n}x_i+\frac{nc}{b},$$

а уравнения правдоподобия таковы:

$$\frac{\partial L}{\partial a}=-n\Psi(a)+\sum_{i=1}^{n}\ln\left(\frac{x_i-c}{b}\right)=0,$$

$$\frac{\partial L}{\partial b} = -\frac{na}{b} + \frac{1}{b^2}\sum_{i=1}^{n}(x_i - c) = 0,$$

$$\frac{\partial L}{\partial c} = -(a-1)\sum_{i=1}^{n}\frac{1}{x_i - c} + \frac{n}{b} = 0,$$

где:

$$\Psi(a) = \frac{d}{da}\ln\Gamma(a).$$

Ясно, что выписанная система нелинейных уравнений не имеет аналитического решения, в отличие от аналогичной системы для семейства нормальных распределений. Построим ОШ-оценки для задачи оценивания трех неизвестных параметров [17].

В качестве начальных оценок $\theta_1(n)$ будем использовать оценки метода моментов (см. раздел 3.1):

$$a^* = 4\frac{s^6}{m_3^2},\quad b^* = \frac{1}{2}\frac{m_3}{s^2},\quad c^* = \bar{x} - a^* b^*,$$

где $\bar{x}$ — выборочное среднее арифметическое, s^2 — выборочная дисперсия, m_3 — выборочный третий центральный момент.

Матрица информации Фишера согласно [16, с.98] при $a > 2$ имеет вид:

$$I(\theta) = I(a,b,c) = \begin{Vmatrix} \frac{d\Psi(a)}{da} & \frac{1}{b} & \frac{1}{b(a-1)} \\ \frac{1}{b} & \frac{a}{b^2} & \frac{1}{b^2} \\ \frac{1}{b(a-1)} & \frac{1}{b^2} & \frac{1}{b^2(a-2)} \end{Vmatrix}. \qquad (10)$$

Вектор-столбец частных производных логарифма плотности вероятности:

$$s(x;\theta) = s(x;a,b,c) = (s(1), s(2), s(3))'$$

имеет координаты:

$$s(1) = -\Psi(a) + \ln\left(\frac{x-c}{b}\right),$$

$$s(2) = -\frac{a}{b} + \frac{x-c}{b^2},$$

$$s(3) = -\frac{a-1}{x-c} + \frac{1}{b}.$$

Таким образом, для получения $s_n(a^*, b^*, c^*)$ необходимо вычислить две суммы:

$$\sum_{i=1}^{n} \ln\left(\frac{x_i - c}{b}\right), \ \sum_{i=1}^{n} \frac{1}{x_i - c}$$

и произвести еще несколько арифметических действий, число которых не зависит от объема выборки.

Одношаговые оценки a_n, b_n, c_n для параметров гамма-распределения вычисляют по формуле:

$$(a_n, b_n, c_n) = (a^*, b^*, c^*) + I^{-1}(a^*, b^*, c^*) s_n(a^*, b^*, c^*),$$

где I^{-1} — обратная матрица к матрице информации Фишера I, заданной формулой (9). Матрицу I^{-1} нетрудно рассчитать аналитически. Формулы для нахождения одношаговых оценок расписаны в [1]. Расчеты облегчает то обстоятельство, что для гамма-распределения вторая координата вектора $s_n(a^*, b^*, c^*)$ тождественно равна 0, т.е. $s_n^{(2)}(a^*, b^*, c^*) \equiv 0$.

При $n \to \infty$ распределение вектора оценок (a_n, b_n, c_n) приближается трехмерным нормальным распределением с математическим ожиданием, равным вектору истинных значений параметров (a, b, c), и ковариационной матрицей $I^{-1}(a_n, b_n, c_n)$. На этом приближении основаны правила расчета доверительных границ для параметров гамма-распределения [1]. Дисперсии оценок неизвестны, но зато имеются известные статистику зависимости этих дисперсий от параметров гамма-распределения. Эти зависимости непрерывные. Они стоят на главной диагонали ковариационной матрицы $I^{-1}(a_n, b_n, c_n)$. Поэтому можно вместо неизвестных параметров подставить в них оценки этих параметров и на основе принципа наследования сходимости (см. главу 14) получить состоятельные оценки дисперсий. Затем на основе оценок дисперсий обычным образом строятся доверительные интервалы для параметров гамма-распределения.

В табл. 1 приведены результаты реализации описанной выше схемы расчетов — точечные и интервальные (при односторонней доверительной вероятности 0,95) оценки параметров гамма-распределения для данных, содержащихся в табл. 2 предыдущего пункта 3.1.

Таблица 1

Одношаговые оценки и доверительные границы
для параметров гамма-распределения

Параметр	Одношаговая оценка	Верхняя доверительная граница	Нижняя доверительная граница
Формы	7,32	16,41	– 1,77
Масштаба	8,77	15,24	2,30
Сдвига	– 11,46	23,28	– 46,20

Приведенные в табл. 1 данные получены на основе асимптотических формул. Из-за конечности объема выборки необходимо внести некоторые коррективы. Поскольку параметр формы всегда положителен, $a > 0$, то нижняя доверительная граница для этого параметра должна быть неотрицательна, следует положить $a_{\text{Н}} = 0$. Поскольку плотность гамма-распределения положительна только правее параметра c, то, очевидно, $c \le x_{\min} = 9{,}00$, верхняя доверительная граница для параметра сдвига должна быть заменена на $c_{\text{В}} = 9{,}00$.

Может ли параметр сдвига быть отрицательным в данной прикладной задаче? Отрицательность параметра сдвига означает, что с положительной вероятностью рассматриваемая случайная величина отрицательна, т.е. наработка резца до предельного состояния отрицательна. Ясно, что такого быть не может, хотя для специалиста по математической статистике отрицательность параметра сдвига вполне приемлема. Однако специалист по прикладной статистике должен признать неотрицательность параметра c при обработке данных, составляющих рассматриваемую выборку. Следовательно, нижнюю доверительную границу для параметра сдвига необходимо заменить на $c_{\text{н}} = 0$.

Как следует из проведенных выше рассуждений и выкладок (см. также [16, с. 98–100]), отношение дисперсий оценок метода моментов и ОШ-оценок имеет вид:

$$\frac{Da_n}{Da*} = \frac{\left\{(a-1)^3 + \frac{1}{5}(a-1)\right\}}{a(a+1)(a+5)}$$

при больших a. Это отношение, как и должно быть из общих соображений, всегда меньше 1. Отношение дисперсий возрастает при приближении к 0 коэффициента асимметрии распределения. Если $a > 39{,}1$ (коэффициент асимметрии меньше 0,102), то эффективность оценки метода моментов превышает 80 %. При $a = 20$ (коэффициент асимметрии 0,20) она равна 65 %. Напомним, что при безграничном росте параметра формы a гамма-распределение приближается к нормальному, для которого оценки метода моментов и ОМП совпадают, а потому имеют равные дисперсии. Поэтому вполне естественно, что отношение дисперсий в формуле (10) стремится к 1 при безграничном росте параметра формы a.

Хотя дисперсии оценок метода моментов, как правило, меньше, чем дисперсии НАН-оценок, таких, как ОШО и ОМП, метод моментов играет большую роль в статистических методах. Во-первых, обычно их расчет проще (в частности, требует меньшего числа компьютерных операций), чем оценок других типов. К тому же оценки находятся с помощью выборочных моментов, которые, как правило, вычисляются на этапе описания статистических данных. Во-вторых, они служат основой для вычисления оценок других типов, например, ОШО. Для запуска итерационных методов нахождения ОМП также нужны начальные значения, и ими обычно являются оценки метода моментов. В-третьих, если учитывать погрешности результатов наблюдений, то оценки метода моментов могут оказаться точнее ОМП и асимптотически эквивалентных им ОШО [18, 19].

Методы оценивания параметров гамма-распределения и примеры расчетов для всех семи постановок, перечисленных в табл. 1 раздела 3.1, приведены в [1]. Большинство из них основано на асимптотических (при $n \to \infty$) теоретических результатах прикладной статистики. Методом статистических испытаний (Монте-Карло) показано, что уже при $n \geq 10$ используемые приближения удовлетворительны. Другими словами, асимптотической нормальностью оценок и другими важными для проведенных выше рассуждений предельными результатами можно пользоваться уже при $n \geq 10$.

Алгоритмическое и программное обеспечение ОШ-оценок для распределения Вейбулла — Гнеденко и гамма-распределения рассмотрено в содержательной монографии [20]. История вопроса освещена в статье [14].

3.3. Асимптотика решений экстремальных статистических задач

Если проанализировать приведенные в [19, 21] постановки и результаты, касающиеся эмпирических и теоретических средних и законов больших чисел (в пространствах произвольной природы), то становится очевидной возможность их обобщения. Так, доказательства теорем практически не меняются, если считать, что функция $f(x,y)$ определена на декартовом произведении бикомпактных различных пространств X и Y, а не на X^2. Тогда можно считать, что элементы выборки лежат в X, а Y — пространство параметров, подлежащих оценке.

Обобщения законов больших чисел. Пусть, например, выборка $x_1 = x_1(\omega)$, $x_2 = x_2(\omega)$, … , $x_n = x_n(\omega)$ взята из распределения с плотностью $p(x,y)$, где y — неизвестный параметр. Если положить:

$$f(x,y) = -\ln p(x,y),$$

то задача нахождения эмпирического среднего:

$$f_n(\omega, y) = \frac{1}{n}\sum_{k=1}^{n} f(x_k(\omega), y) \to \min$$

переходит в задачу оценивания неизвестного параметра y методом максимального правдоподобия:

$$\sum_{k=1}^{n} \ln p(x_k(\omega), y) \to \max .$$

Соответственно законы больших чисел переходят в утверждения о состоятельности этих оценок в случае пространств X и Y общего вида. При такой интерпретации функция $f(x,y)$ уже не является расстоянием или показателем различия, как при определении теоретических и эмпирических средних. Однако для доказательства сходимости оценок к соответствующим значениям параметров это и не требуется. Достаточно, например, непрерывности этой функции на декартовом произведении бикомпактных пространств X и Y.

В случае функции $f(x,y)$ общего вида можно говорить об определении в пространствах произвольной природы оценок минимального контраста и их состоятельности. При этом при каждом конкретном значении параметра y справедливо предельное соотношение:

$$f_n(\omega, y) = \frac{1}{n}\sum_{k=1}^{n} f(x_k(\omega), y) \to Mf(x_1(\omega), y) = g(y),$$

где f — функция контраста. Тогда состоятельность оценок минимального контраста вытекает из справедливости предельного перехода:

$$Arg\min\left\{\frac{1}{n}\sum_{k=1}^{n} f(x_k(\omega), y)\right\} \to Arg\min\{Mf(x_1(\omega), y)\}.$$

Частными случаями оценок минимального контраста являются, устойчивые (робастные) оценки Тьюки — Хубера (см. ниже), а также оценки параметров в задачах аппроксимации (параметрической регрессии) в пространствах произвольной природы.

Можно пойти и дальше в обобщении законов больших чисел. Пусть известно, что при каждом конкретном y при безграничном росте n имеет быть сходимость по вероятности:

$$f_n(\omega, y) \to f(y),$$

где $f_n(\omega, y)$ — последовательность случайных функций на пространстве Y, а $f(y)$ — некоторая функция на $У$. В каких случаях и в каком смысле имеет место сходимость:

$$\text{Argmin } \{f_n(\omega, y), y \in X\} \to \text{Argmin } \{f(y), y \in X\}?$$

Другими словами, когда из поточечной сходимости функций вытекает сходимость точек минимума?

Причем здесь можно под n понимать натуральное число. А можно рассматривать сходимость по направленному множеству (глава 14), или же, что практически то же самое — «сходимость по фильтру» в смысле Картана и Бурбаки [22, с. 118]. В частности, можно описывать ситуацию вектором, координаты которого — объемы нескольких выборок, и все они безгранично растут.

Поскольку основные задачи прикладной статистики можно представить в виде оптимизационных задач, то ответ на поставленный вопрос о сходимости точек минимума дает возможность единообразного подхода к изучению асимптотики решений разнообразных экстремальных статистических задач. Одна из возможных формулировок, основанная на бикомпактности пространств X и $У$ и нацеленная на изучение оценок минимального контраста, кратко рассмотрена выше. Другой подход развит в работе [23]. Он основан на использовании понятий асимптотической равномерной разбиваемости и координатной асимптотической равномерной разбиваемости пространств. С помощью указанных подходов удается стандартным образом обосновывать состоятельность оценок характеристик и параметров в основных задачах прикладной статистики.

Рассматриваемую тематику можно развивать дальше, в частности, рассматривать аналоги законов больших чисел в случае пространств, не являющихся бикомпактными, а также изучать скорость сходимости $\mathrm{Argmin}\{f_n(x(\omega), y), y \in X\}$ к $\mathrm{Argmin}\{f(y), y \in X\}$.

Приведем примеры применения результатов о предельном поведении точек минимума.

Задача аппроксимации зависимости (параметрической регрессии). Пусть X и Y — некоторые пространства. Пусть имеются статистические данные — n пар (x_k, y_k), где $x_k \in X$, $y_k \in Y$, $k = 1, 2, \ldots, n$. Задано параметрическое пространство Θ произвольной природы и семейство функций $g(x,\theta)$: $X \times \Theta \to Y$. Требуется подобрать параметр $\theta \in \Theta$ так. чтобы $g(x_k, \theta)$ наилучшим образом приближали y_k, $k = 1, 2, \ldots, n$. Пусть f_k — последовательность показателей различия в $У$. При сделанных предположениях параметр θ естественно оценивать путем решения экстремальной задачи:

$$\theta_n = Arg\min_{\theta \in \Theta} \sum_{k=1}^{n} f_k(g(x_k, \theta), y_k). \tag{1}$$

Часто, но не всегда, все f_k совпадают. В классической постановке, когда $X = R^k$, $У = R^1$, функции f_k различны при неравноточных наблюдениях, например, когда число опытов меняется от одной точки x проведения опытов к другой.

Если $f_k(y_1,y_2) = f(y_1,y_2) = (y_1 - y_2)^2$, то получаем общую постановку метода наименьших квадратов (см. подробности в главе 6):

$$\theta_n = Arg\min_{\theta \in \Theta} \sum_{k=1}^{n} (g(x_k, \theta) - y_k)^2.$$

В рамках детерминированного анализа данных остается единственный теоретический вопрос — о существовании θ_n. Если все участвующие в формулировке задачи (1) функции непрерывны, а минимум берется по бикомпакту, то θ_n существует. Есть и иные условия существования θ_n [23–25].

При появлении нового наблюдения x в соответствии с методологией восстановления зависимости рекомендуется выбирать оценку соответствующего y по правилу:

$$y^* = g(x, \theta_n).$$

Обосновать такую рекомендацию в рамках детерминированного анализа данных невозможно. Это можно сделать только в вероятностной теории, равно как и изучить асимптотическое поведение θ_n, в частности, доказать состоятельность этой оценки.

Как и в классическом случае, вероятностную теорию целесообразно строить для трех различных постановок.

1. Переменная x — детерминированная (например, время), переменная y — случайная, ее распределение зависит от x.

2. Совокупность (x_k, y_k), $k = 1, 2, \ldots, n$, — выборка из распределения случайного элемента со значениями в $X \times Y$.

3. Имеется детерминированный набор пар (x_{k0}, y_{k0}), $k = 1, 2, \ldots, n$, результат наблюдения (x_k, y_k) является случайным элементом, распределение которого зависит от (x_{k0}, y_{k0}). Это — постановка т.н. конфлюэнтного анализа.

Во всех трех случаях:

$$f_n(\omega,\theta) = \sum_{k=1}^{n} f_k(g(x_k,\theta), y_k),$$

однако случайность входит в правую часть по-разному в зависимости от постановки, от которой зависит и определение предельной функции $f(\theta)$.

Проще всего выглядит $f(\theta)$ в случае второй постановки при $f_k \equiv f$:

$$f(\theta) = Mf(g(x_1,\theta),y).$$

В случае первой постановки:

$$f(\theta) = \lim_{n\to\infty} \sum_{k=1}^{n} Mf_k(g(x_k,\theta), y_k(\omega))$$

в предположении существования указанного предела. Ситуация усложняется для третьей постановки:

$$f(\theta) = \lim_{n\to\infty} \sum_{k=1}^{n} Mf_k(g(x_k(\omega),\theta), y_k(\omega)).$$

Во всех трех случаях на основе общих результатов о поведении решений экстремальных статистических задач можно изучить [23–25] асимптотику оценок θ_n. При выполнении соответствующих внутриматематических условий регулярности оценки оказываются состоятельными, т.е. удается восстановить зависимость.

Аппроксимация и регрессия. Соотношение (1) дает решение задачи аппроксимации. Поясним, как эта задача соотносится с нахождением регрессии. Согласно [26] для случайной величины (ξ, η) со значениями в $X \times Y$ регрессией η на ξ относительно меры близости f естественно назвать решение задачи:

$$Mf(g(\xi), \eta) \to \min_{g}, \qquad (2)$$

где f: $Y \times Y \to R^1$, g: $X \to Y$, минимум берется по множеству всех измеримых функций $\{g\}$.

Можно исходить и из другого определения. Для каждого $x \in X$ рассмотрим случайную величину η(x), распределение которой является условным распределением η при условии $\xi = x$. В соответствии с определением математического ожидания в пространстве общей природы назовем условным математическим ожиданием решение экстремальной задачи:

$$M(\eta \mid \xi = x) = Arg\min\{Mf(y, \eta(x)), y \in Y\}.$$

Оказывается, при обычных предположениях измеримости решение задачи (2) совпадает с $M(\eta \mid \xi = x)$. (Внутриматематические уточнения типа «равенство имеет место почти всюду» здесь опущены.)

Если заранее известно, что условное математическое ожидание $M(\eta \mid \xi = x)$ принадлежит некоторому параметрическому семейству $g(x,\theta)$, то задача нахождения регрессии сводится к оцениванию параметра θ в соответствии с рассмотренной выше второй постановкой вероятностной теории параметрической регрессии. Если же нет оснований считать, что регрессия принадлежит параметрическому семейству, то можно использовать непараметрические оценки регрессии. Они строятся с помощью непараметрических оценок плотности [19, 21]. Способы построения непараметрических оценок плотности рассмотрены в главе 2.

Пусть ν_1 — мера в X, ν_2 — мера в Y, а их прямое произведение $\nu = \nu_1 \times \nu_2$ — мера в $X \times Y$. Пусть $g(x,y)$ — плотность случайного элемента (ξ,η) по мере ν. Тогда условная плотность $g(y|x)$ распределения η при условии $\xi = x$ имеет вид:

$$g(y \mid x) = \frac{g(x, y)}{\int_Y g(x, y)\nu_2(dy)} \qquad (3)$$

(в предположении, что интеграл в знаменателе отличен от 0). Следовательно,

$$Mf(y,\eta(x)) = \int_Y f(y,a)g(a\,|\,x)\nu_2(da),$$

а потому:

$$M(\eta\,|\,\xi = x) = Arg\min_{y\in Y} Mf(y,\eta(x)) = Arg\min_{y\in Y} \int_Y f(y,a)g(a\,|\,x)\nu_2(da).$$

Заменяя $g(x,y)$ в (3) непараметрической оценкой плотности $g_n(x,y)$, получаем оценку условной плотности:

$$g_n(y\,|\,x) = \frac{g_n(x,y)}{\int_Y g_n(x,y)\nu_2(dy)}. \qquad (4)$$

Если $g_n(x,y)$ — состоятельная оценка $g(x,y)$, то числитель (4) сходится к числителю (3). Сходимость знаменателя (4) к знаменателю (3) обосновывается с помощью предельной теории статистик интегрального типа [19]. В итоге получаем утверждение о состоятельности непараметрической оценки (4) условной плотности, заданной формулой (3).

Непараметрическая оценка регрессии ищется как:

$$M_n(\eta\,|\,\xi = x) = Arg\min_{y\in Y} \int_Y f(y,a)g_n(a\,|\,x)\nu_2(da).$$

Состоятельность этой оценки следует из приведенных выше общих результатов об асимптотическом поведении решений экстремальных статистических задач.

Применение к методу главных компонент. Исходные данные — набор векторов ξ_1, ξ_2, … , ξ_n, лежащих в евклидовом пространстве R^k размерности k. Цель состоит в снижении размерности, т.е. в уменьшении числа рассматриваемых показателей. Для этого берут всевозможные линейные ортогональные нормированные центрированные комбинации исходных показателей, получают k новых показателей, из них берут первые m, где $m < k$ (подробности см. в главе 6.5). Матрицу преобразования C выбирают так, чтобы максимизировать информационный функционал:

$$I_n(C) = \frac{s^2(z(1)) + s^2(z(2)) + \ldots + s^2(z(m))}{s^2(x(1)) + s^2(x(2)) + \ldots + s^2(x(k))}, \qquad (5)$$

где $x(i)$, $i = 1, 2, \ldots, k$, — исходные показатели; исходные данные имеют вид $\xi_j = (x_j(1), x_j(2), \ldots, x_j(k))$, $j = 1, 2, \ldots, n$; при этом $z(\alpha)$, $\alpha = 1, 2, \ldots, m$, — комбинации исходных показателей, полученные с помощью матрицы C. Наконец, $s^2(z(\alpha))$, $\alpha = 1, 2, \ldots, m$, $s^2(x(i))$, $i = 1, 2, \ldots, k$, — выборочные дисперсии переменных, указанных в скобках.

Укажем подробнее, как новые показатели (главные компоненты) $z(\alpha)$ строятся по исходным показателям $x(i)$ с помощью матрицы C:

$$z_j(\alpha) = \sum_{\beta=1}^{k} c_{\alpha\beta}(x_j(\beta) - \overline{x(\beta)}),\ \alpha = 1,2,\ldots,m,\ \ j = 1,2,\ldots,n,$$

где

$$\overline{x(\beta)} = \frac{1}{n}\sum_{j=1}^{n} x_j(\beta).$$

Матрица $C = ||c_{\alpha\beta}||$ порядка $m \times k$ такова, что:

$$\sum_{\beta=1}^{k} c_{\alpha\beta}^2 = 1,\ \ \alpha = 1,2,\ldots,m \tag{6}$$

(матрица C является нормированной),

$$\sum_{\beta=1}^{k} c_{\alpha\beta} c_{\gamma\beta} = 0,\ \ \alpha,\gamma = 1,2,\ldots,m,\ \ \alpha \neq \gamma \tag{7}$$

(матрица C является ортогональной).

Решением основной задачи метода главных компонент является:

$$C_n = Arg\min_{C}(-I_n(C)),$$

где минимизируемая функция определена формулой (5), а минимизация проводится по всем матрицам C, удовлетворяющим условиям (6) и (7).

Вычисление матрицы C_n — задача детерминированного анализа данных. Однако, как и в иных случаях, например, для медианы Кемени, возникает вопрос об асимптотическом поведении C_n. Является ли решение основной задачи метода главных компонент устойчивым, т.е. существует ли предел C_n при $n \to \infty$? Чему равен этот предел?

Ответ, как обычно, может быть дан только в вероятностной теории. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые одинаково распределенные случайные вектора. Положим:

$$z_\infty(\alpha) = \sum_{\beta=1}^{k} c_{\alpha\beta}(x_1(\beta) - Mx_1(\beta)),\ \alpha = 1,2,\dots,m,$$

где матрица $C = \|c_{\alpha\beta}\|$ удовлетворяет условиям (6) и (7). Введем функцию от матрицы:

$$I(C) = \frac{D(z_\infty(1)) + D(z_\infty(2)) + \dots + D(z_\infty(m))}{D(x(1)) + D(x(2)) + \dots + D(x(k))}.$$

Легко видеть, что при $n \to \infty$ и любом C:

$$I_n(C) \to I(C).$$

Рассмотрим решение предельной экстремальной задачи:

$$C_\infty = \underset{C}{Arg\min}(-I(C)).$$

Естественно ожидать, что:

$$\lim_{n\to\infty} C_n = C_\infty.$$

Действительно, это соотношение вытекает из приведенных выше общих результатов об асимптотическом поведении решений экстремальных статистических задач.

Таким образом, теория, развитая для пространств произвольной природы, позволяет единообразным образом изучать конкретные процедуры прикладной статистики.

3.4. Робастность статистических процедур

Термин «робастность» (*robustness* — англ.) образован от *robust* — крепкий, грубый (англ.). Сравните с названием одного из сортов кофе — *robusta*. Имеется в виду, что робастные статистические процедуры должны «выдерживать» ошибки, которые теми или иными способами могут попадать в исходные данные или искажать предпосылки используемых вероятностно-статистических моделей.

Термин «робастный» стал популярным в нашей стране в 1970-е гг. Сначала он использовался фактически как сужение термина «устойчивый» на алгоритмы статистического анализа данных классического типа (не включая теорию измерений, статистику нечисловых и интервальных данных). Затем реальная сфера его применения сузилась.

Пусть исходные данные — это выборка, т.е. совокупность независимых одинаково распределенных случайных величин с одной и той же функцией распределения $F(x)$. Наиболее простая модель изучения устойчивости — это модель засорения:

$$F(x) = (1-\varepsilon)F_0(x) + \varepsilon H(x). \tag{1}$$

Эта модель имеются также моделью Тьюки — Хубера. (Джон Тьюки — американский исследователь, П. Хубер, или Хьюбер — швейцарский ученый.) Модель (1) показывает, что с близкой к 1 вероятностью, а именно, с вероятностью $(1-\varepsilon)$, наблюдения берутся из совокупности с функцией распределения $F_0(x)$, которая предполагается обладающей «хорошими» свойствами. Например, она имеет известный статистику вид (хотя бы с точностью до параметров), у нее существуют все моменты, и т.д. Но с малой вероятностью ε появляются наблюдения из совокупности с «плохим» распределением (например, взятые из распределения Коши, не имеющего математического ожидания), резко выделяющиеся аномальные наблюдения, выбросы.

Актуальность модели (1) не вызывает сомнений. Наличие засорений (выбросов) может сильно исказить результаты эконометрического анализа данных. Например, пусть функция распределения элементов выборки имеет вид (1). Причем первое слагаемое соответствует случайной величине с конечным математическим ожиданием, а второе — такой случайной величине, для которой математического ожидания не существует. Тогда для итоговой функции распределения также не существует математического ожидания. Исследователя обычно интересуют характеристики первого слагаемого, но найти их, т.е. освободиться от влияния засорения, не так-то просто. Например, среднее арифметическое результатов наблюдений не будет иметь никакого предела (это — строгое математическое утверждение, вытекающее из того, что математическое ожидание не существует [27]).

Существуют различные способы борьбы с засорением. Эмпирическое правило найдено в фигурном катании: при подведении итогов работы команды судей наибольшая и наименьшая оценки отбрасываются, а по остальным рассчитывается средняя арифметическая. Ясно, что единичное «засорение» окажется среди отброшенных оценок.

Оценивать характеристики и параметры, проверять статистические гипотезы, вообще осуществлять статистический анализ данных все чаще рекомендуют на основе эмпирических квантилей (другими словами, порядковых статистик, членов вариационного ряда), отделенных от концов вариационного ряда. Речь идет об использовании статистик вида:

$$ax(0{,}1n) + bx(0{,}3n) + cx(0{,}5n) + dx(0{,}7n) + ex(0{,}9n),$$

где a, b, c, d, e — заданные числа, $x(0{,}1n)$, $\mathrm{x}(0{,}3n)$, $x(0{,}5n)$, $x(0{,}7n)$, $x(0{,}9n)$ — члены вариационного ряда с номерами, наиболее близкими к числам, указанным в скобках. Так ценой небольшой потери в эффективности оценок избавляемся от влияния засоренности типа, описанной в модели (1).

Вариантом этого подхода является переход к сгруппированным данным. Отрезок прямой, содержащий основную часть наблюдений, разбивается на интервалы, и вместо количественных значений статистик подсчитывает лишь, сколько наблюдений попало в те или иные интервалы. Особое значение приобретают крайние интервалы — к ним относят все наблюдения, которые больше некоторого верхнего порога и меньше некоторого нижнего порога. Любым методам анализа сгруппированных данных резко выделяющиеся наблюдения не страшны.

Можно поставить под сомнение и саму опасность засорения. Дело в том, что практически все реально наблюдающиеся величины ограничены. Все они лежат на каком-то интервале — от и до. Это совершенно ясно, если речь идет о физическом измерении — все результаты измерений укладывается в шкалу прибора. По-видимому, и для иных статистических измерений наибольшие сложности создают не сверхбольшие помехи, а те засорения, что находятся «на грани» между «интуитивно возможными значениями» и «интуитивно невозможными».

Что же это означает для практики статистического анализа данных? Если элементы выборки по абсолютной величине не превосходят числа A, то все возможные засорения могут сдвинуть среднее арифметическое на величину не более εA. Если засорение невелико, то и сдвиг мал.

Построена достаточно обширная и развитая теория, посвященная разработке и изучению методов анализа данных в модели (1). С ней можно познакомиться по монографиям [28–30]. К сожалению, в теории обычно предполагается известной степень засорения ε, а на практике эта величина неизвестна. Кроме того, теория обычно направлена на защиту от воздействий, якобы угрожающих из бесконечности (например, отсутствием математического ожидания), а на самом деле реальные данные финитны (сосредоточены на конечных отрезках). Все это объясняет, почему теория робастности, исходящая из модели (1), популярна среди теоретиков, но мало интересна тем, кто анализирует реальные технические, экономические, медицинские и иные статистические данные.

Рассмотрим несколько более сложную модель. Пусть наблюдаются реализации $x_1, x_2, ..., x_n$ независимых случайных величин с функциями распределения $F_1(x), F_2(x), ..., F_n(x)$ соответственно. Эта модель соответствует гипотезе о том, что в процессе наблюдения (измерения) условия несколько менялись. Естественной представляется модель малых отклонений функций распределений наблюдаемых случайных величин от некоторой «базовой» функции распределения $F_0(x)$. Множество возможных значений функций распределений наблюдаемых случайных величин (т.е. совокупность допустимых отклонений согласно общей схеме устойчивости, рассмотренной в главе 14) описывается следующим образом:

$$E((F_1, F_2, ..., F_n); \varepsilon) = \{(F_1, F_2, ..., F_n) : \sup_x |F_i(x) - F_0(x)| < \varepsilon, i = 1,2,...,n].$$

Следующий тип моделей — это введение малой (т.е. слабой) зависимости между рассматриваемыми случайными величинами. Такие модели рассмотрены, например, в монографии [31]. Ограничения на взаимную зависимость можно задать разными способами. Пусть $F(x_1, x_2, ..., x_n)$ — совместная функция распределения n-мерного случайного вектора, $F_1(x_1)$, $F_2(x_2)$, … , $F_n(x_n)$ — функции распределения его координат. Если все координаты независимы, то $F(x_1, x_2, ..., x_n)$ = $F_1(x_1)F_2(x_2)$ … $F_n(x_n)$. Пусть $\rho(i, j)$ – коэффициент корреляции между i-й и j-й случайными величинами — координатами вектора. Множество возможных совместных функций распределения (т.е. совокупность допустимых отклонений от ситуации независимости координат согласно общей схеме устойчивости, рассмотренной в главе 14) описывается следующим образом:

$$\mathring{A}(F; \varepsilon) = \{F(x_1, x_2, ..., x_n) : P(x_i(\omega) < x) = F_i(x),\ |\rho(i, j)| \le \varepsilon,\ 1 \le i < j \le n\}.$$

Таким образом, фиксируются функции распределения координат, а коэффициенты корреляции предполагаются малыми (по абсолютной величине).

Есть еще целый ряд постановок задач робастности. Если накладывать погрешности непосредственно на результаты наблюдений (измерений) и предполагать лишь, что эти погрешности не превосходят (по абсолютной величине) заданных величин, то получаем постановки задач статистики интервальных данных [19]. При этом каждый результат наблюдения превращается в интервал — исходное значение плюс-минус максимально возможная погрешность.

Разработано много вариантов робастных методов анализа статистических данных (см. монографии [5, 28–30]). Иногда говорят, что робастные методы позволяют использовать информацию о том, что реальные наблюдения лежат «около» тех или иных параметрических семейств, например, нормальных. В этом, дескать, их преимущество по сравнению с непараметрическими методами, которые предназначены для анализа данных, распределенных согласно произвольной непрерывной функции распределения. Однако количественных подтверждений этих уверений энтузиастам робастных методов обычно не удается найти. В основном потому, что термин «около» трудно формализовать.

На примере сравнения различных подходов к изучению робастности статистических процедур оценивания и проверки гипотез видны сложности, связанные с изучением устойчивости. Дело в том, что для каждой конкретной статистической задачи можно самыми разными способами задать совокупность допустимых отклонений. Так, выше кратко рассмотрены четыре такие совокупности, соответствующие:

1) модели засорения Тьюки — Хубера;

2) модели малых отклонений функций распределения;

3) модели слабых связей;

4) модели интервальных данных.

В каждой из этих моделей общая схема устойчивости (глава 14) предлагает для решения целый спектр задач устойчивости. Кроме изучения свойств робастности известных статистических процедур можно в каждой из постановок находить оптимальные процедуры. Однако практическая ценность этих оптимальных процедур, как правило, невелика, поскольку в других постановках оптимальными будут уже другие процедуры.

3.5. Оценивание для сгруппированных данных

При вычислении различных статистических характеристик часто пользуются сгруппированными данными. Мы рассматриваем погрешность группировки как один из видов погрешностей наблюдений. До появления мощных компьютеров группировку проводили для облегчения расчетов. К сожалению, устаревшие рекомендации по обязательному проведению группировки укоренились в курсах по «общей теории статистики» [32]. В отмененном из-за низкого научного уровня ГОСТе [33] группировку предписывалось проводить даже при проверке согласия с помощью критериев Колмогорова и омега-квадрат, что приводит к прямым ошибкам.

В современных условиях нет необходимости сокращать число арифметических операций при анализе статистических данных. Вместо гистограмм для описания распределения в настоящее время рекомендуют использовать непараметрические оценки плотности. Однако с существующей традицией приходится считаться, для борьбы с ней надо знать влияние группировки данных на статистические характеристики.

Другая причина использования сгруппированных данных — неточность измерений, приводящих либо к автоматическому (с помощью средств измерения) округлению, либо к округлению, проводимому специалистом, осуществляющим измерение.

В настоящем разделе рассмотрим статистические методы анализа сгруппированных данных. Начнем с одномерных данных.

Вероятностная модель группировки. Пусть выборка объема n из распределения числовой случайной величины X с плотностью вероятности $f(x)$. Пусть элементы выборки сгруппированы по интервалам длины h, центры которых находятся в точках $a_i = a_0 + ih$, где $i = 0, \pm 1, \pm 2, \ldots$ (Поскольку вероятность попадания в точку соприкосновения интервалов равна 0, то нет необходимости указывать, к какому из интервалов относятся такие точки.) В таких случаях при вычислении моментов и других выборочных характеристик по сгруппированным данным обычно предполагается [4, с. 394–394], что все выборочные значения, принадлежащие некоторому интервалу, совпадают с центром этого интервала. Тогда фактически рассматривается выборка из дискретного распределения, в соответствии с которым случайная величина Y принимает значения $a_i = a_0 + ih$ с вероятностями:

$$p_i = P(Y = a_i) = \int_{a_i - \frac{h}{2}}^{a_i + \frac{h}{2}} f(x)dx .$$

Пользуясь сгруппированными данными, оценивают те или иные выборочные характеристики, например, моменты:

$$M(Y^k) = \sum_{-\infty<i<+\infty} a_i^k p_i .$$

Во многих случаях исследователя интересуют не эти «групповые» моменты, а моменты исходной непрерывно случайной величины X:

$$M(X^k) = \int_{-\infty}^{+\infty} x^k f(x)dx .$$

Поэтому важно изучить соотношения между этими двумя множествами моментов. Оказывается, при некоторых внутриматематических условиях регулярности приближенные значения моментов $M(X^k)$ можно получить, прибавляя некоторые поправки к «групповым» моментам $M(Y^k)$.

Групповые моменты запишем в виде:

$$M(Y^k) = \sum_{-\infty<i<+\infty} a_i^k \int_{a_i-\frac{h}{2}}^{a_i+\frac{h}{2}} f(x)dx = \sum_{-\infty<i<+\infty} g(a_i),$$

где:

$$g(a) = a^k \int_{a-\frac{h}{2}}^{a+\frac{h}{2}} f(x)dx .$$

Для решения поставленной задачи полезной оказывается формула Эйлера — Маклорена [4, с. 394–398]. В настоящем разделе нам понадобится также ее обобщение, полученное в [34].

Обобщение формулы Эйлера — Маклорена. В различных статистических методах анализа данных используются дискретные распределения. Возникает необходимость вычисления сумм следующего вида:

$$S(g, a, h, n) = \sum_{i=1}^{n} g(a + (i-1)h) ,$$

где $g(x)$ — достаточно гладкая функция. Другими словами, требуется просуммировать значения функции $g(x)$ в n точках, отстоящих друг от друга на расстояние h, начиная с точки a. Типичным примером является суммирование биномиальных вероятностей с целью вычисления попадания биномиально распределенной случайной величины в заданный интервал.

Классическая формула Эйлера-Маклорена такова [35, п. 465]:

$$S(g,a,h,n)=\frac{1}{h}\int_{a}^{a+nh}g(x)dx+A_1\{g(a+nh)-g(a)\}+A_2h\{g'(a+nh)-g'(a)\}+\ldots$$
$$+A_mh^{m-1}\{g^{(m-1)}(a+nh)-g^{(m-1)}(a)\}+R_m.$$

Здесь R_m — остаточный член порядка h^m, выражение для которого приведено в [35, с. 540]:

$$A_1=-\frac{1}{2},\ \ A_{2p-1}=0,\ \ p>1,\ \ A_{2p}=(-1)^{p-1}\frac{B_p}{(2p)!},\ \ p\geq 1,$$

где B_p есть p-е число Бернулли [35, п. 449]. В частности:

$$A_2=-\frac{1}{12},\ \ A_4=\frac{1}{720},\ \ A_6=-\frac{1}{30240}.$$

Очевидным недостатком классической формулы Эйлера-Маклорена является ее несимметричность — левый конец отрезка $[a,\ a+nh]$, по которому проводится интегрирование, входит в число точек, значения функции в которых суммируется, а правый — нет. Этот недостаток связан с тем, что каждому слагаемому в рассматриваемой сумме ставится в соответствие отрезок длины h, значение функции $g(x)$ в *левом* конце которого и есть рассматриваемое слагаемое. В результате нет симметричности — формула меняется при изменении направления оси x-ов.

Чтобы избавиться от несимметричности, достаточно слагаемому $g(x_0)$ (где $x_0 = a + (i - 1)h$ при некотором i) поставить в соответствие отрезок $\left[x_0-\frac{h}{2};\ x_0+\frac{h}{2}\right]$. Опишем подход к получению асимптотических разложений, впервые предложенный в статье [34] в связи с изучением скорости сходимости распределения классической статистики омега-квадрат (Крамера — Мизеса —

Смирнова) к предельному распределению. Будем исходить из формулы Тейлора с дополнительным членом в виде определенного интеграла [35, п. 318]:

$$\begin{aligned}g(x) = g(x_0) + \frac{g'(x_0)}{1!}(x-x_0) + \frac{g''(x_0)}{2!}(x-x_0)^2 + \ldots \\ + \frac{g^{(m)}(x_0)}{m!}(x-x_0)^m + \frac{1}{m!}\int_{x_0}^{x} g^{(m+1)}(t)(x-t)^n\,dt\end{aligned}.$$

Проинтегрируем почленно обе части этого равенства по x по отрезку отрезок $\left[x_0 - \frac{h}{2};\ x_0 + \frac{h}{2}\right]$. Получим, что:

$$\begin{aligned}&\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g(x)dx = g(x_0)h + \frac{g'(x_0)}{1!}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}}(x-x_0)dx + \frac{g''(x_0)}{2!}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}}(x-x_0)^2\,dx + \ldots \\ &+ \frac{g^{(m)}(x_0)}{m!}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}}(x-x_0)^m\,dx + \frac{1}{m!}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}}\left\{\int_{x_0}^{x} g^{(m+1)}(t)(x-t)^n\,dt\right\}dx = g(x_0)h + \frac{g''(x_0)}{2!}\frac{h^3}{12} + \\ &+ \frac{g^{(4)}(x_0)}{4!}\frac{h^5}{80} + \ldots + \frac{g^{(m)}(x_0)}{m!}\frac{h^{m+1}}{2^m(m+1)} + \frac{1}{m!}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}}\left\{\int_{x_0}^{x} g^{(m+1)}(t)(x-t)^n\,dt\right\}dx\end{aligned}$$

(при четном m). Весьма важно, что обращаются в 0 все интегралы, соответствующие нечетным степеням.

Из только что полученного соотношения следует, что:

$$g(x_0)h = \int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g(x)dx - \frac{g''(x_0)h^3}{24} - \frac{g^{(4)}(x_0)h^5}{1920} - \ldots - \frac{g^{(m)}(x_0)}{m!}\frac{h^{m+1}}{2^m(m+1)} - Q_m,$$

где Q_m — остаточный член, указанный в предыдущей формуле.

Применим аналогичную процедуру ко второй производной — функции $g''(x)$. Исходим из формулы Тейлора:

$$g''(x) = g''(x_0) + \frac{g^{(3)}(x_0)}{1!}(x-x_0) + \frac{g^{(4)}(x_0)}{2!}(x-x_0)^2 + \ldots$$

Интегрируя, получаем, что

$$\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g''(x)dx = g''(x_0)h + \frac{g^{(4)}(x_0)}{2!}\frac{h^3}{12} + \ldots,$$

откуда:

$$g''(x_0)h = \int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g''(x)dx - \frac{g^{(4)}(x_0)}{2!}\frac{h^3}{12} - \ldots = \left\{g'\left(x_0+\frac{h}{2}\right) - g'\left(x_0-\frac{h}{2}\right)\right\} - \frac{g^{(4)}(x_0)}{2!}\frac{h^3}{12} - \ldots .$$

Подставляя в формулу для $g(x_0)h$, получаем, что

$$g(x_0)h = \int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g(x)dx - \frac{h^2}{24}\left(\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g''(x)dx - \frac{g^{(4)}(x_0)}{2!}\frac{h^3}{12} - \ldots\right) - \frac{g^{(4)}(x_0)h^5}{1920} + O(h^7) =$$

$$= \int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g(x)dx - \frac{h^2}{24}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g''(x)dx + \frac{7g^{(4)}(x_0)h^5}{5760} + O(h^7) = \int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} g(x)dx -$$

$$-\frac{h^2}{24}\left\{g'\left(x_0+\frac{h}{2}\right) - g'\left(x_0-\frac{h}{2}\right)\right\} + \frac{7g^{(4)}(x_0)h^5}{5760} + O(h^7).$$

Дальше можно провести аналогичные выкладки для четвертой производной, и т.д. С помощью продемонстрированного подхода можно получить, например, следующий результат:

$$g(x_0) = \frac{1}{h}\int_{x_0-\frac{h}{2}}^{x_0+\frac{h}{2}} A(x)dx,$$

где:

$$A(x) = g(x) - \frac{h^2}{24}g''(x) + \frac{7h^4}{5760}g^{(4)}(x) - \frac{31h^6}{967680}g^{(6)}(x) +$$

$$+\int_{x_0}^{x} g^{(8)}(t)\left\{-\frac{(x-t)^7}{5040} + \frac{(x-t)^5h^2}{2880} - \frac{7(x-t)^3h^4}{34560} + \frac{31(x-t)h^6}{967680}\right\}dt.$$

Таким образом, справедливо следующее соотношение, напоминающее классическую формулу Эйлера — Маклорена:

$$S(g,a,h,n)=\sum_{i=1}^{n} g(a+(i-1)h)=\frac{1}{h}\int\limits_{a-\frac{h}{2}}^{a+\left(n-\frac{1}{2}\right)h} g(x)dx-\frac{h}{24}\left\{g'\left(a+\left(n-\frac{1}{2}\right)h\right)-g'\left(a-\frac{h}{2}\right)\right\}+$$

$$+\frac{7h^3}{5760}\left\{g^{(3)}\left(a+\left(n-\frac{1}{2}\right)h\right)-g^{(3)}\left(a-\frac{h}{2}\right)\right\}-\frac{31h^5}{967680}\left\{g^{(5)}\left(a+\left(n-\frac{1}{2}\right)h\right)-g^{(5)}\left(a-\frac{h}{2}\right)\right\}+O(h^7).$$

Сравнение с классической формулой Эйлера — Маклорена показывает, что в последней формуле отсутствует первый член асимптотического разложения, соответствующий $A_1\{g(a+nh)-g(a)\}$, а коэффициенты при остальных членах меньше по абсолютной величине:

$$\frac{1}{24}<|A_2|=\frac{1}{12};\ \frac{7}{5760}<A_4=\frac{1}{720};\ \frac{31}{967680}<|A_6|=\frac{1}{30240}.$$

Полученная формула симметрична — не меняется при изменении направлении оси x-ов.

Из сказанного вытекает рекомендация: вместо классической формулы Эйлера-Маклорена употреблять полученную нами в статье [34] формулу, в которой каждому входящему в сумму значению аргумента ставится в соответствие отрезок, для которого это значение является центром. Полученная формула позволила, в частности, построить компактные таблицы для специальных функций, используемых при оценивании параметров гамма-распределения, и вывести формулы для использования вне таблиц, при разработке государственного стандарта ГОСТ 11.011-83 (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) [1]. Многомерный аналог полученной формулы, в котором каждой точке целочисленной решетки ставился в соответствие прямоугольный параллелепипед, для которого эта точка являлась центром, позволил разработать оригинальный метод оценивания скорости сходимости распределений непараметрических статистик типа Колмогорова, Смирнова и омега-квадрат [34].

Поправки Шеппарда. С помощью формулы Эйлера — Маклорена групповые моменты $M(Y^k)$ могут быть выражены линейными функциями от «истинных» моментов $M(X^k)$ (в предположении, что остаточным членом можно пренебречь). Решая последовательно уравнения относительно «истинных» моментов $M(X^k)$, получаем [4, с. 395]:

$$\begin{aligned}
&M(X) = M(Y),\\
&M(X^2) = M(Y^2) - \frac{1}{12}h^2,\\
&M(X^3) = M(Y^3) - \frac{1}{4}M(Y)h^2,\\
&M(X^4) = M(Y^4) - \frac{1}{2}M(Y^2)h^2 + \frac{7}{240}h^4,\ldots
\end{aligned}$$

Таким образом, получены поправки к результатам расчетов по сгруппированным данным, позволяющие более точно оценить моменты исходного распределения. Эти поправки были впервые выведены В.Ф. Шеппардом в 1898 г. [36] и поэтому называются в статистической литературе *поправками Шеппарда*. Общая формула имеет вид [37]:

$$M(X^k) = \sum_{j=0}^{k} C_k^j \left(2^{1-j} - 1\right) B_j M(Y^{k-j}) h^j ,$$

где B_j — числа Бернулли [35, п. 449].

Хотя поправки Шеппарда получены еще в XIX веке, статистическая теория обработки сгруппированных данных продолжает развиваться. В данном разделе рассмотрена непараметрическая многомерная постановка, в которой границы группировки задаются статистиком. Параметрическую теорию развивал Г. Кулльдорф [38]. Модель группировки со случайным сдвигом начала координат изучал Н.А. Бодин [39–41].

Многомерная группировка. Дадим описание группировки данных в случае m-мерного случайного вектора $Z = (Z_1, Z_2, \ldots, Z_m)$ с плотностью распределения $p(x_1, x_2, \ldots, x_m)$. Пусть выборочные значения координаты Z_i группируются по интервалам длины h_i со средними точками (центрами) $a_{i,k} = a_i + kh_i$, $k = 0, \pm 1, \pm 2, \ldots$ Вектор $h = (h_1, h_2, \ldots, h_m)$ назовем шагом группировки. При вычислении выборочных характеристик предполагается, что все выборочные значения вектора Z, координаты которого удовлетворяют неравенствам:

$$a_{i,k(i)} - \frac{h_i}{2} < Z_i \le a_{i,k(i)} + \frac{h_i}{2}, \quad i = 1,2,\ldots,m,$$

совпадают с центром ($a_{1,k(1)}$, $a_{2,k(2)}$, …, $a_{m,k(m)}$) данного m-мерного прямоугольного параллелепипеда, т.е. фактически мы обрабатываем выборочные значения дискретного m-мерного вектора $W = (W_1, W_2, \ldots, W_m)$ с распределением:

$$P\{W = (a_{1,k(1)}, a_{2,k(2)}, \ldots, a_{m,k(m)})\} = \int_A p(x_1, x_2, \ldots, x_m)dx_1 dx_2 \ldots dx_m,$$

где

$$A = \left\{(x_1, x_2, \ldots, x_m):\ a_{i,k(i)} - \frac{h_i}{2} < x_i \le a_{i,k(i)} + \frac{h_i}{2},\ i = 1,2,\ldots,m\right\}.$$

Представляется естественным разложить разность между характеристиками вектора W и характеристиками вектора Z по степеням координат шага группировки и оценить возникающий при этом остаточный член. Таким образом, мы хотим получить обобщение поправок Шеппарда для моментов.

Будем рассматривать характеристики вида $Mf(Z)$, где f — достаточно гладкая функция.

Поправки на группировку для коэффициента корреляции. В качестве примера рассмотрим коэффициент корреляции $\rho(Z_1, Z_2)$ (точнее, линейный парный коэффициент корреляции Пирсона — см. раздел 6.1). Как известно, для удовлетворяющей некоторым условиям регулярности [4, п. 27.9] плотности $p(x_1, x_2)$ справедливы соотношения:

$$M(W_1) = M(Z_1),\ M(W_2) = M(Z_2),\ M(W_1W_2) = M(Z_1Z_2),$$
$$D(W_1) = D(Z_1) + \frac{h_1^2}{12},\ D(W_2) = D(Z_2) + \frac{h_1^2}{12}$$

при достаточно малых h_1 и h_2 с точностью до членов более высокого порядка по h_1 и h_2. Из приведенных соотношений и определения коэффициента корреляции $\rho(Z_1, Z_2)$ с помощью элементарных преобразований получаем, что:

$$\rho(W_1, W_2) - \rho(Z_1, Z_2) = \rho(Z_1, Z_2)\left(\frac{1}{\sqrt{1 + \frac{h_1^2}{12D(Z_1)}}\sqrt{1 + \frac{h_2^2}{12D(Z_2)}}} - 1\right)$$

с той же точностью. Поскольку при малых y:

$$\frac{1}{\sqrt{1+y}} = 1 - \frac{y}{2} + O(y^2)\,,$$

то при достаточно малых h_1 и h_2 с точностью до членов более высокого порядка:

$$\rho(W_1, W_2) - \rho(Z_1, Z_2) = \rho(Z_1, Z_2)\left(-\frac{h_1^2}{24D(Z_1)} - \frac{h_2^2}{24D(Z_2)}\right).$$

Из последней формулы вытекает, что:

$$\rho(Z_1, Z_2) = \rho(W_1, W_2)\left(1 + \frac{h_1^2}{24D(Z_1)} + \frac{h_2^2}{24D(Z_2)}\right).$$

с той же точностью. Воспользовавшись приведенными выше соотношениями для вторых моментов координат двумерного вектора, получаем окончательную формулу:

$$\rho(Z_1, Z_2) = \rho(W_1, W_2)\left(1 + \frac{h_1^2}{24D(W_1)} + \frac{h_2^2}{24D(W_2)}\right),$$

в которой поправка на группировку определяется только по сгруппированным данным.

Полученная поправка на группировку для коэффициента корреляции рекомендуется для использования в практических вычислениях. Отметим, что группировка приводит к уменьшению коэффициента корреляции (по абсолютной величине):

$$|\,\rho(Z_1, Z_2)\,| > |\,\rho(W_1, W_2)\,|\,.$$

Поправка на группировку для коэффициента корреляции была получена в заметке [42]. Эта заметка возникла в порядке полемики со статьей [43], в которой без какого-либо основания утверждалось, что при группировке коэффициент корреляции увеличивается. Несмотря на элементарность выкладок, Поправка на группировку для коэффициента корреляции не была ранее известна, что и отмечено в реферате заметки [42] в основном реферативном журнале по математике *Mathematical Review*.

Асимптотические поправки на группировку. Перейдем к общему случаю. Результаты были получены в статье [44]. Полное изложение дано в [45].

Будем использовать следующий многомерный аналог классической формулы Эйлера-Маклорена:

$$\sum_{v\in Z^m} g(a_1+v_1h_1,\ldots,a_m+v_mh_m)=\int_{R^m} g(a_1+x_1h_1,\ldots,a_m+x_mh_m)dx_1\ldots dx_m+$$
$$+\sum_{t=1}^{m}(-1)^t\int_{R^m}\sum_{\mu\in C(t,m)}\left(\prod_{i\in\mu}P_1(x_i)\frac{\partial}{\partial x_i}\right)g(a_1+x_1h_1,\ldots,a_m+x_mh_m)dx_1\ldots dx_m,$$

где Z^m — множество точек с целыми координатами m-мерного пространства R^m; функция $g:R^m\to R^1$ абсолютно интегрируема; $C(t,m)$ — совокупность всех подмножеств мощности t множества $\{1, 2, \ldots, m\}$ и:

$$P_1(q)=\sum_{j=1}^{\infty}\frac{\sin(2\pi qj)}{j\pi}.$$

Справедлив следующий многомерный аналог полученного нами ранее обобщения одномерной классической формулы Эйлера-Маклорена:

$$Mf(W)=Mf(Z)+\sum_{i=1}^{m}\frac{h_i^2}{24}\int_{R^m}p(x)\frac{\partial^2 f(x)}{\partial x_i^2}dx+\sum_{i=1}^{m}\frac{h_i^4}{1920}\int_{R^m}p(x)\frac{\partial^4 f(x)}{\partial x_i^4}dx+$$
$$+\sum_{\substack{i,j=1\\ i\neq j}}^{m}\frac{h_i^2h_j^2}{3456}\int_{R^m}p(x)\frac{\partial^4 f(x)}{\partial x_i^2\partial x_j^2}dx+\ldots$$

(в этой формуле $x = (x_1, x_2, \ldots, x_m)$, соответственно, $dx = dx_1dx_2\ldots dx_m$, используются введенные ранее обозначения). Описан закон получения членов разложения и указана оценка остаточного члена [44]. Однако в общем случае они весьма громоздки, поэтому ограничимся несколькими доведенными до числа примерами.

Пусть $m = 2$, вектор (Z_1, Z_2) имеет нормальное распределение. Тогда:

$$M(Z_j)=M(W_j)-R(M(Z_j)),\ \ D(W_j)=D(Z_j)+\frac{h_j^2}{12}+R(D(Z_j)),\ \ j=1,2,$$
$$\mathrm{cov}(W_1,W_2)=\mathrm{cov}(Z_1,Z_2)+R.$$

На основе описанных выше общих результатов можно получить точные формулы и оценки остаточных членов в этих формулах. Оказывается, остаточные члены $R(M(Z_j))$ и $R(D(Z_j))$, $j = 1, 2$, весьма малы даже при достаточно большом шаге группировки. Так,

$$R(M(Z_j)) = \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \exp\left\{-2\left(\frac{j\pi\sigma_j}{h_j}\right)^2\right\} \sin\left(2j\pi \frac{M(Z_j)-a_j}{h_j}\right)$$

и $|R(M(Z_j))|$ не превосходит $10^{-9}\sigma_j$ при $h_j \le \sigma_j$ и не превосходит $5\times10^{-3}\sigma_j$ при $h_j \le 2\sigma_j$, где $\sigma_j^2 = D(Z_j)$, а a_j — начало отсчета по j-й координате, $j = 1,2$.

Полученные оценки зависят от $M(Z_j) - a_j$. Так, при $a_j - M(Z_j) = h_j/4$ относительная погрешность определения истинной дисперсии по сгруппированной с использованием поправок Шеппарда, т.е.

$$\left|\frac{D(Z_j) - D(W_j) + \frac{h_j^2}{12}}{D(Z_j)}\right| = \left|\frac{R(D(Z_j))}{D(Z_j)}\right|,$$

не превосходит $6{,}8\times10^{-2}$ при $h_j \le 3\sigma_j$, меньше $0{,}25\times10^{-2}$ при $h_j \le 2\sigma_j$ и $6{,}8\times10^{-11}$ при $h_j \le \sigma_j$. Таким образом (это подтверждается равномерной оценкой остаточного члена), при шаге группировки, равном среднеквадратическому отклонению, формула:

$$D(W_j) = D(Z_j) + \frac{h_j^2}{12}$$

становится практически точной.

Иным оказывается положение при расчете по сгруппированным данным коэффициента ковариации. При малых и средних значениях коэффициент корреляции ρ (при $|\rho| \le 0{,}8$) относительная погрешность невелика даже при довольно большом шаге группировки (так, $\left|\frac{R}{\rho\sigma_1\sigma_2}\right| \le 10^{-3}$ при $h_1 = \sigma_1$ и $h_2 = \sigma_2$. При больших коэффициентах корреляции, если ставится цель добиться достаточно хорошего приближения, шаг группировки необходимо уменьшать, поскольку при $|\rho| \to 1$ основной член в оценке относительной погрешности таков:

$$\left|\frac{R}{\rho\sigma_1\sigma_2}\right| \approx 4\exp\left\{-2(1-|\rho|)\pi^2\left[\left(\frac{\sigma_1}{h_1}\right)^2+\left(\frac{\sigma_2}{h_2}\right)^2\right]\right\}.$$

В фундаментальных и прикладных научных исследованиях часто оказывается необходимым использовать коэффициент корреляции ρ в случае произвольного распределения случайного вектора при естественном и обычно выполняющемся условии существования достаточного числа моментов и гладкости плотности вероятности. На основе описанного выше общего подхода получена следующая приближенная формула для выражения истинного коэффициента корреляции $\rho(Z)=\rho(Z_1,Z_2)$ через «сгруппированный» коэффициент корреляции $\rho(W)=\rho(W_1,W_2)$ и «сгруппированные» дисперсии $D(W_1)$ и $D(W_2)$:

$$\rho(Z_1,Z_2)=\rho(W_1,W_2)\left(1+\frac{h_1^2}{24D(W_1)}+\frac{h_2^2}{24D(W_2)}-\frac{1}{288}\frac{h_1^2h_2^2}{D(W_1)D(W_2)}\right)=\rho_1(W_1,W_2).$$

Эта формула уточняет ранее приведенную формулу, полученную элементарными средствами. Оценена относительная погрешность определения коэффициента корреляции:

$$\left|\frac{\rho(Z_1,Z_2)-\rho_1(W_1,W_2)}{\rho(Z_1,Z_2)}\right|\le R_D+\frac{1}{120}\left(\frac{h_1^2}{D(W_1)}+\frac{h_2^2}{D(W_2)}\right)+\left|\frac{R}{\rho(Z_1,Z_2)\sigma_1\sigma_2}\right|,$$

где R_D — максимум относительных погрешностей в определении истинных дисперсий $D(Z_1)$ и $D(Z_2)$ через «сгруппированные» дисперсии $D(W_1)$ и $D(W_2)$ с использованием поправок Шеппарда, R — абсолютная погрешность при определении истинного коэффициента ковариации через «сгруппированный» (см. выше).

Подводя итоги раздела, следует отметить, что хотя теория статистического анализа сгруппированных данных в настоящее время недостаточно развита и систематизирована, принципиальных трудностей на пути расчета поправок на группировку нет. Во многих случаях шаг группировки порядка среднего квадратического отклонения исходных случайных величин дает с учетом поправок вполне приемлемую точность. По сравнению с другими видами погрешностей, в частности, рассматриваемыми в статистике интервальных данных ([46, глава 12]), группировка сравнительно слабо влияет на свойства статистических процедур.

Литература

1. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения = Applied statistics. Regulations for determinations of estimates and confidence limits for parameters of gamma distribution : национальные стандарт Союза ССР : издание официальное : утвержден и введен в действие постановлением Государственного комитета СССР по стандартам от 27 июня 1983 г. / А.И. Орлов, Н.Г. Миронова, М.Б. Невельсон и др. — Москва : Изд-во стандартов, 1984. — 53 с. — Переиздание : Москва : Изд-во стандартов, 1985. — 50 с. (В настоящее время отменен как нормативный документ, но может использоваться как научная публикация.)

2. *Рао, С.Р.* Линейные статистические методы и их применения / С.Р. Рао. — Москва : Наука, 1968. — 548 с.

3. *Вентцель, Е.С.* Теория вероятностей / Е.С. Вентцель. — Москва : Наука, 1964. — 576 с.

4. *Крамер, Г.* Математические методы статистики / Г. Крамер. — Москва : Мир, 1975. — 648 с.

5. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 396 с.

6. *Лумельский, Я.П.* К вопросу сравнения несмещенных и других оценок / Я.П. Лумельский // Прикладная статистика. — Москва : Наука, 1983. — С. 316–319.

7. ГОСТ 11.010-81. Прикладная статистика. Правила определения оценок параметров и доверительных границ для биномиального и отрицательного биномиального распределений = Applied statistics. Point and interval estimators for parameters of binomial and negative binomial distribution : национальные стандарт Союза ССР : издание официальное : утвержден и введен в действие постановлением Государственного комитета СССР по стандартам от 30 марта 1981 г. — Москва : Изд-во стандартов, 1982. — 32 с. (В настоящее время отменен как нормативный документ, но может использоваться как научная публикация.)

8. *Сатаров, Г.А.* Новая статистическая модель парных сравнений / Г.А. Сатаров, Д.С. Шмерлинг // Экспертные оценки в задачах управления. — Москва : Изд-во Института проблем управления АН СССР, 1982. — С. 67–79.

9. *Лапига, А.Г.* Многокритериальные задачи управления качеством: построение прогноза качества в балльной шкале / А.Г. Лапига // Заводская лаборатория. — 1983. — Т. 49. — № 7. — С. 55–59.

10. *Закс, Ш.* Теория статистических выводов / Ш. Закс. — Москва : Мир, 1975. — 776 с.

11. *Бахмутов, В.О.* Использование метода максимального правдоподобия для оценки однородности результатов усталостных испытаний / В.О. Бахмутов, Л.Н. Косарев // Заводская лаборатория. — 1986. — Т. 52. — № 5. — С. 52–57.

12. *Резникова, А.Я.* Оценивание параметров вероятностных моделей парных и множественных сравнений / А.Я. Резникова, Д.С. Шмерлинг // Статистические методы оценивания и проверки гипотез : межвузовский сборник научных трудов. — Пермь : Изд-во Пермского госуниверситета, 1984. — С. 110–120.

13. *Ибрагимов, И.А.* Асимптотическая теория оценивания / И.А. Ибрагимов, Р.З. Хасьминский. — Москва : Наука, 1979. — 528 с.

14. *Орлов, А.И.* О нецелесообразности использования итеративных процедур нахождения оценок максимального правдоподобия / А.И. Орлов // Заводская лаборатория. — 1986. — Т. 52. — № 5. — С. 67–69.

15. *Боровков, А.А.* Математическая статистика : учебное пособие для вузов / А.А. Боровков. — Москва : Наука, 1984. — 472 с.

16. *Кендалл, М.Дж.* Статистические выводы и связи / М.Дж. Кендалл, А. Стьюарт — Москва : Наука, 1973. — 900 с.

17. *Орлов, А.И.* Одношаговые оценки для параметров гамма-распределения / А.И. Орлов, Н.Г. Миронова // Надежность и контроль качества. — 1988. — № 9. — С. 18–22.

18. *Орлов, А.И.* Интервальная статистика: метод максимального правдоподобия и метод моментов / А.И. Орлов // Статистические методы оценивания и проверки гипотез : межвузовский сборник научных трудов. — Пермь : Изд-во Пермского государственного университета, 1995. — С. 114–124.

19. *Орлов, А.И.* Нечисловая статистика / А.И. Орлов. — Москва : Экзамен, 2007.

20. *Петрович, М.Л.* Статистическое оценивание и проверка гипотез на ЭВМ / М.Л. Петрович, М.И. Давидович. — Москва : Финансы и статистика, 1989. — 191 с.

21. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — Изд. 2-е, испр. и доп. — Москва : Экзамен, 2003. — 576 с.

22. Келли Дж. Общая топология / Дж. Келли. — Москва : Наука, 1968. — 384 с.

23. *Орлов, А.И.* Асимптотика решений экстремальных статистических задач / А.И. Орлов // Анализ нечисловых данных в системных исследованиях : сборник трудов. — Вып. 10. — Москва : Всесоюзный научно-исследовательский институт системных исследований, 1982. — С. 4–12.

24. *Орлов, А.И.* Общий взгляд на статистику объектов нечисловой природы / А.И. Орлов // Анализ нечисловой информации в социологических исследованиях. — Москва : Наука, 1985. — С. 58–92.

25. *Орлов, А.И.* Некоторые неклассические постановки в регрессионном анализе и теории классификации / А.И. Орлов // Программно-алгоритмическое обеспечение анализа данных в медико-биологических исследованиях. — Москва : Наука, 1987. — С. 27–40.

26. *Орлов, А.И.* Статистика объектов нечисловой природы и экспертные оценки / А.И. Орлов // Экспертные оценки. Вопросы кибернетики. — Вып. 58. — Москва : Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1979. — С. 17–33.

27. *Гнеденко, Б.В.* Курс теории вероятностей : учебник / Б.В. Гнеденко. — 7-е изд., испр. — Москва : Эдиториал УРСС, 2001. — 320 с.

28. *Смоляк, С.А.* Устойчивые методы оценивания: Статистическая обработка неоднородных совокупностей / С.А. Смоляк, Б.П. Титаренко. — Москва : Статистика, 1980. — 208 с.

29. Хьюбер П. Робастность в статистике / П. Хьюбер. — Москва : Мир, 1984. — 304 с.

30. Робастность в статистике. Подход на основе функций влияния / Ф. Хампель, Э. Рончетти, П. Рауссеу, В. Штаэль. — Москва : Мир, 1989. — 512 с.

31. *Эльясберг П.Е.* Измерительная информация. Сколько ее нужно, как ее обрабатывать? / П.Е. Эльясберг. — Москва : Наука, 1983. — 208 с.

32. *Пасхавер, И.С.* Общая теория статистики: Для программированного обучения : учебное пособие / И.С. Пасхавер, А.Л. Яблочник. — Изд. 2-е, перераб. и доп. — Москва : Финансы и статистика, 1983. — 432 с.

33. ГОСТ 11.006-74 (СТ СЭВ 1190-78). Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим = Applied statistics. The tests for goodness of fit distributions of empirical : национальные стандарт Союза ССР : издание официальное : утвержден и введен в действие постановлением Государственного комитета СССР по стандартам от 20 декабря 1974 г. — Москва : Изд-во стандартов, 1975. — 26 с. (Отменен в 1987 г.)

34. *Орлов, А.И.* Скорость сходимости распределения статистики Мизеса — Смирнова / А.И. Орлов // Теория вероятностей и ее применения. — 1974. — Т. 19. — № 4. — С. 766–786.

35. *Фихтенгольц, Г.М.* Курс дифференциального и интегрального исчисления / Г.М. Фихтенгольц. — Т. 2. — Москва : Наука, 1966. — 800 с.

36. *Sheppard, W.F.* On the calculation of the most probable values of frequency constants for data arranged according to equidistant divisions of a scale /

W.F. Sheppard // Proceedings of the London Mathematical Society. — 1898. — V. 29. — P. 353.

37. *Wold, H.* Sulla correzione di Sheppard / H. Wold // Giorn. Ist. Italiano d. Attari. — 1934. — V. 5. — P. 304.

38. *Кулдорф, Г.* Введение в теорию оценивания по группированным и частично группированным выборкам / Г. Кулдорф. — Москва : Наука, 1966. — 196 с.

39. *Бодин, Н.А.* Об ошибках округления при многомерных измерениях / Н.А. Бодин // Труды Математического института им. В.А. Стеклова. — Т. 79. — Москва : Ленинград : Наука, 1965. — С. 76–105.

40. *Бодин, Н.А.* Оценка параметров распределения по группированным выборкам / Н.А. Бодин // Труды Математического института им. В.А. Стеклова. — Т. 111. — Ленинград : Наука, 1970. — С. 110–154.

41. *Бодин, Н.А.* Об ошибках округления при измерениях с помощью случайно сдвигающейся шкалы / Н.А. Бодин // Труды Математического института им. В.А. Стеклова. — Т. 111. — Ленинград : Наука, 1970. — С. 155–162.

42. *Орлов, А.И.* Поправка на группировку для коэффициента корреляции / А.И. Орлов // Экономика и математические методы. — 1980. — № 4. — С. 800–801.

43. *Свободин, В.А.* К вопросу о корреляционных моделях по усредненным данным / В.А. Свободин, И.К. Сирожитдинов // Экономика и математические методы. — 1980. — № 4. — С. 796–800.

44. *Орлов, А.И.* О поправках на группировку // Прикладной многомерный статистический анализ / А.И. Орлов, И.В. Орловский. — Москва : Наука, 1978. — С. 339–342.

45. *Орловский, И.В.* Поправки на группировку и остаточные члены для некоторых статистик : дипломная работа (руководитель — А.И. Орлов) / И.В. Орловский ; Механико-математический ф-т МГУ им. М.В. Ломоносова. — Москва, 1977.

46. *Орлов, А.И.* Прикладная статистика : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 672 с.

47. *Орлов, А.И.* Оценивание параметров: одношаговые оценки предпочтительнее оценок максимального правдоподобия / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 109. — С. 208–237.

48. *Орлов, А.И.* Предельная теория решений экстремальных статистических задач / А.И. Орлов // Научный журнал КубГАУ. — 2017. — № 133. — С. 579–600.

49. *Орлов, А.И.* Новый подход к изучению устойчивости выводов в математических моделях / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 100. — С. 146–176.

50. *Орлов, А.И.* Свойства общей схемы устойчивости / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 161. — С. 121–149.

51. *Орлов, А.И.* Статистическое оценивание для сгруппированных данных / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 98. — С. 1–13.

Контрольные вопросы и задачи

1. Чем задачи оценивания параметров распределения отличаются от задач оценивания характеристик распределения?

2. С помощью метода линеаризации обоснуйте вид асимптотических дисперсий оценок метода моментов для параметров гамма-распределения (табл. 3 раздела 3.1).

3. Почему одношаговые оценки предпочтительнее оценок максимального правдоподобия?

4. Как связаны законы больших чисел в пространствах произвольной природы и утверждения об асимптотическом поведении решений экстремальных статистических задач?

5. Как соотносятся параметрическая регрессия и непараметрическая регрессия?

6. Сопоставьте различные постановки задач изучения робастности статистических процедур.

7. В чем основная идея поправок Шеппарда?

Темы докладов, рефератов, исследовательских работ

1. Квантильные оценки.

2. Минимизация расстояния как способ построения оценок параметров.

3. Одношаговые оценки параметров распределения Вейбулла — Гнеденко.

4. Оптимизационные постановки основных задач прикладной статистики.

5. Роль функции влияния при изучении робастности в модели засорения Тьюки — Хубера.

6. На основе четырех описанных в разделе 3.4 моделей с помощью общей схемы устойчивости (глава 14) опишите спектр возможных постановок задач устойчивости статистических процедур.

7. Различные подходы к анализу сгруппированных данных.

ГЛАВА 4. ПРОВЕРКА ГИПОТЕЗ

4.1. Метод моментов проверки гипотез

К методу моментов относят все статистические процедуры, основанные на использовании выборочных моментов и функций от них. Метод моментов оценивания параметров распределения рассмотрен в главе 3. В непараметрической статистике на основе выборочных моментов проводится точечное и интервальное оценивание характеристик распределения, таких, как математическое ожидание, дисперсия, среднее квадратическое отклонение, коэффициент вариации (глава 5.1). Для проверки гипотез в непараметрической статистике также используется метод моментов. Примером является критерий Крамера-Уэлча, предназначенный для проверки равенства математических ожиданий по двум независимым выборкам (глава 5.1).

В практике применения статистических методов (согласно классическим схемам) довольно часто возникает необходимость проверки гипотезы о том, что функция распределения результатов наблюдений $X_1, X_2, \ldots , X_n$ принадлежит параметрическому семейству распределений $\{F(x, \theta), \theta \in \Theta\}$, где $\Theta \subseteq R^k$. Как проверять эту гипотезу?

Давно разработан универсальный метод — критерий минимума хи-квадрат [1]. Однако у него имеется существенный недостаток — необходимость группирования наблюдений, что приводит к потере информации. Как хорошо известно [2], это приводит к существенному снижению мощности критерия минимума хи-квадрат по сравнению с критериями типа Колмогорова и типа омега-квадрат. Кроме того, нахождение минимума статистики хи-квадрат — достаточно сложная вычислительная процедура. Поэтому иногда вместо оценок, получаемых при указанной оптимизации, подставляют оценки максимального правдоподобия или какие-либо еще. Такая замена приводит к тому, что распределение рассматриваемой статистики существенно отличается от классического, причем различие не исчезает при росте объема выборки. Предложенная член-корр. АН СССР Л.Н. Большевым и проф. М.С. Никулиным [3] модификация критерия минимума хи-квадрат не снимает недостатков, связанных с группированием и необходимостью существенной вычислительной работы.

Общий подход, основанный на дистанционном методе, предложен Дж. Вольфовицем (США) в 1950-х годах. Согласно этому методу, следует основываться на том или ином расстоянии между эмпирической функцией распределения и параметрическим семейством распределений (как многообразием в пространстве всех функций распределения). Конкретная реализация этого подхода приводит к критериям типа Колмогорова и типа омега-квадрат. Однако для каждого конкретного параметрического семейства приходится разрабатывать самостоятельную теорию и рассчитывать только ему соответствующие предельные и точные распределения [4, 5]. Предельные распределения найдены лишь для нескольких семейств, а точных почти ничего не известно. До сих пор часто делают ошибку, применяя для произвольных семейств предельные распределения, найденные для проверки согласия с фиксированным распределением (см. подробности в главе 2.4).

Отметим, что критерии минимума хи-квадрат и аналогичные им не являются состоятельными, поскольку вероятности попадания в области группирования не задают однозначно функцию распределения. С этим недостатком можно бороться, увеличивая число интервалов группирования вместе с ростом объема выборки, однако на этом пути еще не выработаны рекомендации, пригодные для широкого практического использования. Критерии типа Колмогорова и типа омега-квадрат — состоятельные, т.е. любую альтернативную функцию распределения, не входящую в рассматриваемое параметрическое семейство, они отвергают с вероятностью, стремящейся к 1 при росте объема выборки.

Для конкретности обсудим проверку согласие результатов наблюдений с трехпараметрическим семейством гамма-распределений с плотностями:

$$f(x;a,b,c)=\begin{cases}\dfrac{1}{\Gamma(a)}(x-c)^{a-1}b^{-a}\exp\left[-\dfrac{x-c}{b}\right], & x\ge c,\\ 0, & x<c.\end{cases} \qquad (1)$$

Здесь a>2 — параметр формы, b>0 — параметр масштаба и c — параметр сдвига, $\Gamma(a)$ — одна из используемых в математике специальных функций, так называемая «гамма-функция». Критерий минимума хи-квадрат имеет указанные выше недостатки. Критерии типа Колмогорова и типа омега-квадрат для этого случая не разработаны.

В подобных ситуациях целесообразно строить критерии согласия на основе функций от выборочных моментов, т.е. пользоваться методом моментов. Для оценивания параметров метод моментов хорошо известен и обычно рассматривается в учебной литературе по теории вероятностей и математической статистике. Реализацией метода моментов для проверки нормальности являются известные критерии асимметрии и эксцесса [6].

Пример 1. Если случайная величина X имеет нормальное распределение с математическим ожиданием a и дисперсией σ^2, то, как известно [6],

$$\delta = \frac{M|X-a|}{\sigma} = \sqrt{\frac{2}{\pi}} = 0{,}79788, \ \gamma_1 = \frac{M(X-a)^3}{\sigma^3} = 0, \ \beta_1 = \frac{M(X-a)^4}{\sigma^4} = 3,$$

где δ — нормированное среднее абсолютное отклонение, γ_1 — коэффициент асимметрии и (β_1 — 3) — коэффициент эксцесса. Таким образом, если выборочные оценки указанных моментных отношений существенно отличаются от соответствующих теоретических значений, то следует признать, что распределение результатов наблюдений отлично от нормального. Так как указанные выше значения моментных отношений могут приниматься и для распределений, отличных от нормальных, то близость выборочных значений к только что выписанным не обязательно свидетельствует о нормальности распределения результатов наблюдений. Критерии, полученные методом моментов, служат не столько для проверки нормальности, сколько для выявления отклонений распределения от нормального, или, точнее, для проверки гипотез $\delta \neq \sqrt{2/\pi}$, $\gamma_1 \neq 0$, $\beta_1 \neq 3$. Рассматриваемые критерии построены на основе выборочных моментных отношений:

$$d = \frac{1}{ns}\sum_{k=1}^{n}|X_k - \overline{X}|, \ g_1 = \frac{1}{ns^3}\sum_{k=1}^{n}(X_k - \overline{X})^3, \ b_1 = \frac{1}{ns^4}\sum_{k=1}^{n}(X_k - \overline{X})^4 .$$

Здесь, как обычно, $\overline{X}$ — выборочное среднее арифметическое и s^2 — выборочная дисперсия, соответственно, s — выборочное среднее квадратическое отклонение. Как вытекает из результатов главы 14, все три статистики являются асимптотически нормальными. Выражения для параметров их асимптотических распределений приведены в [6]. Процентные точки распределений рассматриваемых выборочных моментных отношений при конечных объемах выборки найдены в предположении нормальности результатов наблюдений [6].

Как и критерии минимума хи-квадрат, критерии метода моментов никогда не являются состоятельными. Однако они, как и в случае критериев асимметрии и эксцесса, позволяют в ряде случаев отвергнуть гипотезу согласия. Использование несостоятельных критериев часто встречается в прикладной статистике. Отметим, например, что применение критерия Вилкоксона для проверки гипотезы однородности двух выборок широко распространено, хотя против общей альтернативы он является несостоятельным (см. главу 5.3).

Критерии метода моментов основаны на использовании функций от выборочных моментов, имеющих асимптотически нормальные распределения, параметры которых легко могут быть вычислены по методике, описанной в главе 14. Метод моментов по сравнению с другими методами проверки согласия требует существенно меньше вычислений (число операций пропорционально объему выборки). Поэтому он может быть рекомендован для использования при проверке согласия с семействами распределений, для которых не разработаны более совершенные методы, а также в качестве быстрого (экспрессного) метода. Что же касается хорошо изученных семейств, например, нормального, то основанные на использовании моментов критерии асимметрии и эксцесса применять для проверки нормальности нецелесообразно. Судя по специальным исследованиям, следует рекомендовать критерий W Шапиро — Уилка.

Продемонстрируем применение метода моментов на примере проверки гипотезы согласия с двухпараметрическим семейством гамма-распределений без сдвига, т.е. выделяемого из семейства (1) условием c=0. Поскольку для трехпараметрического семейства гамма-распределений (1):

$$M(X) = ab + c,\ D(X) = ab^2,\ \mu_3 = M(X - M(X))^3 = 2ab^3,$$

то при справедливости гипотезы H_0: $c = 0$ выполнено соотношение:

$$\frac{M(X)\mu_3}{2\sigma^4} - 1 = 0. \qquad (2)$$

Для специалистов по техническим наукам большое значение имеет альтернативная гипотеза:

$$H_1: c > 0.$$

В частности, она связана с дискуссией о выборе нормируемых показателей надежности технических устройств. Альтернативная гипотеза соответствует предположению, что в течение некоторого времени (до момента $c > 0$) отказы невозможны, а нулевая — с отрицанием этого предположения и признанием того, что отказы возможны в любой момент.

При справедливости альтернативной гипотезы:

$$\frac{M(X)\mu_3}{2\sigma^4} - 1 = \frac{c}{ab} > 0,$$

поэтому для проверки гипотезы согласия в рассматриваемой постановке целесообразно использовать критерий со статистикой:

$$Z = \frac{\overline{X} m_3}{2s^4} - 1 .$$

С помощью описанной в главе 14 методики вычисления предельного распределения функции от выборочных моментов можно установить, что при $n \to \infty$ распределение статистики $\sqrt{n}Z$ сходится к нормальному, причем при справедливости нулевой гипотезы, т.е. соотношения (2), асимптотическое распределение имеет нулевое математическое ожидание и дисперсию:

$$\frac{1}{2a}(3a^2 + 13a + 10) . \qquad (3)$$

Поскольку параметр формы a неизвестен статистику, необходимо в выражении (3) заменить a на его состоятельную оценку, например, на оценку метода моментов (см. главу 3.1):

$$a^* = \frac{\left(\overline{X}\right)^2}{s^2} .$$

Рассмотрим критерий с критической областью вида:

$$\left\{ Z: \ Z > u(1-\alpha) \frac{3(a^*)^2 + 13a^* + 10}{2a^* \sqrt{n}} \right\}, \qquad (4)$$

где $u(1 - \alpha)$ — квантиль порядка $1 - \alpha$ стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. При $n \to \infty$ уровень значимости этого критерия стремится к α.

Если альтернативная гипотеза является двусторонней, т.е. $H_1^{'}: c \neq 0$, то аналогично строится двусторонняя критическая область.

Критерий (4) состоятелен против альтернативы H_1: $c > 0$, а также против непараметрической альтернативы:

$$\frac{M(X)\mu_3}{2\sigma^4} > 1 ,$$

в которой не предполагается, что функция распределения элементов выборки имеет гамма-распределение (1) с какими-либо конкретными значениями параметров, но не является состоятельным против общей альтернативы.

Пример 2. Применим критерий (4) для проверки согласия с гамма-распределением при $c = 0$, т.е. с двухпараметрическим семейством, данных о наработке $n = 50$ резцов до предельного состояния (в часах), приведенных в табл. 2 раздела 3.1.

Для рассматриваемых данных $\overline{X} = 57{,}88$, $s^2 = 663{,}00$, выборочный третий центральный момент $m_3 = 14927{,}91$, откуда $Z = -0{,}01719$. При этом $a^* = 5{,}05$, и потому:

$$\frac{3(a^*)^2 + 13a^* + 10}{2a^*\sqrt{n}} = 0{,}4488\,.$$

Следовательно, гипотеза согласия рассматриваемых данных с двухпараметрическим гамма-распределением не отвергается на любом из обычно используемых уровней значимости, как для односторонней критической области, так и для двухсторонней.

4.2. Неустойчивость параметрических методов отбраковки выбросов

При обработке реальных технических, экономических, медицинских и иных данных, полученных в процессе наблюдений, измерений, расчетов, иногда один или несколько результатов наблюдений резко выделяются, т.е. далеко отстоят от основной массы данных. Такие резко выделяющиеся результаты наблюдений часто считают содержащими грубые погрешности, соответственно называют промахами или выбросами. В рассматриваемых случаях возникает естественная мысль о том, что подобные наблюдения не относятся к изучаемой совокупности, поскольку содержат грубую погрешность, а получены они в результате ошибки, промаха. В справочнике по метрологии об этом явлении говорится так: «Грубые погрешности и промахи возникают из-за ошибок или неправильных действий оператора (его психофизиологического состояния, неверного отсчета, ошибок в записях или вычислениях, неправильного включения приборов и т.п.). А также при резких кратковременных изменениях условий проведения измерений (в результате вибрации, поступления холодного воздуха, толчка прибора оператором и т.п.). Если грубые погрешности и промахи обнаруживают в процессе измерений, то результаты, содержащие их, отбрасывают. Однако чаще всего их выявляют только при окончательной обработке результатов измерений с помощью специальных критериев оценки грубых погрешностей» [7, с. 46–47].

Есть два подхода к обработке данных, которые могут быть искажены грубыми погрешностями и промахами:

1) отбраковка резко выделяющихся результатов наблюдений, т.е. обнаружение наблюдений, искаженных грубыми погрешностями и промахами, и исключение их из дальнейшей статистической обработки;

2) применение устойчивых (робастных) методов обработки данных, на результаты работы которых мало влияет наличие небольшого числа грубо искаженных наблюдений (см. подраздел 3.4).

Обсудим методы отбраковки. Наиболее изучена ситуация, когда результаты наблюдений — числа x_1, x_2, ..., x_n, среди них резко выделяется один результат наблюдения, для определенности, максимальный $x_{\max}$.

Простейшая вероятностно-статистическая модель такова [6]. При нулевой гипотезе H_0 результаты наблюдения x_1, x_2, ..., x_n рассматриваются как реализация независимых одинаково распределенных случайных величин числа X_1, X_2, ..., X_n с функцией распределения $F(x)$. При альтернативной гипотезе H_1 случайные величины X_1, X_2, ..., X_n также независимы, X_1, X_2, ..., X_{n-1} имеют распределение $F(x)$, а X_n — распределение $G(x)$, оно «существенно сдвинуто вправо» относительно $F(x)$, например, $G(x)=F(x - A)$, где A достаточно велико. Если альтернативная гипотеза справедлива, то при $A \to \infty$ вероятность равенства:

$$X_n = \max(X_1, X_2, ..., X_n)$$

стремится к 1, поэтому естественно применять решающее правило следующего вида:

$$\begin{aligned} &\text{если } x_{max}. > d, \text{ то принять } H_1, \\ &\text{если } x_{max}. \leq d, \text{ то принять } H_0, \end{aligned} \qquad (1)$$

где d — параметр решающего правила, который следует определять из вероятностно-статистических соображений.

При справедливости нулевой гипотезы:

$$P\{\max_{1 \leq i \leq n} X_i \leq d\} = \{F(d)\}^n.$$

Статистический критерий проверки гипотезы H_0, основанный на решающем правиле вида (1), имеет уровень значимости α, если:

$$P\{\max_{1\le i\le n} X_i > d\} = 1 - \{F(d)\}^n = \alpha,$$

т.е.

$$F(d) = \sqrt[n]{1-\alpha}. \qquad (2)$$

Из соотношения (2) определяют граничное значение $d = d(\alpha,\ n)$ в решающем правиле (1).

При больших n и малых α согласно известным результатам математического анализа:

$$F(d) = \sqrt[n]{1-\alpha} = 1 - \frac{\alpha}{n} + O\left(\frac{\alpha^2}{n^2}\right), \qquad (3)$$

поэтому в качестве хорошего приближения к $d(\alpha,\ n)$ рассматривают $(1 - \alpha/n)$ — квантиль распределения $F(x)$.

Пусть правило отбраковки задано в соответствии с соотношениями (1) и (2) с некоторой функцией распределения F, однако выборка берется из функции распределения G, мало отличающейся от F в смысле расстояния Колмогорова:

$$\rho(F,G) = \sup_x |F(x) - G(x)| \le \delta. \qquad (4)$$

С помощью соотношения (3) получаем, что величина $\gamma = G(d)$ для d из уравнения (2) находится между $\gamma_1 = \max(0,\ 1 - \frac{\alpha}{n} - \delta)$ и $\gamma_2 = \min(1 - \frac{\alpha}{n} + \delta,\ 1)$. Таким образом, уровень значимости критерия, построенного для F, при применении к наблюдениям из G есть $1\text{-}\gamma^n$ и может принимать любые значения в отрезке $[1-\gamma_2;\ 1-\gamma_1]$.

В частности, при δ = 0,01, α = 0,05, n = 5 возможные значения уровня значимости заполняют отрезок [0; 0,1], т.е. уровень значимости может быть в 2 раза выше номинального. А если n возрастает до 30, то максимальный уровень значимости есть 0,297, т.е. почти в 6 раз выше номинального. При дальнейшем росте n верхняя граница для уровня значимости, как нетрудно видеть, приближается к 1.

Рассмотрим и другой вопрос — насколько правило отбраковки с уровнем значимости α для G может отличаться от такового для F при справедливости неравенства (4). С использованием соотношения (3) заключаем, что из:

$$G(d) = 1 - \frac{\alpha}{n} \qquad (5)$$

следует, что $\gamma_1 \le F(d) \le \gamma_2$, где γ_1 и γ_2 выписаны выше. Решение уравнения (5) может принимать любое значение в отрезке $[F^{-1}(\gamma_1); F^{-1}(\gamma_2)]$. В частности, при α=0,05 и n = 5 для стандартного нормального распределения F имеем $d(\alpha,\ n)$ = 2,319, при δ=0,01 решение уравнения (5) может принимать любое значение в отрезке [2,054; + ∞], при δ =0,005 — любое значение в [2,170; 2,576].

При использовании любого другого расстояния между функциями распределения выводы о неустойчивости правил отбраковки также справедливы. Отметим, что проведенные рассмотрения выполнены в рамках «общей схемы устойчивости» (см. главу 14.6).

Рассмотренные примеры показывают, что при конкретном значении δ = 0,01 в неравенстве (4) весьма неустойчивы как уровни значимости при фиксированном правиле отбраковки, так и параметр d правила отбраковки при фиксированном уровне значимости. Обсудим, насколько реалистично определение функции распределения с точностью $\delta \le 0{,}01$.

Есть два подхода к определению функции распределения результатов наблюдений: эвристический подбор с последующей проверкой с помощью критериев согласия и вывод из некоторой вероятностной модели.

Пусть с помощью критерия согласия Колмогорова проверяется гипотеза о том, что выборка взята из распределения F. Пусть функции распределения F и G удовлетворяют соотношению (4). Пусть на самом деле выборка взята из распределения G, а не F. При каких δ не удастся различить F и G? Для определенности, при каких δ гипотеза согласия с F будет приниматься не менее чем в 50 % случаев?

Критерий согласия Колмогорова основан на статистике:

$$\lambda_n = \sqrt{n}\rho(F_n, H), \qquad (6)$$

где расстояние ρ между функциями распределения определено выше в формуле (4); H — та функция распределения, согласие с которой проверяется, а F_n — эмпирическая функция распределения (т.е. $F_n(x)$ равно доле наблюдений, меньших x, в выборке объема n). Как показал А.Н. Колмогоров в 1933 г., функция распределения случайной величины λ_n при росте объема выборки n сходится к некоторой функции распределения $K(x)$, которую ныне называют функцией Колмогорова. При этом $K(1{,}36)= 0{,}95$ и $K(0{,}83)=0{,}50$.

Поскольку выборка взята из распределения G, то с вероятностью 0,50:

$$\rho(F_n, G) < 0{,}83/\sqrt{n} \ (7)$$

(при больших n). Тогда для рассматриваемой выборки с учетом неравенства (4) и неравенства треугольника для расстояния Колмогорова и симметричности этого расстояния имеем:

$$\rho(F_n, F) \le \rho(F_n, G) + \rho(G\ , F) = \rho(F_n, G) + \rho(F\ , G) < 0{,}83/\sqrt{n} + \delta.$$

Если:

$$0{,}83/\sqrt{n} + \delta \le 1{,}36/\sqrt{n},$$

т.е.

$$\delta\sqrt{n} \le 0{,}53, \qquad (8)$$

то, согласно формуле (6), гипотеза согласия принимается (на уровне значимости 0,95) по крайней мере с той же вероятностью, с которой выполнено неравенство (7), т.е. с вероятностью не менее 0,50. Для $\delta = 0{,}01$ это условие выполняется при $n \le 2809$. Таким образом, для определения функции распределения с точностью $\delta \le 0{,}01$ с помощью критерия согласия Колмогорова необходимо несколько тысяч наблюдений, что для большинства задач прикладной статистики нереально.

При втором из названных выше подходов к определению функции распределения ее конкретный вид выводится из некоторой системы аксиом, в частности, из некоторой модели порождения соответствующей случайной величины. Например, из модели суммирования вытекает нормальное распределение. А из мультипликативной модели (т.е. модели перемножения) — логарифмически нормальное распределение. Как правило, при выводе используется предельный переход. Так, из Центральной Предельной Теоремы теории вероятностей вытекает, что сумма независимых случайных величин может быть приближена нормальным распределением. Однако более детальный анализ, в частности, с помощью неравенства Берри — Эссеена (см. раздел 2.1) показывает, что для гарантированного достижения точности $\delta \leq 0{,}01$ необходимо более полутора тысяч слагаемых. Такого количества слагаемых реально, конечно, указать почти никогда нельзя. Это означает, что при решении практических статистических задач теория дает возможность лишь сформулировать гипотезу о виде функции распределения, а проверять ее надо с помощью анализа реальной выборки объема, как показано выше, не менее нескольких тысяч.

Таким образом, в большинстве реальных ситуаций определить функцию распределения с точностью $\delta \leq 0{,}01$ невозможно.

Итак, показано, что правила отбраковки, основанные на использовании конкретной функции распределения, являются крайне неустойчивыми к отклонениям от нее распределения элементов выборки, а гарантировать отсутствие подобных отклонений почти всегда невозможно. Поэтому отбраковка по классическим правилам математической статистики [6] не является научно обоснованной, особенно при больших объемах выборок. Указанные правила целесообразно применять лишь для выявления «подозрительных» наблюдений, вопрос об отбраковке которых должен решаться из соображений соответствующей предметной области, а не из формально-математических соображений.

Выше для простоты изложения рассмотрен лишь случай полностью известного распределения F, для которого изучено правило отбраковки, заданное формулами (1) и (2). Аналогичные выводы о крайней неустойчивости правил отбраковки справедливы, если «истинное распределение» принадлежит какому-либо параметрическому семейству, например, нормальному, Вейбулла — Гнеденко, гамма.

Параметрическим методам отбраковки, основанным на моделях тех или иных параметрических семейств распределений, посвящены тысячи книг и статей. Приходится признать, что они имеют в основном внутриматематический интерес. При обработке реальных данных следует применять устойчивые методы (см. разделы 3.4 и 14.6). Прежде всего можно рекомендовать непараметрические методы, а среди них — ранговые (т.е. инвариантные в порядковой шкале).

4.3. Предельная теория непараметрических критериев

В прикладной статистике широко используются статистики типа омега-квадрат и типа Колмогорова-Смирнова. Они применяются для проверки согласия с фиксированным распределением или семейством распределений, для проверки однородности двух выборок, симметрии распределения относительно 0, при оценивании условной плотности и регрессии в пространствах произвольной природы и т.д.

Статистики интегрального типа и их асимптотика. Рассмотрим статистики интегрального типа:

$$\xi_\alpha = \xi(f_\alpha, F_\alpha) = \int_X f_\alpha(x,\omega) dF_\alpha(x,\omega), \qquad (1)$$

где X — некоторое пространство, по которому происходит интегрирование (например, $X = [0; 1]$, $X = R^1$ или $X = R^k$). Здесь $\{\alpha\}$ — направленное множество, переход к пределу, где обозначен как $\alpha\to\infty$ (см. главу 14). Случайные функции f_α: $X\times\Omega \to Y$ обычно принимают значения, являющиеся числами. Но иногда рассматривают и постановки, в которых $Y = R^k$ или Y — банахово пространство (т.е. полное нормированное пространство [8]). Наконец, $F_\alpha(x,\omega)$ — случайная функция распределения или случайная вероятностная мера; в последнем случае используют также обозначение $dF_\alpha(x,\omega) = F_\alpha(dx,\omega)$.

Предполагаются выполненными необходимые для корректности внутриматематические предположения измеримости, например, сформулированные в [9, 10].

Пример 1. Рассмотрим критерий Лемана — Розенблатта, т.е. критерий типа омега-квадрат для проверки однородности двух независимых выборок (см. главу 5). Его статистика имеет вид:

$$A = \frac{mn}{m+n}\int_{-\infty}^{\infty}(F_m(x) - G_n(x))^2 dH_{m+n}(x),$$

где $F_m(x)$ — эмпирическая функция распределения, построенная по первой выборке объема m, $G_n(x))$ — эмпирическая функция распределения, построенная по второй выборке объема n, а $H_{m+n}(x)$ — эмпирическая функция распределения, построенная по объединенной выборке объема $m+n$. Легко видеть, что:

$$H_{m+n}(x) = \frac{m}{m+n}F_m(x) + \frac{n}{m+n}G_n(x).$$

Ясно, что статистика A имеет вид (1). При этом x — действительное число, $X = Y = R^1$, в роли α выступает пара (m, n), и $\alpha\to\infty$ означает, что $\min(m, n) \to \infty$. Далее,

$$f_\alpha(x,\omega) = \frac{mn}{m+n}\left(F_m(x) - G_n(x)\right)^2.$$

Наконец, $F_\alpha(x,\omega) = H_{m+n}(x)$.

Теперь обсудим асимптотическое поведение функций $f_\alpha(x,\omega)$ и $F_\alpha(x,\omega)$, с помощью которых определяется статистика A. Ограничимся случаем, когда справедлива гипотеза однородности, функции распределения, соответствующие генеральным совокупностям, из которых взяты выборки, совпадают. Их общую функцию распределения обозначим $F(x)$. Она предполагается непрерывной.

Введем в рассмотрение выборочные процессы:

$$\xi_m(x) = \sqrt{m}(F_m(x) - F(x)),\ \ \eta_n(x) = \sqrt{n}(G_n(x) - F(x)).$$

Нетрудно проверить, что:

$$f_\alpha(x,\omega) = \left(\sqrt{\frac{n}{m+n}}\xi_m(x) - \sqrt{\frac{m}{m+n}}\eta_n(x)\right)^2.$$

Сделаем замену переменной $t = F(x)$. Тогда выборочные процессы переходят в соответствующие эмпирические (см. главу 14):

$$f_\alpha(F^{-1}(t),\omega) = \left(\sqrt{\frac{n}{m+n}}\xi_m(t) - \sqrt{\frac{m}{m+n}}\eta_n(t)\right)^2,\ \ 0 \le t \le 1.$$

Конечномерные распределения этого процесса, т.е. распределения случайных векторов:

$$(f_\alpha(F^{-1}(t_1),\omega), f_\alpha(F^{-1}(t_2),\omega), \dots, f_\alpha(F^{-1}(t_k),\omega))$$

для всех возможных наборов $(t_1, t_2, \dots, t_k)$, сходятся к конечномерным распределениям квадрата броуновского моста $\xi^2(t)$. В соответствии с разделом 14.5 рассматриваемая сходимость по распределению обозначается так:

$$f_\alpha(F^{-1}(t_1),\omega) \Rightarrow \xi^2(t),\ \ 0 \le t \le 1. \tag{2}$$

Нетрудно видеть, что:

$$F_\alpha(x,\omega) = H_{m+n}(x) \to F(x)$$

при $\alpha\to\infty$. С помощью замены переменной $t = F(x)$ получаем, что:

$$F_\alpha(F^{-1}(t),\omega) = H_{m+n}(F^{-1}(t)) \to t \qquad (3)$$

при $\alpha\to\infty$. Из соотношений (2) и (3) хотелось бы сделать вывод, что в случае статистики Лемана — Розенблатта типа омега-квадрат:

$$\xi_\alpha = \int_X f_\alpha(x,\omega)dF_\alpha(x,\omega) = A \Rightarrow \int_0^1 \xi^2(t)dt\,,$$

т.е. предельным распределением этой статистики является классическое распределение [6], найденное как предельное для одновыборочной статистики критерия согласия омега-квадрат Крамера — Мизеса — Смирнова.

Действительно, сформулированное утверждение справедливо. Однако доказательство нетривиально.

Так, может показаться очевидным следующее утверждение.

Утверждение 1. Пусть f: $[0;\ 1] \to R^1$ — ограниченная функция, $G_n(x)$ и $G(x)$ — функции распределения, $G_n(0) = G(0) = 0$, $G_n(1) = G(1) = 1$, причем $G_n(x) \to G(x)$ при всех x. Тогда:

$$\lim_{n\to\infty}\int_0^1 f(x)d(G_n(x)-G(x)) = 0\,. \qquad (4)$$

Это утверждение неверно (ср. [5, с. 42]). Действительно, пусть $f(x) = 1$, если x рационально, и $f(x) = 0$, если x иррационально, $G(x) = x$, а $G_n(x)$ имеет скачки величиной 2^{-n} в точках $m/2^n$, $m = 1, 2, \ldots, 2^n$ при всех $n = 1, 2, \ldots$ Тогда $G_n(x) \to G(x)$ при всех x, однако:

$$\int_0^1 f(x)dG_n(x) = 1,\ \int_0^1 f(x)dG(x) = 0$$

при всех n =1, 2, … Следовательно, вопреки сформулированному выше утверждению 1,

$$\int_0^1 f(x)d(G_n(x)-G(x))=1,$$

т.е. соотношение (4) неверно.

Итак, сформулируем проблему. Пусть известно, что последовательность случайных функций $f_\alpha(x, \omega)$ сходится по распределению при $\alpha\to\infty$ к случайной функции $f(x,\omega)$. Пусть последовательность случайных мер $F_\alpha(A,\omega)$ сходится по распределению к вероятностной мере $F(A)$ при $\alpha\to\infty$. Если речь идет о конечномерном пространстве и меры задаются функциями распределения, то сходимость $F_\alpha(x,\omega)$ к $F(x)$ должна иметь место во всех точках непрерывности $F(x)$. В каких случаях можно утверждать, что при $\alpha\to\infty$ справедлив предельный переход:

$$\xi_\alpha=\xi(f_\alpha,F_\alpha)=\int_X f_\alpha(x,\omega)dF_\alpha(x,\omega)\Rightarrow\xi=\xi(f,F)=\int_X f(x,\omega)dF(x)\ ?$$

Выше показано, что, например, ограниченности $f_\alpha(x, \omega)$ для этого недостаточно.

Метод аппроксимации ступенчатыми функциями. Пусть $T = \{C_1, C_2, \ldots , C_k\}$ — разбиение пространства X на непересекающиеся подмножества. Пусть в каждом элементе C_j разбиения T выделена точка x_j, $j = 1, 2, \ldots , k$. На множестве функций f: $X \to Y$ введем оператор A_T: если $x\in C_j$, то:

$$A_T f(x) = f(x_j), j = 1, 2, \ldots . k. \tag{5}$$

Тогда $A_T f$ — аппроксимация функции f ступенчатыми (кусочно-постоянными) функциями.

Пусть $f_\alpha(x,\omega)$ — последовательность случайных функций на X, а $K(\cdot)$ — функционал на множестве всех возможных их траекторий как функций от x. Для изучения распределения $K(f_\alpha)$ методом аппроксимации ступенчатыми функциями используют разложение:

$$K(f_\alpha) = K(A_T f_\alpha) + \{K(f_\alpha) - K(A_T f_\alpha)\}. \tag{6}$$

Согласно (5) распределение первого слагаемого в (6) определяется конечномерным распределением случайного элемента, а именно, распределением вектора:

$$(f_\alpha(x_1,\omega), f_\alpha(x_2,\omega), \ldots f_\alpha(x_k,\omega)). \tag{7}$$

В обычных постановках предельной теории непараметрических критериев распределение вектора (7) сходится при $\alpha\to\infty$ к соответствующему конечномерному распределению предельной случайной функции $f(x,\omega)$, т.е. к распределению случайного вектора:

$$(f(x_1,\omega), f(x_2,\omega), \dots f(x_k,\omega)). \tag{8}$$

В соответствии с теорией наследования сходимости (глава 1.4) при слабых условиях на функционал $K(\cdot)$ из сходимости по распределению вектора (7) к вектору (8) следует сходимость по распределению $K(A_T f_\alpha)$ к $K(A_T f)$.

Используя аналогичное (6) разложение

$$K(f) = K(A_T f) + \{K(f) — K(A_T f)\}, \tag{9}$$

можно устанавливать сходимость по распределению $K(f_\alpha)$ к $K(f)$ при $\alpha\to\infty$ в два этапа: сначала выбрать разбиение T так, чтобы вторые слагаемые в правых частях соотношений (6) и (9) были малы, а затем при фиксированном операторе A_T воспользоваться сходимостью по распределению $K(A_T f_\alpha)$ к $K(A_T f)$.

Рассмотрим простой пример применения метода аппроксимации ступенчатыми функциями.

Обобщение теоремы Хелли. Пусть f: $[0; 1] \to R^1$ — измеримая функция, $F_n(x)$ — функции распределений, сосредоточенных на отрезке [0; 1]. Пусть $F_n(x)$ сходятся в основном к функции распределения $F(x)$, т.е.:

$$\lim_{n\to\infty} F_n(x) = F(x) \tag{10}$$

для всех x, являющихся точками непрерывности $F(x)$.

Утверждение 2. Если $f(x)$ — непрерывная функция, то:

$$\lim_{n\to\infty} \int_0^1 f(x)dF_n(x) = \int_0^1 f(x)dF(x) \tag{11}$$

(рассматриваются интегралы Лебега — Стилтьеса).

Утверждение 2 известно в литературе как первая теорема Хелли [8, с. 344–346], вторая теорема Хелли [11, с. 174–175], лемма Хелли — Брея [12, с. 193–194].

Естественно поставить вопрос: при каких f из (10) следует (11)? Необходимо ввести условия и на F_n: если $F_n \equiv F$, то соотношение (11) верно для любой измеримой функции f, для которой интеграл в (11) существует. Поэтому рассмотрим следующую постановку.

Постановка 1. Пусть функция f такова, что для *любой* последовательности F_n, удовлетворяющей (10), справедливо (11). Что можно сказать о функции f?

В работах [9, 10] найдены следующие необходимые и достаточные условия на функцию f.

Теорема 1. Пусть ограниченная на [0; 1] функция f интегрируема по Риману-Стилтьесу по функции распределения $F(x)$. Тогда для *любой* последовательности функций распределения F_n, сходящейся в основном к F, имеет место предельный переход (11).

Теорема 2. Пусть функция f не интегрируема по Риману — Стилтьесу по функции распределения $F(x)$. Тогда *существует* последовательность функций распределения F_n, сходящаяся в основном к F, для которой соотношение (11) не выполнено.

Теоремы 1 и 2 в совокупности дают необходимые и достаточные условия для f в постановке 1. А именно, необходимо и достаточно, чтобы ограниченная на [0; 1] функция f была интегрируема по Риману — Стилтьесу по F.

Напомним определение интегрируемости функции f по Риману-Стилтьесу по функции распределения F [8, с. 341]. Рассмотрим разбиение $T = \{C_1, C_2, \ldots , C_k\}$, где:

$$\begin{gathered} C_i = [y_{i-1}, y_i),\ i = 1, 2, \ldots, m-1,\ C_m = [y_{m-1}, y_m], \\ 0 = y_0 < y_1 < y_2 < \ldots < y_m = 1. \end{gathered} \tag{12}$$

Выберем в C_i произвольную точку x_i, $i = 1, 2, \ldots, m$, и составим сумму:

$$S(T) = \sum_{i=1}^{m} f(x_i)[F(y_i) - F(y_{i-1})].$$

Если при $\max(y_i - y_{i-1}) \to 0$ эти суммы стремятся к некоторому пределу (не зависящему ни от способа дробления отрезка [0; 1], ни от выбора точек x_i в каждом из элементов разбиения), то этот предел называется интегралом Римана-Стилтьеса от функции f по функции F по отрезку [0; 1] и обозначается символом, приведенным в правой части равенства (11).

Рассмотрим суммы Дарбу — Стилтьеса:

$$S_H(T)=\sum_{i=1}^{m} m_i[F(y_i)-F(y_{i-1})],\quad S_B(T)=\sum_{i=1}^{m} M_i[F(y_i)-F(y_{i-1})],$$

где

$$m_i=\inf\{f(x),x\in X_i\},\quad M_i=\sup f\{f(x),x\in X_i\}.$$

Ясно, что

$$S_H(T)\le S(T)\le S_B(T).$$

Необходимым и достаточным условием интегрируемости по Риману — Стилтьесу является следующее: для любой последовательности разбиений T_k, $k = 1, 2, 3, \ldots$ вида (12) такой, что $\max(y_i - y_{i-1}) \to 0$ при $k\to\infty$, имеем:

$$\lim_{k\to\infty}[S_B(T_k)-S_H(T_k)]=0. \tag{13}$$

Напомним, что согласно подразделу 1.4.3 колебанием $\delta(f, B)$ функции f на множестве B называется $\delta(f, B) = \sup\{|f(x) - f(y)|, x\in B, y\in B\}$. Поскольку

$$\delta(f, C_i) = M_i - m_i,$$

то условие (13) можно записать в виде:

$$\lim_{k\to\infty}\sum_{C\in T_k}\delta(f,C)F(C)=0. \tag{14}$$

Условие (14), допускающее обобщение с $X = [0; 1]$ и f: $[0; 1] \to R^1$ на X и f более общего вида, и будем использовать при доказательстве теорем 1 и 2.

Доказательство теоремы 1. Согласно методу аппроксимации ступенчатыми функциями рассмотрим оператор A_T. Как легко проверить, имеет место разложение:

$$\begin{gathered}\beta_n=\int_0^1 f(x)dF_n(x)-\int_0^1 f(x)dF(x)=\int_0^1\{f(x)-A_T f(x)\}dF_n(x)+\\ +\int_0^1\{A_T f(x)-f(x)\}dF(x)+\left\{\int_0^1 A_T f(x)dF_n(x)-\int_0^1 A_T f(x)dF(x)\right\}\end{gathered}. \tag{15}$$

Поскольку:

$$|f(x) - A_T f(x)| \le \delta(f, X_i),\ x \in C_i,$$

то первое слагаемое в правой части (15) не превосходит:

$$\sum_{C \in T} \delta(f, C) F_n(C), \qquad (16)$$

а второе не превосходит:

$$\sum_{C \in T} \delta(f, C) F(C).$$

Согласно определению оператора A_T , третье слагаемое в (15) имеет вид:

$$\sum_{i=1}^{m} f(x_i)(F_n(C_i) - F(C_i)).$$

Очевидно, оно не превосходит по модулю:

$$\sup_{x \in X} |f(x)| \sum_{C \in T} |F_n(C) - F(C)|$$

(здесь используется ограниченность f на X).

Согласно (16) первое слагаемое в правой части (15) не превосходит:

$$\sum_{C \in T} \delta(f, C) F(C) + \sum_{C \in T} \delta(f, C) |F_n(C) - F(C)|.$$

Поскольку

$$\delta(f, C) \le 2 \sup_{x \in X} |f(x)|,$$

то первое слагаемое в правой части (15) не превосходит:

$$\sum_{C \in T} \delta(f, C) F(C) + 2 \sup_{x \in X} |f(x)| \sum_{C \in T} |F_n(C) - F(C)|.$$

Из оценок, относящихся к трем слагаемым в разложении (15), следует, что:

$$|\beta_n| \le 2\sum_{C\in T}\delta(f,C)F(C) + 3\sup_{x\in X}|f(x)|\sum_{C\in T}|F_n(C)-F(C)|. \qquad (17)$$

Используя оценку (17), докажем, что $\beta_n \to 0$ при $n \to \infty$. Пусть дано $\varepsilon > 0$. Согласно условию интегрируемости функции f по Риману-Стилтьесу, т.е. условию (14), можно указать разбиение $T = T(\varepsilon)$ такое, что:

$$\sum_{C\in T(\varepsilon)}\delta(f,C)F(C) < \frac{\varepsilon}{4}, \qquad (18)$$

и в точках y_i, $i = 1, 2, \ldots, m - 1$ (см. (12)), функция F непрерывна.

Поскольку

$$F_n(X_i) = F_n(y_i) - F_n(y_{i-1}),$$

то из (10) следует, что существует число $n = n(\varepsilon)$ такое, что при $n > n(\varepsilon)$ справедливо неравенство:

$$\sum_{C\in T(\varepsilon)}|F_n(C)-F(C)| < \frac{\varepsilon}{6}\left(\sup_{x\in X}|f(x)|\right)^{-1}. \qquad (19)$$

Из (17), (18) и (19) следует, что при $n > n(\varepsilon)$ справедливо неравенство:

$$\left|\int_0^1 f(x)dF_n(x) - \int_0^1 f(x)dF(x)\right| < \varepsilon,$$

что и требовалось доказать.

Обсудим условие ограниченности f. Если оно не выполнено, то из (10) не всегда следует (11).

Пример 2. Пусть $f(x) = 1/x$ при $x > 0$ и $f(0)=0$. Пусть $F(0,5) = 0$, т.е. предельное распределение сосредоточено на [1/2; 1]. Пусть распределение F_n на [0; S) имеет единственный атом в точке $x = 1/n$ величиной $n^{-1/2}$, а на [1/2; 1] справедливо (10). Тогда по причинам, изложенным при доказательстве теоремы 1:

$$\lim_{n\to\infty}\int_{1/2}^1 f(x)dF_n(x) = \int_{1/2}^1 f(x)dF(x),$$

однако:

$$\int_0^{1/2} f(x)dF_n(x) = \sqrt{n}, \quad \int_0^{1/2} f(x)dF(x) = 0,$$

т.е. соотношение (11) не выполнено.

Условие ограниченности подынтегральной функции f можно заменить, как это сделано, например, в [9], на условие строгого возрастания функции распределения F.

Лемма. Пусть функции распределения F всюду строго возрастает, т.е. из $x_1 < x_2$ вытекает $F(x_1) < F(x_2)$. Пусть функция f интегрируема по Риману — Стилтьесу по F, т.е. выполнено (14). Тогда функция f ограничена.

Доказательство. Рассмотрим точки $0 = y_0 < y_1 < y_2 <\ldots< y_{2m} = 1$ и два разбиения:

$$T_1 = \{[0;y_1),[y_1;y_3),[y_3;y_5),\ldots,[y_{2m-1};1]\},\ T_2 = \{[0;y_2),[y_2;y_4),[y_4;y_6),\ldots,[y_{2m-2};1]\}.$$

Тогда для любых двух точек x и x' можно указать конечную последовательность точек $x_1 = x, x_2, x_3, \ldots, x_s, x_{s+1} = x'$ такую, что любые две соседние точки x_i, x_{i+1}, $i = 1, 2, \ldots, s$, одновременно принадлежат некоторому элементу C_i разбиения T_1 или разбиения T_2, причем $C_i \neq C_j$ при $i \neq j$. Действительно, пусть $x \in [y_p; y_{p+1}), x' \in [y_q; y_{q+1})$. Пусть для определенности $q > p$. Тогда можно положить $x_2 = y_{p+1}, x_3 = y_{p+2}, \ldots, x_s = y_q$. Поскольку среди элементов разбиений T_1 и T_2 есть $C_1 = [y_p; y_{p+2})$, то $x \in x_1 \in C_1, x_2 = y_{p+1} \in C_1$. Далее, $x_2 \in [y_{p+1}, y_{p+3}) = C_2, x_3 \in C_2$, и т.д.

Из указанных выше свойств последовательности $x_1 = x, x_2, x_3, \ldots, x_s, x_{s+1} = x'$ следует, что:

$$| f(x) - f(x') | \le \sum_{i=1}^{s} | f(x_{i+1}) - f(x_i) | \le \sum_{C \in T_1} \delta(f,C) + \sum_{C \in T_2} \delta(f,C).$$

Пусть теперь число $\max(y_i - y_{i-2})$ настолько мало, что согласно (14):

$$\sum_{C \in T_1} \delta(f,C)F(C) < 1,\ \sum_{C \in T_2} \delta(f,C)F(C) < 1.$$

Тогда согласно двум последним соотношениям:

$$| f(x) - f(x') | \le 2[\min\{F(C) : C \in T_1 \bigcup T_2\}]^{-1},$$

что и доказывает лемму.

Доказательство теоремы 2. Пусть условие (14) не выполнено, т.е. существуют число $\gamma > 0$ и последовательность разбиений T_n, $n = 1, 2, \ldots$, такие, что $\max(y_i - y_{i-1}) \to 0$ при $n \to \infty$ и при всех n:

$$\sum_{C \in T_n} \delta(f, C) F(C) \geq \gamma . \tag{20}$$

Для доказательства теоремы построим две последовательности функций распределения F_{1n} и F_{2n}, $n = 1, 2, \ldots$, для которых выполнено (10), но последовательность:

$$\delta_n = \int_0^1 f(x) dF_{1n}(x) - \int_0^1 f(x) dF_{2n}(x)$$

не стремится к 0 при $n \to \infty$. Тогда (11) не выполнено хотя бы для одной из последовательностей F_{1n} и F_{2n}.

Для любого C — элемента некоторого разбиения T — можно указать, как вытекает из определения $\delta(f, C)$, точки $x_1(C)$ и $x_2(C)$ такие, что:

$$f(x_1(C)) - f(x_2(C)) > S\,\delta(f, C). \tag{21}$$

Построим F_{1n} и F_{2n} следующим образом. Пусть $F_{1n}(C) = F_{2n}(C) = F(C)$ для любого C из T_n. При этом F_{1n} имеет в C один атом в точке $x_1(C)$ величиной $F(C)$, а F_{2n} имеет в C также один атом в точке $x_2(C)$ той же величины $F(C)$. Другими словами, распределение F_{1n} в C сосредоточено в одной точке, а именно, в $x_1(C)$, а распределение F_{2n} сосредоточено в $x_2(C)$. Тогда:

$$\delta_n = \sum_{C \in T_n} (f(x_1(C)) - f(x_2(C))) F(C) . \tag{22}$$

Из (20), (21) и (22) следует, что:

$$\delta_n \geq \frac{1}{2} \sum_{C \in T_n} \delta(f, C) F(C) \geq \frac{\gamma}{2} .$$

Остается показать, что для последовательностей функций распределения F_{1n} и F_{2n} выполнено (10). Пусть x — точка непрерывности F. Пусть

$$y_1(x, T) = \max\{y_{kn}: y_{kn} < x\},\ y_2(x, T) = \min\{ y_{kn}: y_{kn} > x\},$$

где y_{kn} — точки, определяющие разбиения T_n согласно (12). В соответствии с определением F_{in}:

$$F_{in}(y_j(x, T_n)) = F(y_j(x, T_n)),\ i = 1, 2, j = 1, 2,$$

а потому:

$$|F_{in}(\mathrm{x}) — F(x)| \le F(y_2(x, T_n)) - F(y_1(x, T_n)),\ i = 1, 2.$$

В силу условия $\max(y_{kn} - y_{(k-1)n}) \to 0$ и непрерывности F в точке x правая часть последнего соотношения стремится к 0 при $n \to \infty$, что и заканчивает доказательство теоремы 2.

Теоремы 1 и 2 демонстрируют основные идеи предельной теории статистик интегрального типа и непараметрических критериев в целом. Как показывают эти теоремы, основную роль в рассматриваемой теории играет предельное соотношение (14). Отметим, что если $\delta(f, T_n) \to 0$ при $n \to \infty$, то (14) справедливо, но, вообще говоря, не наоборот. Возникает еще ряд постановок. Пусть (14) выполнено для f_1 и f_2. При каких функциях h это соотношение выполнено для $h(x, f_1(x), f_2(x))$? В прикладной статистике вместо $f(x)$ рассматривают $f_\alpha(x, \omega)$ и $f(x, \omega)$, а вместо интегрирования по функциям распределения $F_n(x)$ — интегрирование по случайным мерам $F_\alpha(\omega)$. Как меняются формулировки в связи с такой заменой? В связи со слабой сходимостью (т.е. сходимостью по распределению) $A_T f_\alpha$ к A_T и переходом от $f_\alpha(x, \omega)$ к $h_\alpha(x, f_{1\alpha}(x, \omega), f_{2\alpha}(x, \omega))$ возникает следующая постановка. Пусть последовательность случайных элементов κ_α слабо сходится к случайному элементу κ при $\alpha \to \infty$. Когда распределения случайных элементов $g_\alpha(\kappa_\alpha)$ сближаются с распределениями случайных элементов $g_\alpha(\kappa)$? Полным ответом на последний вопрос являются необходимые и достаточные условия наследования сходимости. Они приведены в главе 14.

Основные результаты. Наиболее общая теорема типа теоремы 1 выглядит так [10].

Теорема 3. Пусть существует последовательность разбиений T_n, $n = 1, 2, \ldots$, такая, что при $n \to \infty$ и $\alpha \to \infty$:

$$\Delta(f_\alpha, T_n) = \sum_{C \in T_n} \delta(f_\alpha, C) F(C) \to 0. \tag{23}$$

Пусть для любого C, входящего хотя бы в одно из разбиений T_n,

$$F_\alpha(C, \omega) \to F(C) \tag{24}$$

при $\alpha\to\infty$ (сходимость по вероятности). Пусть f_α асимптотически ограничены по вероятности при $\alpha\to\infty$. Тогда:

$$\xi(f_\alpha,F_\alpha)-\xi(f_\alpha,F)\to 0 \qquad (25)$$

при $\alpha\to\infty$ (сходимость по вероятности).

Как известно, полное сепарабельное метрическое пространство называется польским. Это понятие понадобится для формулировки аналога теоремы 2.

Теорема 4. Пусть X — польское пространство, У конечномерно, существует измельчающаяся последовательность T_n разбиений, для которой соотношение (23) не выполнено. Тогда существует удовлетворяющая (24) последовательность F_α, для которой соотношение (25) неверно, хотя F_α слабо сходится к F при $\alpha\to\infty$.

Условие (23) естественно назвать условием римановости, поскольку в случае, рассмотренном в теореме 1, оно является условием интегрируемости по Риману-Стилтьесу. Рассмотрим *наследуемость римановости* при переходе от $f_{1\alpha}(x, \omega)$ со значениями в $У_1$ и $f_{2\alpha}(x, \omega)$ со значениями в $У_2$, удовлетворяющих (23), к $h_\alpha(x, f_{1\alpha}(x, \omega), f_{2\alpha}(x, \omega))$ со значениями в $У_3$.

Положим:

$$Y_k(a,\varepsilon)=\{(y,y'):y\in Y_k, y'\in Y_k, \|y\|_k<a, \|y'\|<a, \|y-y'\|<\varepsilon\}, k=1,2,$$

где $\|\cdot\|_k$ — норма (т.е. длина вектора) в пространстве Y_k, k = 1, 2. Рассмотрим также множества:

$$A(C,a,\varepsilon)=\{(x,x',y_1,y_1^*,y_2,y_2^*):x,x'\in C,(y_k,y_k^*)\in Y_k(a,\varepsilon),k=1,2\}$$

и функции:

$$q_\alpha(x,x',y_1,y_1^*,y_2,y_2^*)=h_\alpha(x,y_1,y_2)-h_\alpha(x',y_1^*,y_2^*).$$

Наконец, понадобится измеритель колеблемости:

$$c(h_\alpha,T,a,\varepsilon)=\sum_{C\in T}\sup_{A(C,a,\varepsilon)}\|q_\alpha\|_3\,F(C)$$

и множество:

$$Z(a)=X\times\{y_1:\|y_1\|<a\}\times\{y_2:\|y_2\|<a\}.$$

Теорема 5. Пусть h_α асимптотически (при $\alpha\to\infty$) ограничены на $Z(a)$ при любом положительном a, функции $f_{1\alpha}$ и $f_{2\alpha}$ асимптотически ограничены по вероятности и удовлетворяют условию (23). Пусть для участвующей в (23) последовательности T_n:

$$c(h_\alpha, T_n, a, \varepsilon) \to 0 \tag{26}$$

при $\alpha\to\infty$, $n\to\infty$, $\varepsilon\to 0$ и любом положительном a. Тогда $f_{3\alpha}(x, \omega) = h_\alpha(x, f_{1\alpha}(x, \omega), f_{2\alpha}(x, \omega))$ удовлетворяют условию (23) и асимптотически ограничены по вероятности.

Теорема 6. Пусть условие (26) не выполнено для h_α. Тогда существуют детерминированные ограниченные функции $f_{1\alpha}$ и $f_{2\alpha}$ такие, что соотношение (23) выполнено для $f_{1\alpha}$ и $f_{2\alpha}$ и не выполнено для $f_{3\alpha}$.

Пример 3. Пусть $X = [0; 1]^k$, пространства Y_1 и Y_2 конечномерны, функция $h_\alpha \equiv h(x, y_1, y_2)$ непрерывна. Тогда условие (26) выполнено.

С помощью теорем 3 и 5 и результатов о наследовании сходимости можно изучить асимптотическое поведение статистик интегрального типа

$$\xi_\alpha = \int_X h_\alpha(x, f_{1\alpha}(x,\omega), f_{2\alpha}(x,\omega)) F_\alpha(dx,\omega)$$

со значениями в банаховом пространстве Y.

Теорема 7. Пусть для некоторой последовательности T_n разбиений X справедливы соотношения (23) для $f_{1\alpha}$ и $f_{2\alpha}$ и (24) для F_α. Пусть последовательность функций h_α удовлетворяет условию в теореме 5, конечномерные распределения $(f_{1\alpha}(x, \omega), f_{2\alpha}(x, \omega))$ слабо сходятся к конечномерным распределениям $(f_1(x, \omega), f_2(x, \omega))$, причем для f_1 и f_2 справедливо соотношение (23). Тогда:

$$\lim_{\alpha\to\infty} L(\xi_\alpha, \eta_\alpha) = 0,$$

где L — расстояние Прохорова (см. раздел 14.3),

$$\eta_\alpha = \int_X h_\alpha(x, f_1(x,\omega), f_2(x,\omega)) F(dx).$$

Теорема 7 дает общий метод получения асимптотических распределений статистик интегрального типа. Важно, что соотношение (23) выполнено для эмпирического процесса и для процессов, связанных с оцениванием параметров при проверке согласия [9].

Один из выводов общей теории состоит в том, что в качестве F_α можно использовать практически любую состоятельную оценку истинной функции распределения. Этот вывод использовался при построении критерия типа омега-квадрат для проверки симметрии распределения относительно 0 и обнаружения различий в связанных выборках (см. главу 5).

Асимптотическое поведение критериев типа Колмогорова может быть получено с помощью описанного выше метода аппроксимации ступенчатыми функциями. Этот метод не требует обращения к теории сходимости вероятностных мер в функциональных пространствах. Для критериев Колмогорова и Смирнова достаточно использовать лишь свойства эмпирического процесса и броуновского моста. В случае проверки согласия добавляется необходимость изучения еще одного случайного процесса. Он является разностью между двумя функциями распределения. Одна — функция распределения элементов выборки. Вторая — член параметрического семейства распределений, полученный путем подстановки оценок параметров вместо их истинных значений.

4.4. Метод проверки гипотез по совокупности малых выборок

Одна из областей применения прикладной статистики — статистические методы управления качеством продукции (глава 10 и [13, гл. 13]). К ним относится статистический приемочный контроль, в котором по результатам испытаний элементов выборки делается вывод о качестве партии продукции. В простейшем варианте проводится контроль по альтернативному признаку, при котором возможны лишь два результата контроля конкретной единицы продукции — «соответствует требованиям» или «не соответствует требованиям», короче — «да» или «нет».

Рассмотрим статистический приемочный контроль по двум альтернативным признакам одновременно. На основе теории люсианов обсудим проблему проверки независимости двух альтернативных признаков. Ее приходится проводить по совокупности малых выборок, т.е. в так называемой асимптотике А.Н. Колмогорова, когда число неизвестных параметров распределения не является постоянным, а растет пропорционально объему данных.

Испытания по двум альтернативным признакам. При статистическом контроле качества продукции, в частности, при сертификации, чаще всего используют контроль по альтернативным признакам. При этом устанавливается, соответствует ли контролируемый параметр единицы продукции (изделия, детали) заданным в нормативно-технической документации требованиям или не соответствует. Если соответствует — единица продукции признается годной. Примем для определенности, что в этом случае результат контроля кодируется символом 0. Если же не соответствует — единица продукции признается дефектной, а результат контроля кодируется символом 1.

Таким образом, в рассматриваемой нами математической модели контроля альтернативный признак — это функция $X = X(w)$, определенная на множестве единиц продукции $W = \{w\}$ и принимающая два значения 0 и 1. Причем $X(w) = 0$ означает, что единица продукции w является годной, а $X(w) = 1$ — что она является дефектной.

Методы статистического контроля, в частности, включенные в государственные стандарты и иную нормативно-техническую документацию (НТД), как правило, используют контроль по одному признаку. В НТД указывают правила выбора планов контроля и расчета различных их характеристик, приводят графики оперативных характеристик и т.п.

Однако на производстве контроль нередко проводится по нескольким альтернативным признакам. Возникает проблема выбора плана контроля и расчета его характеристик.

Рассмотрим сначала контроль по двум альтернативным признакам $X(w)$ и $Y(w)$. В вероятностной модели $X(w)$ и $Y(w)$ — случайные величины, принимающие два значения — 0 и 1. Пусть, пользуясь стандартной (для статистических методов управления качеством) терминологией,

$$p_1 = P(X(w) = 1)$$

— входной уровень дефектности для первого признака, а

$$p_2 = P(Y(w) = 1)$$

— для второго. Вероятности результатов контроля по двум признакам одновременно описываются четырьмя числами:

$$P(X(w) = 0,\ Y(w) = 0) = p_{00},\ P(X(w) = 1,\ Y(w) = 0) = p_{10},$$
$$P(X(w) = 0,\ Y(w) = 1) = p_{01},\ P(X(w) = 1,\ Y(w) = 1) = p_{11}.$$

При этом справедливы соотношения:

$$p_{00} + p_{10} + p_{01} + p_{11} = 1,\ p_{10} + p_{11} = p_1,\ p_{01} + p_{11} = p_2.$$

С прикладной точки зрения наиболее интересна вероятность p_{00} того, что единица продукции является годной (по всем параметрам), и вероятность ее дефектности $(1 - p_{00})$, т.е. входной уровень дефектности для изделия в целом.

В табл. 1 сведены вместе введенные выше вероятности.

Таблица 1

Вероятности результаты испытаний при контроле по двум альтернативным признакам

Значения признаков	X**=0**	X**=1**	**Всего**
$Y=0$	p_{00}	p_{10}	$1 - p_2$
$Y=1$	p_{01}	p_{11}	p_2
Всего	$1 - p_1$	p_1	1

Есть три важных частных случая — поглощения, несовместности и независимости дефектов. Другими словами, поглощения, несовместности и независимости событий $\{w: X(w) = 1\}$ и $\{w: Y(w) = 1\}$. В случае поглощения одно из этих событий содержит другое, а потому:

$$p_{00} = 1 - \max(p_1 , p_2).$$

В случае несовместности:

$$p_{00} = 1 - p_1 - p_2.$$

В случае независимости:

$$p_{00} = (1 - p_1)(1 - p_2) = 1 - p_1 - p_2 + p_1 p_2.$$

Очевидно, что вероятность годности изделия всегда заключена между значениями, соответствующими случаям поглощения и несовместности. Кроме того, известно, что при большом числе признаков и малой вероятности дефектности по каждому из них случаи поглощения и независимости дают (в асимптотике) крайние значения для вероятности годности изделия, т.е. формулы, соответствующие независимости и несовместности, асимптотически совпадают.

Причина этого явления состоит в том, что при малости p_1 и p_2 их произведение p_1p_2является бесконечно малой более высокого порядка по сравнения с p_1 и p_2.

Рассмотрим несколько примеров. Пусть некоторая продукция, скажем, гвозди, контролируются по двум альтернативным признакам, для определенности, по весу и длине. Пусть результаты контроля 1 000 единиц продукции представлены в табл. 2.

Таблица 2

Результаты 1 000 испытаний по двум альтернативным признакам (случай поглощения)

Значения признаков	***X*=0**	***X*=1**	**Всего**
У=0	952	0	952
У=1	0	48	48
Всего	952	48	1 000

Судя по данным табл. 2, дефекты всегда встречаются парами — если есть один, то есть и другой. Входной уровень дефектности как по каждому показателю, так и по обоим вместе — один и тот же, а именно, 0,048. Получив по результатам статистического наблюдения данные типа приведенных в табл. 2, целесообразно перейти к контролю только одного показателя, а не двух. Каково именно? Видимо, того, контроль которого дешевле. Однако совсем иная ситуация в случае несовместности дефектов (табл. 3).

Таблица 3

Результаты 1 000 испытаний по двум альтернативным признакам (случай несовместности)

Значения признаков	***X*=0**	***X*=1**	**Всего**
У=0	904	48	952
У=1	48	0	48
Всего	952	48	1 000

Судя по данным табл. 3, дефекты всегда встречаются поодиночке — если есть один, то другого нет. В результате входной уровень дефектности по каждому признаку по-прежнему равен 0,048, в то время как доля дефектных изделий (т.е. имеющих хотя бы один дефект) вдвое выше, т.е. входной уровень дефектности для изделия в целом равен 0,096.

Случай независимости результатов контроля по двум независимым признакам (табл. 4) лежит между крайними случаями поглощения и несовместности. Независимость альтернативных признаков обосновывается путем статистической проверки с помощью описанного ниже критерия $n^{1/2}V$.

Таблица 4

Результаты 1 000 испытаний по двум альтернативным признакам (случай независимости)

Значения признаков	X=0	X=1	**Всего**
$У$=0	909	43	952
$У$=1	43	5	48
Всего	952	48	1 000

Согласно данным табл. 4, входной уровень дефектности для каждого из двух альтернативных признаков по-прежнему равен 0,048, в то время как для изделий в целом он равен 0,091, т.е. на 5,2 % меньше, чем в случае несовместности, и на 89,6 % больше, чем в случае поглощения.

Проблема состоит в том, что таблицы и стандарты по статистическому приемочному контролю относятся обычно к случаю одного контролируемого параметра. А как быть, если контролируемых параметров несколько? Приведенные выше примеры показывают, что входной уровень дефектности изделия в целом не определяется однозначно по входным уровням дефектности отдельных его параметров.

Гипотеза независимости. Как должны соотноситься характеристики планов контроля по отдельным признакам с характеристиками плана контроля по двум (или многим) признакам одновременно? Рассмотрим распространенную рекомендацию — складывать уровни дефектности, т.е. считать, что уровень дефектности изделия в целом равен сумме уровней дефектности по отдельным его параметрам. Она, очевидно, опирается на гипотезу несовместности дефектов, а потому во многих случаях преувеличивает дефектность, следовательно, ведет к использованию излишне жестких планов контроля, что экономически невыгодно.

Зная специфику применяемых технологических процессов, в ряде конкретных случаев можно предположить, что дефекты по различным признакам возникают независимо друг от друга. Это предположение необходимо обосновывать по статистическим данным. Если же оно обосновано, следует рассчитывать входной уровень дефектности по формуле:

$$1 - p_{00} = p_1 + p_2 - p_1 p_2,$$

соответствующей независимости признаков.

Итак, необходимо уметь проверять по статистическим данным гипотезу независимости двух альтернативных признаков. Речь идет о статистической проверке нулевой гипотезы:

$$H_0\colon p_{11} = p_1 p_2 \qquad (1)$$

(что эквивалентно проверке равенства $p_{00} = (1 - p_1)(1 - p_2)$). Нетрудно проверить, что гипотеза о справедливости равенства (1) эквивалентна гипотезе

$$H_0 : p_{00}\, p_{11} - p_{10}\, p_{01} = 0. \qquad (2)$$

В простейшем случае предполагается, что проведено n независимых испытаний $(Xi,\ Y_i)$, $i = 1, 2, ..., n$, в каждом из которых проконтролированы два альтернативных признака, а вероятности результатов контроля не меняются от испытания к испытанию. Общий вид статистических данных приведен в табл. 5.

Таблица 5

Общий вид результатов контроля по двум альтернативным признакам

Значения признаков	***X*=0**	***X*=1**	**Всего**
$У = 0$	a	b	$a + b$
$У = 1$	c	d	$c + d$
Всего	$a + c$	$b + d$	n

В табл. 5 величина a — число испытаний, в которых $(Xi,\ Y_i) = (0{,}0)$, величина b — число испытаний, в которых $(Xi,\ Y_i) = (1{,}0)$, и т.д.

Случайный вектор $(a,\ b,\ c,\ d)$ имеет мультиномиальное распределение с числом испытаний n и вектором вероятностей исходов $(p_{00}, p_{10}, p_{01}, p_{11})$. Состоятельными оценками этих вероятностей являются дроби a/n, b/n, c/n, d/n соответственно. Следовательно, критерий проверки гипотезы (2) может быть основан на статистике:

$$Z = ad - bc. \tag{3}$$

Как вытекает из известной формулы для ковариаций мультиномиального вектора (см., например, формулу (6.3.5) в учебнике С. Уилкса [14] на с. 153),

$$M(Z) = n\,(p_{10}\,p_{01} - p_{00}\,p_{11}), \tag{4}$$

что равно 0 при справедливости гипотезы независимости (2).

Связь между переменными X и Y обычно измеряется коэффициентом, отличающимся от Z нормирующим множителем:

$$V = (ad - bc)\{(a + b)(a + c)(b + d)(c + d)\}^{-1/2}$$

(см. классическую монографию М. Дж. Кендалла и А. Стьюарта [15, с. 723]). При справедливости гипотезы H_0 и больших n случайная величина nV^2 имеет хи-квадрат распределение с одной степенью свободы, а $n^{1/2}V$ имеет стандартное нормальное распределение с математическим ожиданием 0 и дисперсией 1 (см. [15, с. 736]). Значение $n^{1/2}V$ для данных табл. 4 равно 1,866, т.е. на уровне значимости 0,05 гипотезу независимости следует принять.

Рассмотрим еще один пример. Пусть проведено 100 испытаний, результаты которых описаны в табл. 6. Тогда:

$$V = (50 \cdot 20 - 10 \cdot 20)\,(60 \cdot 70 \cdot 30 \cdot 40)^{-1/2} = (1000 - 200) \cdot 5940000^{-1/2} =$$
$$= 800 / 2245 = 0{,}35635,\ n^{1/2}V = 3{,}5635.$$

Таблица 6

Результаты 100 испытаний по двум альтернативным признакам

Значения признаков	***X*=0**	***X*=1**	**Всего**
У=0	50	10	60
У=1	20	20	40
Всего	70	30	100

Поскольку полученное значение $n^{1/2}V$ превышает критическое значение при любом применяемом в статистике уровне значимости, то гипотезу о независимости признаков необходимо отклонить.

Проверка гипотез по совокупности малых выборок. К сожалению, приведенный простой метод годится не всегда. При статистическом анализе реальных данных возникают проблемы, связанные с отсутствием достаточно больших однородных выборок, т.е. выборок, в которых постоянны параметры вероятностных распределений. Реально единицы продукции представляются на контроль партиями, из каждой партии контролируются лишь несколько изделий, т.е. малая выборка. При этом от партии к партии меняются параметры p_{00}, p_{10}, p_{01}, p_{11}, описывающие уровень дефектности. Поэтому необходимы статистические методы, позволяющие проверять гипотезу независимости признаков по совокупности малых выборок. Построим один из возможных методов.

Рассмотрим вероятностную модель совокупности k малых выборок объемов n_1 , n_2 ,..., n_k соответственно. Пусть j-я выборка $(X_{jt}$, $Y_{jt})$, $t = 1, 2, ..., n_j$ имеет распределение, задаваемое вектором параметров $(p_{00j}$, p_{10j}, p_{01j}, $p_{11j})$ в соответствии с ранее введенными обозначениями, $j = 1,2,...,k$. Будем проверять гипотезу:

$$H_0\colon p_{11j} = (p_{10j} + p_{11j})\,(p_{01j} + p_{11j}),\ j = 1, 2, ..., k, \tag{5}$$

или в эквивалентной формулировке:

$$H_0\colon p_{11j}\, p_{00j} - p_{10j}\, p_{01j}\ ,\ j = 1, 2, ..., k\,. \tag{6}$$

Основная идея состоит в нахождении асимптотического распределения статистики типа $n^{1/2}V$ при росте числа k малых выборок, а именно, статистики:

$$S = g_1 Z_1 + g_2 Z_2 + ... + g_k Z_k\ , \tag{7}$$

где Z_1, Z_2, ..., Z_k — статистики, рассчитанные по формуле (3) для каждой из k выборок, т.е. $Z_j = a_j d_j - b_j c_j$, $j = 1,2, ..., k$, а g_1 , g_2 , ..., g_k — некоторые весовые коэффициенты, которые, в частности, могут совпадать. Поскольку:

$$M(S) = g_1 M(Z_1) + g_2 M(Z_2) + ... + g_k M(Z_k),$$

то при справедливости гипотезы независимости (5)–(6) имеем $M(S) = 0$ согласно соотношению (4). Поскольку слагаемые в сумме (7) независимы, то при росте k случайная величина S в силу Центральной Предельной Теоремы является асимптотически нормальной. Дисперсия этой величины равна сумме дисперсий слагаемых:

$$D(S) = g_1^2 D(Z_1) + g_2^2 D(Z_2) + ... + g_k^2 D(Z_k). \tag{8}$$

Для оценивания дисперсии S необходимо использовать **несмещенные** оценки дисперсий в каждой из k выборок (и в этом одна из основных «изюминок» разбираемого метода). Предположим, что построены статистики Tj такие, что:

$$M(Tj) = D(Zj)\,,\ j = 1, 2, ..., k. \tag{9}$$

Тогда при некоторых математических «условиях регулярности», на которых нет необходимости здесь останавливаться, несмещенная оценка дисперсии статистики S, имеющая согласно формулам (8) и (9) вид:

$$L = g_1^2 T_1 + g_2^2 T_2 + ... + g_k^2 T_k\,,$$

в силу закона больших чисел такова, что дробь $D(S)/L$ приближается к 1 при росте числа выборок (сходимость по вероятности). Отсюда следует, что распределение случайной величины $Q = SL^{-1|2}$ приближается при росте числа выборок к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Следовательно, критерий проверки гипотезы (5)–(6) независимости признаков, состоящий в том, что при $(-1{,}96) < Q < 1{,}96$ гипотеза принимается, а при Q , выходящих за пределы интервала (– 1,96; 1,96), гипотеза отклоняется, имеет уровень значимости, приближающийся к 0,05 при росте числа выборок. Мощность этого критерия зависит от величины $M(S)D(S)^{-1|2}$ при альтернативе.

Для реализации намеченного плана осталось научиться несмещенно оценивать $D(Zj)$. К сожалению, в литературе по несмещенному оцениванию не рассматривают случай мультиномиального распределения, поэтому кратко опишем процедуру построения несмещенной оценки $D(Zj)$. Поскольку согласно формулам (3) и (4):

$$\begin{aligned} D(Z_j) &= M(Z_j^2) - (M(Z_j))^2 = M(a_j^2 d_j^2) - 2M(a_j b_j c_j d_j) + \\ &+ M(b_j^2 c_j^2) + n_j^2 (p_{00j} p_{11j} - p_{01j} p_{10j})^2, \end{aligned} \tag{10}$$

то для вычисления $D(Z_j)$ достаточно найти входящие в правую часть формулы (10) начальные смешанные моменты мультиномиального распределения (четвертого порядка). Теоретически это просто — известен вид характеристической функции мультиномиального распределения (см., например, формулу (6.3.4) в монографии [14, с. 152]), а начальные смешанные моменты равны значениям ее соответствующих производных в 0, деленным на нужную степень мнимой единицы (формула (5.2.3) в монографии [14, с. 131]). Например, с помощью описанной процедуры после некоторых вычислений получаем, что (для упрощения записи здесь и далее опустим индекс j):

$$M(a^2d^2) = n(n-1)(n-2)(n-3)p_{11}^2p_{00}^2 + \\ + n(n-1)(n-2)(p_{11}^2p_{00} + p_{11}p_{00}^2) + n(n-1)p_{11}p_{00}. \tag{11}$$

Формула (11) показывает, что начальные смешанные моменты мультиномиального распределения являются многочленами от параметров p_{11}, p_{00}, p_{10}, p_{01} этого распределения, однако конкретный вид этих многочленов достаточно громоздок, поэтому не будем их здесь выписывать, ограничившись формулой (11) в качестве образца.

Как вытекает из формул (10) и (11), для построения несмещенной оценки $D(Z_j)$ достаточно научиться несмещенно оценивать произведения типа $p_{11}^r p_{00}^m$, где целые неотрицательные числа r, m не превосходят 2. Эта задача решается, начиная с меньших степеней. Известно, что для ковариации мультиномиального вектора:

$$M(ad) = -n\, p_{00}\, p_{11} \tag{12}$$

(см., например, формулу (6.3.5) в монографии [14, с. 153]), а потому несмещенной оценкой для $p_{00}p_{11}$ является (- ad/n). Далее, поскольку справедлива аналогичная (11) формула:

$$M(a^2d) = n(n-1)(n-2)p_{11}p_{00}^2 + n(n-1)p_{11}p_{00}, \tag{13}$$

то с помощью формулы (12) преобразуем формулу (13) к виду:

$$M(a^2d + (n-1)ad) = n(n-1)(n-2)p_{11}p_{00}^2, \tag{14}$$

т.е. несмещенной оценкой $p_{11}p_{00}^2$ является $ad(a + n - 1)\{n(n-1)(n-2)\}^{-1}$.

Следующий шаг — аналогичным образом с помощью формул (12) и (14) получаем несмещенную оценку для $p_{11}^2 p_{00}^2$, а затем и для $D(Z_j)$. Промежуточные формулы опущены из-за громоздкости. Окончательный результат таков:

$$T_j = (b_j + d_j)(c_j + d_j)(a_j + c_j)(a_j + b_j)(n-1)^{-1}.$$

Как легко видеть,

$$\frac{Z_j}{\sqrt{T_j}} = V_j \sqrt{n_j - 1},$$

т.е. в случае одной выборки предлагаемый метод совпадает с классическим.

Общая идея рассматриваемого метода проверки гипотез по совокупности малых выборок состоит в том, что подбирается статистика, математическое ожидание которой для каждой малой выборки равно 0 при справедливости проверяемой гипотезы. Затем для каждой выборки строится несмещенная оценка дисперсии этой статистики. Итоговая статистика критерия для проверки гипотезы — это сумма рассматриваемых статистик для всех малых выборок, деленная на квадратный корень из суммы всех несмещенных оценок дисперсий рассматриваемых статистик. При справедливости нулевой гипотезы эта итоговая статистика имеет в асимптотике стандартное нормальное распределение (при выполнении некоторых математических «условий регулярности», которые обычно выполняются при анализе реальных статистических данных).

Впервые такой способ проверки гипотез по совокупности малых выборок был предложен в монографии [16, раздел 4.5]. Нестандартность постановки состоит в том, что число неизвестных параметров растет пропорционально объему данных, т.е. имеет место т.н. «асимптотика Колмогорова», или асимптотика растущей размерности. Дальнейшее развитие применительно к данным типа «да» — «нет» (или «годен» — «дефектен») шло в рамках теории люсианов как части статистики объектов нечисловой природы (см. главу 8 монографии [13]).

4.5. Проблема множественных проверок статистических гипотез

Практика применения методов прикладной статистики часто выходит за границы классической математико-статистической теории. В качестве примера рассмотрим проверку статистических гипотез.

Базовая теоретическая модель касается проверки одной-единственной статистической гипотезы. На практике же при выполнении того или иного прикладного исследования гипотезы зачастую проверяют неоднократно. При этом, как правило, остается неясным, как влияют результаты предыдущих проверок на характеристики (уровень значимости, мощность) последующих проверок. Есть ли вообще влияние? Как его оценить? Как его учесть при формулировке окончательных выводов?

Изучены лишь некоторые схемы множественных проверок, например, схема последовательного анализа А. Вальда или схема оценивания степени полинома в регрессии путем последовательной проверки адекватности модели (см. главу 6). В таких исключительных постановках удается рассчитать характеристики статистических процедур, включающих множественные проверки статистических гипотез.

Однако в большинстве важных для практики случаев статистические свойства процедур анализа данных, основанных на множественных проверках, остаются пока неизвестными. Примерами являются процедуры нахождения информативных подмножеств признаков (коэффициенты для таких и только таких признаков отличны от 0) в регрессионном анализе или выявления отклонений параметров в автоматизированных системах управления.

В таких системах происходит слежение за большим числом параметров. Резкое изменение значения параметра свидетельствует об изменении режима работы системы, что, как правило, требует управляющего воздействия. Существует теория для определения границ допустимых колебаний одного или фиксированного числа параметров. Например, можно использовать контрольные карты Шухарта или кумулятивных сумм, а также их многомерные аналоги (см. главу 13 в [13]). В подавляющем большинстве постановок, согласно обычно используемым вероятностным моделям, для каждого параметра, находящемся в стабильном («налаженном») состоянии, существует хотя и малая, но положительная вероятность того, что его значение выйдет за заданные границы. Тогда система зафиксирует резкое изменение значения параметра («ложная разладка»). При достаточно большом числе параметров с вероятностью, близкой к 1, будет обнаружено несколько «случайных сбоев», среди которых могут «затеряться» и реальные отказы подсистем. Можно доказать, что при большом числе параметров имеется два крайних случая — независимых (в совокупности) параметров и функционально связанных параметров, а для всех остальных систем вероятность обнаружения резкого отклонения хотя бы у одного параметра лежит между соответствующими вероятностями для этих двух крайних случаев.

Почему трудно изучать статистические процедуры, использующие множественные проверки гипотез? Причина состоит в том, что результаты последовательно проводящихся проверок, как правило, не являются независимыми (в смысле независимости случайных элементов). Более того, последовательность проверок зачастую задается исследователем произвольно.

Проблема множественных проверок статистических гипотез — часть более общей проблемы «стыковки» (сопряжения, последовательного выполнения) статистических процедур. Дело в том, что каждая процедура может применяться лишь при некоторых условиях, а в результате применения предыдущих процедур эти условия могут нарушаться. Например, часто рекомендуют перед восстановлением зависимости (регрессионным анализом) разбить данные на однородные группы с помощью какого-либо алгоритма классификации, а затем строить зависимости для каждой из выделенных групп отдельно. Здесь идет речь о «стыковке» алгоритмов классификации и регрессии. Как вытекает из рассмотрений статьи [17], попадающие в одну однородную группу результаты наблюдений зависимы и их распределение не является нормальным (гауссовым), поскольку они лежат в ограниченной по некоторым направлениям области, причем границы зависят от всей совокупности результатов наблюдений. При этом при росте объема выборки зависимость уменьшается, но ненормальность остается. Распределение результатов наблюдений, попавших в одну группу, приближается не к нормальному, а к усеченному нормальному. Следовательно, алгоритмами регрессионного анализа, основанными на «нормальной теории», пользоваться некорректно. Целесообразно применять непараметрическую или робастную регрессию.

Проблема «стыковки» статистических процедур обсуждается давно. По проблеме «стыковки» проведен ряд исследований, результаты некоторые из которых упомянуты выше, но сколько-нибудь окончательных результатов получено не было. По нашему мнению, на скорое решение проблемы «стыковки» рассчитывать нельзя. Возможно, она является столь же «вечной», как и проблема выбора между средним арифметическим и медианой как характеристиками «центра» выборки.

В качестве примера обсудим одно интересное исследование по проблеме повторных проверок статистических гипотез — работу С.Г. Корнилова [18].

Как уже отмечалось, теоретическое исследование является весьма сложным, сколько-нибудь интересные результаты удается получить лишь для отдельных постановок. Поэтому вполне естественно, что С.Г. Корнилов применил метод статистического моделирования на ЭВМ. Однако нельзя забывать о проблеме качества псевдослучайных чисел. Достоинства и недостатки различных алгоритмов получения псевдослучайных чисел много лет обсуждаются в различных изданиях (см. главу 11 в [13]).

В работе С.Г. Корнилова хорошо моделируется *мышление* статистика-прикладника. Видно, насколько мешает устаревшее представление о том, что для проверки гипотез необходимо задавать определенный уровень значимости. Особенно оно мешает, если в дальнейшем понадобятся дальнейшие проверки. Гораздо удобнее использовать «достигаемый уровень значимости», т.е. вероятность того, что статистика критерия покажет большее отклонение от нулевой гипотезы, чем то отклонение, которое соответствует имеющимся экспериментальным данным. Если есть желание, можно сравнивать «достигаемый уровень значимости» с заданными значениями 0,05 или 0,01. Так, если «достигаемый уровень значимости» меньше 0,01, то нулевая гипотеза отвергается на уровне значимости 0,01, в противном случае — принимается. Следует рассчитывать «достигаемый уровень значимости» всегда, когда для этого есть вычислительные возможности.

Переход к «достигаемому уровню значимости» может избавить прикладника от еще одной трудности, связанной с использованием непараметрических критериев. Дело в том, что их распределения, как правило, дискретны, поскольку эти критерии используют только ранги наблюдений. Поэтому обычно невозможно построить критерий с заданным номинальным уровнем значимости — реальный уровень значимости может принимать лишь конечное число значений, среди которых, как правило, нет ни 0,05, ни 0,01, ни других популярных номинальных значений.

Невозможность построения критических областей критериев с заданными уровнями значимости затрудняет сравнение критериев по мощности, как это продемонстрировано в работе [19]. Есть формальный способ достичь заданного номинального уровня значимости — провести рандомизацию, т.е. при определенном (граничном) значении статистики критерия провести независимый случайный эксперимент, в котором одни исходы (с заданной суммарной вероятностью) приводят к принятию гипотезы, а остальные — к ее отклонению. Однако подобную процедуру рандомизации прикладнику трудно принять — как оправдать то, что одни и те же экспериментальные данные могут быть основанием как для принятия гипотезы, так и для ее отклонения? Вспоминается обложка журнала «Крокодил», на которой один хозяйственник говорит другому:

«Бросим монетку. Упадет гербом — будем строить завод, а упадет решкой — нет». Описанная процедура рандомизации имеет практический смысл лишь при массовой рутинной проверке гипотез, например, при статистическом контроле больших выборок изделий или деталей.

При использовании все еще распространенных критерия Стьюдента и других параметрических статистических критериев — свои проблемы. Такие критерии построены исходя из предположения о том, что функции распределения результатов наблюдений входят в определенные параметрические семейства небольшой размерности. Наиболее распространена гипотеза нормальности распределения. Однако давно известно, что подавляющее большинство реальных распределений результатов измерений не являются нормальными. Об этом говорится, например, в классической для инженеров и организаторов производства монографии проф. В.В. Налимова [20]. Ряд недавно полученных конкретных экспериментальных фактов и теоретических соображений, подтверждающих точку зрения В.В. Налимова, рассмотрен в главе 2.1.

Как же быть? Проверять нормальность распределения своих данных? Но это дело непростое, можно допустить те или иные ошибки, в частности, применяя критерии типа Колмогорова или типа омега-квадрат. Как уже говорилось (в главе 1.2), одна из наиболее распространенных ошибок состоит в том, что в статистики вместо неизвестных параметров подставляют их оценки, но при этом пользуются критическими значениями, рассчитанными для случая, когда параметры полностью известны. Кроме того, для сколько-нибудь надежной проверки нормальности нужны тысячи наблюдений (см. подраздел 2.1). Поэтому в подавляющем большинстве реальных задач нет оснований принимать гипотезу нормальности. В лучшем случае можно говорить о том, что распределение результатов наблюдений мало отличается от нормального.

Как влияют отклонения от нормальности на свойства статистических процедур? Для разных процедур — разный ответ. Если речь идет об отбраковке выбросов — влияние отклонений от нормальности настолько велико, что делает процедуру отбраковки с практической точки зрения эвристической, а не научно обоснованной (см. главу 4.2). Если же речь идет о проверке однородности двух выборок с помощью критерия Стьюдента (при априорном предположении о равенстве дисперсий) или Крамера-Уэлча (при отсутствии такого предположения), то при росте объемов выборок влияние отклонений от нормальности убывает, как это подробно показано в главе 5. Это вытекает из Центральной Предельной Теоремы. Правда, при этом оказывается, что процентные точки распределения Стьюдента не приносят реальной пользы, достаточно использовать процентные точки предельного нормального распределения.

Весьма важна обсуждаемая, в частности, в работе [18] постоянно встающая перед исследователем проблема выбора того или иного статистического критерия для решения конкретной прикладной задачи. Например, как проверять однородность двух независимых выборок числовых результатов наблюдений? Известны параметрические критерии: Стьюдента, Лорда; непараметрические: Крамера — Уэлча, Вилкоксона, Ван-дер-Вардена, Сэвиджа, Мартынова, Смирнова, типа омега-квадрат (Лемана — Розенблатта) и многие другие (см., например, главу 3.1 и справочник [6]). Какой из них выбрать для конкретных расчетов?

Некоторые авторы предлагают формировать технологию принятия статистического решения, согласно которой решающее правило формируется на основе комбинации нескольких критериев. Например, технология может предусматривать проведение «голосования»: если из 5 критериев большинство «высказывается» за отклонение гипотезы, то итоговое решение — отвергнуть ее, в противном случае — принять. Однако в таком подходе нет ничего принципиально нового, просто к уже имеющимся критериям добавляются их комбинации — очередные варианты критериев, тем или иным образом выделяющие критические области в пространствах возможных значений результатов измерений, т.е. попросту увеличивается число рассматриваемых критериев.

Итак, имеется некоторая совокупность критериев. У каждого — свой набор значений уровней значимости и мощностей на возможных альтернативах. Математическая статистика демонстрирует в этой ситуации виртуозную математическую технику для анализа частных случаев и полную беспомощность при выдаче практических рекомендаций. Так, оказывается, что практически каждый из известных критериев является оптимальным в том или ином смысле для какого-то набора нулевых гипотез и альтернатив. Математики изучают асимптотическую эффективность в разных смыслах — по Питмену, по Бахадуру и т.д., но — для узкого класса альтернативных гипотез, обычно для альтернативы сдвига. При попытке переноса асимптотических результатов на конечные объемы выборок возникают новые нерешенные проблемы, связанные, в частности, с численным оцениванием скорости сходимости (см. подраздел 1.4.7). В целом эта область математической статистики может активно развиваться еще многие десятилетия, выдавая превосходные теоремы (которые могут послужить основанием для защит кандидатских и докторских диссертаций, выборов в академики РАН и т.д.), но не давая ничего практике. Хорошо бы, чтобы этот пессимистический прогноз не вполне оправдался!

С точки зрения прикладной статистики необходимо изучать проблему выбора критерия проверки однородности двух независимых выборок. Такое изучение было проведено, в том числе методом статистических испытаний, и в результате был получен вывод о том, что наиболее целесообразно применять критерий Лемана — Розенблатта типа омега-квадрат (см. главу 5).

В литературе по прикладным статистическим методам, как справедливо замечает С.Г. Корнилов в работе [1], имеется масса ошибочных рекомендаций. Чего стоят хотя бы принципиально неверные государственные стандарты СССР по статистическим методам, а также соответствующие им стандарты СЭВ и ИСО, т.е. Международной организации по стандартизации (о них см. главу 13 учебника [13], а также статью [21]). Особо выделяются своим количеством ошибочные рекомендации по применению критерия типа Колмогорова для проверки нормальности. Ошибки есть и в научных статьях, и в нормативных документах (государственных стандартах), и в методических разработках, и даже в вузовских учебниках. К сожалению, нет способа оградить инженера и научного работника, экономиста и менеджера, нуждающихся в применении статистических методов, от литературных источников и нормативно-технических и инструктивно-методических документов с ошибками, неточностями и погрешностями. Единственный способ — либо постоянно поддерживать профессиональные контакты с квалифицированными специалистами по прикладной статистике, либо самому стать таким специалистом.

Как оценить достигаемый уровень значимости конкретного критерия, предусматривающего повторные проверки? Сразу ясно, что в большинстве случаев никакая современная теория математической статистики не поможет. Остается использовать современные компьютеры. Методика статистического моделирования может стать ежедневным рабочим инструментом специалиста, занимающегося применением методов анализа данных. Для этого она должна быть реализована в виде соответствующей диалоговой программной системы. Современные персональные компьютеры позволяют проводить статистическое моделирование весьма быстро (за доли секунд). Можно использовать различные модификации бутстрепа — одного из вариантов применения статистического моделирования (см. [13, 24, 25]).

Проведенное обсуждение показывает, как много нерешенных проблем стоит перед специалистом, занимающимся, казалось бы, рутинным применением стандартных статистических процедур. Теория статистических методов —

молодая наука, ее основные проблемы, по нашему мнению, еще не до конца решены. Много работы как в сравнительно новых областях, например, в анализе нечисловых и интервальных данных, так и в классических.

Литература

1. *Крамер, Г.* Математические методы статистики / Г. Крамер ; перевод с англ. — 2-е изд. — Москва : Мир, 1975. — 648 с.

2. Орлов, А.И. Метод моментов проверки согласия с параметрическим семейством распределений / А.И. Орлов. —Заводская лаборатория. — 1989. — Т. 55. — № 10. — С. 90–93.

3. *Большев, Л.Н.* Избранные труды. Теория вероятностей и математическая статистика / Л.Н. Большев. — Москва : Наука, 1987. — 286 с.

4. *Тюрин, Ю.Н.* Исследования по непараметрической статистике: Непараметрические методы и линейная модель : автореферат диссертации на соискание ученой степени доктора физико-математических наук / Ю.Н. Тюрин ; Московский государственный университет им. М.В. Ломоносова. — Москва, 1985. — 33 с.

5. *Мартынов, Г.В.* Критерии омега-квадрат / Г.В. Мартынов. — Москва : Наука, 1978. — 80 с.

6. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1983. — 416 с.

7. *Артемьев, Б.Г.* Справочное пособие для работников метрологических служб / Б.Г. Артемьев, С.М. Голубов. — Москва : Изд-во стандартов, 1982. — 280 с.

8. *Колмогоров, А.Н.* Элементы теории функций и функционального анализа : учебник / А.Н. Колмогоров, С.В. Фомин. — Москва : Наука, 1972. — 496 с.

9. *Орлов, А.И.* Асимптотическое поведение статистик интегрального типа / А.И. Орлов // Доклады АН СССР. — 1974. — Т. 219. — № 4. — С. 808–811.

10. *Орлов, А.И.* Асимптотическое поведение статистик интегрального типа / А.И. Орлов // Вероятностные процессы и их приложения : межвузовский сборник научных трудов. — Москва : МИЭМ, 1989. — С. 118–123.

11. *Гнеденко, Б.В.* Курс теории вероятностей : учебник / Б.В. Гнеденко. — 7-е изд., испр. — Москва : Эдиториал УРСС, 2001. — 320 с.

12. *Лоэв, М.* Теория вероятностей / М. Лоэв. — Москва : ИЛ, 1962. — 720 с.

13. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 2-е изд., испр. и доп. — Москва : Экзамен, 2003. — 576 с.

14. *Уилкс, С.* Математическая статистика / С. Уилкс. — Москва : Наука, 1967. — 632 с.

15. *Кендалл, М.Дж.* Статистические выводы и связи / М.Дж. Кендалл, А. Стьюарт. — Москва : Наука, 1973. — 900 с.

16. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

17. *Орлов, А.И.* Некоторые вероятностные вопросы теории классификации / А.И. Орлов // Прикладная статистика. Ученые записки по статистике. — Т. 45. — Москва : Наука, 1983. — С. 166–179.

18. *Корнилов, С.Г.* Накопление ошибки первого рода при повторной проверке статистических гипотез. Регламент повторных проверок / С.Г. Корнилов // Заводская лаборатория. — 1996. — Т. 62. — № 5. — С. 45–51.

19. *Камень, Ю.Э.* Реальные и номинальные уровни значимости в задачах проверки статистических гипотез / Ю.Э. Камень, Я.Э. Камень, А.И. Орлов // Заводская лаборатория. — 1986. — Т. 52. — № 12. — С. 55–57.

20. *Налимов, В.В.* Применение математической статистики при анализе вещества / В.В. Налимов. — Москва : Физматгиз, 1960. — 430 с.

21. *Орлов, А.И.* Сертификация и статистические методы (обобщающая статья) / А.И. Орлов. // Заводская лаборатория. — 1997. — Т. 63. — № 3. — С. 55–62.

22. *Орлов, А.И.* Предельная теория непараметрических статистик / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 100. — С. 31–52.

23. *Орлов, А.И.* Метод проверки гипотез по совокупности малых выборок и его применение в теории статистического контроля / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 104. — С. 38–52.

24. *Орлов, А.И.* Взаимосвязь предельных теорем и метода Монте-Карло / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 114. — С. 27–41.

25. *Орлов, А.И.* Применение метода Монте-Карло при изучении свойств статистических критериев однородности двух независимых выборок / А.И. Орлов // Научный журнал КубГАУ. — 2019. — № 154. — С. 55–83.

Контрольные вопросы

1. Сколько выборочных моментов необходимо использовать для проверки согласия с двухпараметрическим семейством функций распределения?

2. Почему методы отбраковки резко выделяющихся результатов наблюдений, основанные на предположении нормальности, нельзя считать научно обоснованными?

3. Какую роль играет условие интегрируемости по Риману — Стилтьесу в предельной теории статистик интегрального типа?

4. Как проверяют независимость альтернативных признаков с помощью таблиц 2×2?

5. Как влияет предварительное выделение однородных групп на проведение регрессионного анализа?

6. Как повлияет проверка однородности двух совокупностей (с помощью критерия Лемана — Розенблатта) на последующую оценку дисперсии по объединенной выборке (в случае подтверждения однородности)?

Темы докладов, рефератов, исследовательских работ

1. На основе метода моментов разработайте критерий согласия с семейством экспоненциальных распределений.

2. Методы отбраковки выбросов и их анализ с точки зрения теории устойчивости статистических процедур.

3. С помощью метода аппроксимации ступенчатыми функциями найдите асимптотическое распределение статистики Колмогорова.

4. Статистический приемочный контроль по альтернативным признакам.

5. Асимптотика Колмогорова в задачах прикладной статистики.

6. Проблема «стыковки» алгоритмов в технологиях обработки статистических данных.

7. Статистическая теория множественных проверок гипотез о разладке с помощью независимых датчиков.

ЧАСТЬ 2. КОНКРЕТНЫЕ СТАТИСТИЧЕСКИЕ МЕТОДЫ

ГЛАВА 5. СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ЧИСЛОВЫХ ВЫБОРОК

Рассмотрим несколько типовых задач анализа числовых данных, часто встречающихся при применении статистических методов в различных областях научных исследований и отраслях народного хозяйства. В настоящей главе выборка моделируется как совокупность независимых одинаково распределенных *числовых* случайных величин.

5.1. Оценивание основных характеристик распределения

Раздел посвящен непараметрическому точечному и интервальному оцениванию характеристик распределения (математического ожидания, медианы, дисперсии, среднего квадратического отклонения, коэффициента вариации) по выборке результатов измерений. Выборочные значения рассматриваются как реализации независимых одинаково распределенных случайных величин с произвольной функцией распределения.

Существенная часть алгоритмов статистического анализа данных исходит из предположения о нормальности распределения результатов наблюдений. Между тем специально проведенные исследования (сводка дана в главе 2.1) показывают, что распределения погрешностей физических измерений, как правило, отличны от нормальных. Из-за отклонений от нормальности свойства алгоритмов могут в одних случаях измениться сравнительно слабо, как при проверке гипотезы однородности математических ожиданий для выборок равного объема (см. следующий раздел), но иногда изменения таковы, что алгоритмы из научных переходят в эвристические. Например, свойства алгоритмов отбраковки выбросов (резко выделяющихся наблюдений) крайне неустойчивы по отношению к отклонениям от нормальности: если зафиксировать правило отбраковки, то крайне неустойчив уровень значимости, а если зафиксировать уровень значимости, то крайне неустойчиво критическое значение (см. раздел 4.2). Поэтому Российской академией статистических методов в 1998 г. была выдвинута задача изучения влияния отклонения от нормальности на свойства всех практически используемых алгоритмов статистического анализа.

Одна из основных задач в области статистических методов — оценивание по выборочным данным характеристик генеральной совокупности, таких, как математическое ожидание, медиана, дисперсия, среднее квадратическое отклонение, коэффициент вариации. Точечные оценки строятся очевидным образом — используют выборочные аналоги теоретических характеристик. Для получения интервальных оценок приходится использовать асимптотическую нормальность выборочных моментов и функций от них.

Пусть исходные данные — это выборка $x_1, x_2, \dots, x_n$, где n — объем выборки. В вероятностной модели выборочные значения $x_1, x_2, \dots, x_n$ рассматриваются как реализации независимых одинаково распределенных случайных величин $X_1, X_2, \dots, X_n$ с общей функцией распределения $F(x) = P(X_i < x)$, $i = 1,2, \dots, n$. Поскольку функция распределения произвольна (с точностью до условий регулярности типа существования моментов), то рассматриваемые задачи доверительного оценивания характеристик распределения являются *непараметрическими*. Существование моментов является скорее математическим ограничением, чем реальным, поскольку практически все реальные статистические данные финитны (т.е. ограничены сверху и снизу, например, шкалой прибора). Для простоты изложения примем это предположение финитности, из которого вытекает существование теоретических моментов любого порядка.

В расчетах будут использоваться выборочное среднее арифметическое:

$$\overline{X} = (X_1 + X_2 + \dots + X_n)/n,$$

выборочная дисперсия:

$$s_0^2 = \{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \dots + (X_n - \overline{X})^2\} / (n-1),$$

выборочное среднее квадратическое отклонение s_0 (квадратный корень из выборочной дисперсии) и некоторые другие выборочные характеристики, которые введем позже.

Точечное и интервальное оценивание математического ожидания. Точечной оценкой для математического ожидания в силу закона больших чисел является выборочное среднее арифметическое $\overline{X}$. В некоторых случаях могут быть использованы и другие оценки. Например, если известно, что распределение симметрично относительно своего центра, то центр распределения является не только математическим ожиданием, но и медианой, а потому для его оценки можно использовать выборочную медиану.

Нижняя доверительная граница для математического ожидания имеет вид:

$$\overline{X} - U(p)\, s_0 / n^{1/2},$$

где: p — доверительная вероятность (истинное значение математического ожидания находится между нижней доверительной границей и верхней доверительной границей с вероятностью, асимптотически равной доверительной); $U(p)$ — число, заданное равенством $\Phi(U(p)) = (1+ p)/2$, где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Например, при p = 95 % (т.е. при p = 0,95) имеем $U(p)$ = 1,96. Функция $U(p)$ имеется в большинстве литературных источников по теории вероятностей и математической статистике (см., например, [1]).

Верхняя доверительная граница для математического ожидания имеет вид:

$$\overline{X} + U(p)\, s_0/ n^{1/2}.$$

Выражения для верхней и нижней доверительных границ получены с помощью центральной предельной теоремы теории вероятностей и теоремы о наследовании сходимости (см. главу 14). Они являются асимптотическими, т.е. становятся тем точнее, чем больше объем выборки. В частности, вероятность попадания истинного значения математического ожидания в интервал между нижней и верхней доверительными границами асимптотически приближается к доверительной вероятности. Но при конечном объеме выборки может незначительно отличаться от нее. Это — недостатки непараметрического подхода. Достоинством же является то, что его можно применять всегда, когда случайная величина имеет математическое ожидание и дисперсию, что в силу финитности (ограниченности шкал) имеет быть практически всегда в реальных ситуациях.

Сопоставим с параметрическим подходом. Обычно в таких случаях предполагают нормальность результатов наблюдений (которой, как уже было отмечено, практически никогда нет). Тогда формулы нижней и верхней доверительных границ для математического ожидания имеют похожий вид, только вместо $U(p)$ стоят квантили распределения Стьюдента. Как известно, при росте объема выборки квантили распределения Стьюдента сходятся к соответствующим квантилям стандартного нормального распределения, так что при больших объемах выборок оба подхода дают близкие результаты. Отметим, что классические доверительные интервалы несколько длиннее, поскольку квантили распределения Стьюдента больше квантилей стандартного нормального распределения, хотя это различие и невелико.

Пример 1. Рассмотрим данные о наработке резцов до отказа, приведенные в табл. 2 раздела 3.1. Для них выборочное среднее арифметическое $\overline{X}$ = 57,88 (это и есть точечная оценка для математического ожидания), выборочная дисперсия s_0^2 = 663,00, объем выборки n = 50. Следовательно, выборочное среднее квадратическое отклонение s_0 = $\sqrt{663{,}00} = 25{,}75$ и согласно приведенным выше формулам при доверительной вероятности p = 0,95 нижняя доверительная граница для математического ожидания такова:

$$57{,}88 - 1{,}96 \cdot 25{,}75 / \sqrt{50} = 57{,}88 - 7{,}14 = 50{,}74,$$

а верхняя доверительная граница есть 57,88 + 7,14 = 65,02.

Если заранее известно, что результаты наблюдения имеют нормальное распределение, то нижняя и верхняя доверительная границы для математического ожидания определяются по формулам:

$$\overline{X} - t(p, n-1)\, s_0 / \sqrt{n}\,,\ \overline{X} + t(p, n-1)\, s_0 / \sqrt{n}$$

соответственно. Эти формулы отличаются от предыдущих тем, что квантиль нормального распределения $U(p)$ заменена на аналогичную квантиль распределения Стьюдента с $(n-1)$ степенью свободы. Другими словами, $t(p, n-1)$ — это число, заданное равенством $ST_{n-1}(p) = (1 + p)/2$, где $ST_{n-1}(x)$ — функция распределения Стьюдента с $(n-1)$ степенью свободы.

Для доверительной вероятности p = 0,95 при объеме выборки n = 50 согласно [1] имеем $t(p, n - 1)$ = 2,0096. Следовательно, нижняя доверительная граница для математического ожидания такова:

$$57{,}88 - 2{,}0096 \times 25{,}75 / \sqrt{50} = 57{,}88 - 7{,}32 = 50{,}56,$$

а верхняя доверительная граница есть 57,88 + 7,32 = 65,20. Таким образом, длина доверительного интервала увеличилась с 14,28 до 14,64, т.е. на 2,5 %.

Отметим, что согласно расчетам, проведенным в главе 4.1, рассматриваемые данные согласуются с гамма-распределением, а не с нормальным распределением, поэтому использование распределения Стьюдента для получения доверительных границ некорректно.

Иногда рекомендуют сначала проверить нормальность результатов наблюдений, а потом, в случае принятия гипотезы нормальности, рассчитывать доверительные границы с использованием квантилей распределения Стьюдента. Однако проверка нормальности — более сложная статистическая процедура, чем оценивание математического ожидания. Кроме того, применение одной статистической процедуры, как правило, нарушает предпосылки следующей процедуры, в частности, независимость результатов наблюдений (см. главу 4.5). Поэтому цепочка статистических процедур, следующих друг за другом, как правило, образует статистическую технологию, свойства которой неизвестны на современном уровне развития статистических методов.

Итак, только непараметрическую статистическую процедуру следует применять для анализа реальных данных. Как правило, встречающиеся на практике распределения не являются нормальными, а потому использование квантилей распределения Стьюдента неправомерно.

Точечное и интервальное оценивание медианы. Точечной оценкой для медианы является выборочная медиана.

Пример 2. Для данных о наработке резцов до отказа (табл. 2 раздела 3.1) объем выборки — четное число, поэтому выборочной медианой является полусумма 25-го и 26-го членов вариационного ряда, т.е. (56 + 56,5)/2 = 56,25.

Чтобы построить доверительные границы для медианы, по доверительной вероятности p находят $U(p)$. Затем вычисляют натуральное число:

$$C(p) = [n/2 - U(p)n^{1/2}/2],$$

где [.] — знак целой части числа. Нижняя доверительная граница для медианы имеет вид (при $C(p) \geq 1$; если $p = 0,95$ и $U(p) = 1,96$, то $C(p) \geq 1$ при $n \geq 8$):

$$X(C(p)),$$

где $X(i)$ — член с номером i вариационного ряда, построенного по исходной выборке (т.е. i-я порядковая статистика). Верхняя доверительная граница для медианы имеет вид:

$$X(n + 1 - C(p)).$$

Теоретическое основание для приведенных доверительных границ содержится в литературе по порядковым статистикам (см., например, монографию [2, с. 68]).

Пример 3. Для данных о наработке резцов до отказа $n = 50$. Рассмотрим как обычно, доверительную вероятность $p = 0,95$. Тогда:

$$C(p) = [50/2 - 1,96\sqrt{50}/2] = [18,07] = 18.$$

Следовательно, нижней доверительной границей является $X(18) = 47{,}5$, а верхней доверительной границей $X(50 + 1 - 18) = X(33) = 61{,}5$.

Поскольку в случае нормального распределения медиана совпадает с математическим ожиданием, то каких-либо специальных способов ее оценивания в классическом случае нет.

Точечное и интервальное оценивание дисперсии. Точечной оценкой дисперсии является выборочная дисперсия s_0^2. Эта оценка является несмещенной и состоятельной. Доверительные границы находятся с помощью величины:

$$d^2 = (m_4 - ((n-1)/n)^4 s_0^4)/n,$$

где m_4 — выборочный четвертый центральный момент, т.е.

$$m_4 = \{(X_1 - \overline{X})^4 + (X_2 - \overline{X})^4 + \ldots + (X_n - \overline{X})^4\}/n.$$

Нижняя доверительная граница для дисперсии такова:

$$s_0^2 - U(p)d,$$

где s_0^2 — выборочная дисперсия; $U(p)$ — квантиль нормального распределения порядка $(1+p)/2$, а d — положительный квадратный корень из величины d^2, введенной выше.

Верхняя доверительная граница для дисперсии имеет вид:

$$s_0^2 + U(p)d.$$

При выводе приведенных соотношений используется асимптотическая нормальность выборочной дисперсии, установленная, например, в монографии [3, с. 419]. Соответственно, непараметрический доверительный интервал является асимптотическим. В классическом случае точечная оценка имеет тот же вид, а вот доверительные границы находят с помощью квантилей распределения хи-квадрат с числом степеней свободы, на 1 меньшим объема выборки. Отметим, что в случае нормального распределения четвертый момент в 3 раза больше квадрата дисперсии, а потому можно оценить d^2 как $2s_0^4/n$. Это дает быстрый способ для интервальной оценки дисперсии в нормальном случае.

Пример 4. Для данных о наработке резцов до отказа объем выборки n = 50, выборочная дисперсия s_0^2 = 663,00, четвертый выборочный момент m_4 = 1702050,71. Поэтому:

$$d^2 = (1702050{,}71 - ((50-1)/50)^4 663{,}00^2)/50 = 25932{,}13.$$

Тогда d = 161,03. Для доверительной вероятности p =0,95 нижняя доверительная граница для дисперсии случайной величины такова:

$$663{,}00 - 1{,}96 \cdot 161{,}03 = 663{,}00 - 315{,}63 = 347{,}37,$$

а верхняя доверительная граница для дисперсии есть 663,00 + 315,63 = 978,63.

Пример 5. В случае нормального распределения с целью быстрого получения доверительного интервала величина d^2 оценивается как:

$$(2s_0^4)/n = (2 \cdot 663{,}00^2)/50 = 17582{,}76,$$

а потому d = 132,6. Для доверительной вероятности p =0,95 нижняя доверительная граница для дисперсии заменяется на:

$$663{,}00 - 1{,}96 \cdot 132{,}6 = 663{,}00 - 259{,}90 = 403{,}10,$$

а верхняя доверительная граница — на 663,00 + 259,90 = 922,9.

Сужение границ для дисперсии вполне естественно. Данные о наработке резцов до предельного состояния (т.е. до отказа) соответствуют гамма-распределению, а это распределение является асимметричным, с «тяжелым» правым «хвостом». Последнее означает, что плотность убывает заметно медленнее, чем для нормального распределения. Как следствие, четвертый момент заметно больше, чем для нормального распределения с теми же математическим ожиданием и дисперсией. А потому больше и параметр d. Из проведенных расчетов видно, что использование алгоритмов расчетов, соответствующих нормальному распределению, в ситуации, когда распределение результатов наблюдений существенно отличается от нормального, может привести к заметному искажению выводам.

Пример 6. В классическом случае нормального распределения исходят из того, что величина $(n-1)\ s_0^2/\sigma^2$ имеет распределение хи-квадрат с $(n-1)$ степенью свободы. Для доверительной вероятности p =0,95 следует рассмотреть неравенство:

$$31{,}555 < (n-1)\ s_0^2/\sigma^2 < 70{,}222,$$

справедливое с вероятностью 0,95, поскольку

$$F(31{,}555) = 0{,}025,\ F(70{,}222) = 0{,}975,$$

где $F(x)$ — функция хи-квадрат распределения с 49 степенями свободы. Следовательно, нижняя доверительная граница для дисперсии нормально распределенной случайной величины такова:

$$(n-1)\ s_0^2/70{,}222 = (49 \cdot 663{,}00)/70{,}222 = 462{,}63,$$

а верхняя доверительная граница есть

$$(n-1)\ s_0^2/31{,}555 = (49 \cdot 663{,}00)/31{,}555 = 1029{,}54.$$

Полученный доверительный интервал не является симметричным относительно точечной оценки. Нижняя доверительная граница больше, чем в примерах 4 и 5, но и верхняя доверительная граница тоже больше. Несимметричность доверительного интервала в примере 6 приводит к тому, что его трудно сопоставить с симметричными интервалами примеров 4 и 5. Что же касается практических рекомендаций, то они однозначны: поскольку обычно нет основания считать, что результаты измерений имеют нормальное распределение, то при анализе реальных данных надо пользоваться непараметрическими методами.

Точечное и интервальное оценивание среднего квадратического отклонения. Точечной оценкой является выборочное среднее квадратическое отклонение, т.е. неотрицательный квадратный корень из выборочной дисперсии. Дисперсия рассматриваемой случайной величины — выборочного среднего квадратического отклонения s_0 — оценивается как дробь

$$d^2 / (4 s_0^2).$$

Нижняя доверительная граница для среднего квадратического отклонения исходной случайной величины имеет вид:

$$s_0 - U(p)d / (2s_0),$$

где s_0^2 — выборочная дисперсия, $U(p)$ — квантиль нормального распределения порядка $(1+p)/2$, а d — положительный квадратный корень из величины d^2, введенной выше при оценивании дисперсии.

Верхняя доверительная граница для среднего квадратического отклонения исходной случайной величины имеет вид:

$$s_0 + U(p)d\,/\,(2s_0).$$

Пример 7. Для данных о наработке резцов до отказа точечной оценкой для среднего квадратического отклонения является $s_0 = \sqrt{663{,}00} = 25{,}75$. При доверительной вероятности p = 0,95 нижняя доверительная граница такова:

$$25{,}75 - 1{,}96 \times 161{,}03\ /\ (2 \times 25{,}75) = 25{,}75 - 6{,}13 = 19{,}62.$$

Соответственно верхняя доверительная граница симметрична нижней относительно точечной оценки и равна = 25,75 + 6,13 = 31,88.

Правила интервального оценивания для среднего квадратического отклонения получены из аналогичных правил для оценивания дисперсии с помощью метода линеаризации (см. главу 14 или, в наиболее общем виде, [4, п. 2.4]). Доверительный интервал является симметричным, непараметрическим и асимптотическим.

Есть и другой способ доверительного оценивания. Поскольку среднее квадратическое отклонение — это квадратный корень их дисперсии, то доверительные границы можно получить, извлекая квадратные корни из одноименных границ для дисперсии.

Пример 8. Для данных о наработке резцов до отказа при доверительной вероятности p = 0,95 согласно примеру 4 доверительный интервал для дисперсии — это [347,37; 978,63]. Извлекая квадратные корни, получаем доверительный интервал [18,64; 31,28] для среднего квадратического отклонения, соответствующий тому же значению доверительной вероятности. Он не является симметричным относительно точечной оценки. Его длина 12,64 несколько больше длины симметричного доверительного интервала 12,26 в примере 7.

Подход, основанный на гипотезе нормальности распределения результатов наблюдения, связан с использованием распределения хи-квадрат и сводится к извлечению квадратных корней из доверительных границ для дисперсии.

Пример 9. Формально применяя классический подход к данным о наработке резцов до отказа, исходим из доверительного интервала для дисперсии [462,63; 1029,54], соответствующего доверительной вероятности p = 0,95. Извлекая квадратные корни, находим доверительный интервал для среднего квадратического отклонения [21,51; 32,09]. Как и следовало ожидать, длина этого несимметричного интервала 10,58 меньше длины непараметрического доверительного интервала, равной 12,68.

Точечное и интервальное оценивание коэффициента вариации. Коэффициент вариации $V = \sigma/M(X)$ широко используется при анализе конкретных технических, экономических, социологических, медицинских и иных данных (поскольку они, как правило, положительны). Точечной оценкой теоретического коэффициента вариации V является выборочный коэффициент вариации:

$$V_n = s_0 / \overline{X}.$$

Дисперсия выборочного коэффициента вариации состоятельно оценивается с помощью вспомогательной величины:

$$D^2 = (V_n^4 - V_n^2 / 4 + m_4 / (4 s_0^2\ \overline{X}^2) - m_3 / \overline{X}^3) / n,$$

где $\overline{X}$ — выборочное среднее арифметическое, s_0^2 — выборочная дисперсия, m_3 — выборочный третий центральный момент, т.е.:

$$m_3 = \{(X_1 - \overline{X})^3 + (X_2 - \overline{X})^3 + \ldots + (X_n - \overline{X})^3\} / n,$$

где m_4 — выборочный четвертый центральный момент (см. выше), V_n — выборочный коэффициент вариации, n — объем выборки.

Нижняя доверительная граница для (теоретического) коэффициента вариации исходной случайной величины имеет вид:

$$V_n - U(p)\, D,$$

где V_n — выборочный коэффициент вариации, $U(p)$ — квантиль нормального распределения порядка $(1+p)/2$ (как и ранее), D — положительный квадратный корень из величины D^2, введенной выше.

Верхняя доверительная граница для (теоретического) коэффициента вариации исходной случайной величины имеет вид:

$$V_n + U(p)\, D.$$

Как и в предыдущих случаях, доверительный интервал является непараметрическим и асимптотическим. Он получен в результате применения специальной технологии вывода асимптотических соотношений прикладной статистики. Эта технология в качестве первого шага использует многомерную центральную предельную теорему, примененную к сумме векторов, координаты которых — степени исходных случайных величин. Второй шаг — преобразование предельного многомерного нормального вектора с целью получения интересующего исследователя вектора. При этом используются соображения линеаризации и отбрасываются бесконечно малые величины. Третий шаг — строгое обоснование полученных результатов на стандартном для асимптотических математико-статистических рассуждений уровне. При этом обычно приходится использовать необходимые и достаточные условия наследования сходимости, полученные в монографии [4, п. 2.4] (см. подробные формулировки в главе 14). Именно таким образом были получены приведенные выше результаты для выборочного коэффициента вариации. Формулы оказались существенно более сложными, чем в предыдущих случаях. Это объясняется тем, что выборочный коэффициент вариации — функция двух выборочных моментов, а ранее рассматривались либо выборочные моменты поодиночке, либо функция от одного выборочного момента — выборочной дисперсии.

Пример 10. Для данных о наработке резцов до отказа выборочное среднее арифметическое $\overline{X}$ = 57,88, выборочная дисперсия s_0^2 = 663,00, выборочное среднее квадратическое отклонение s_0 = 25,75, выборочный третий центральный момент m_3 = 14 927,91, выборочный четвертый центральный момент m_4 = 1 702 050,71. Следовательно, выборочный коэффициент вариации таков:

$$V_n = 25{,}75 / 57{,}88 = 0{,}4449.$$

Рассчитаем значение вспомогательной величины:

$$D^2 = ((0{,}4449)^4 - (0{,}4449)^2/4 + 1\ 702\ 050{,}71/\ (4 \times 663{,}00 \times (57{,}88)^2) -$$
$$- 14\ 927{,}91/(57{,}88)^3)/50 = (0{,}0392 - 0{,}0495 + 0{,}1916 - 0{,}0770)/50 =$$
$$= 0{,}1043/50 = 0{,}002086.$$

Следовательно, D = 0,04567. При доверительной вероятности p = 0,95 нижняя доверительная граница для теоретического коэффициента вариации имеет вид:

$$0{,}4449 - 1{,}96 \times 0{,}04567 = 0{,}4449 - 0{,}0895 = 0{,}3554,$$

а верхняя доверительная граница такова:

$$0,4449 + 0,0895 = 0,5344.$$

Среди классических результатов математической статистики, основанных на гипотезе нормальности результатов наблюдений, нет методов нахождения доверительных границ для коэффициента вариации, поскольку задача построения таких границ не выражается в терминах обычно используемых распределений, например, распределений Стьюдента, Фишера и хи-квадрат.

Примеры применения доверительных границ для коэффициентов вариации при решении прикладных задач приведены, например, в работе [5], посвященной анализу технических характеристик и показателей качества эластомерных (резинотехнических) материалов и изделий.

Итак, сформулированы правила непараметрического оценивания обычно используемых характеристик распределения случайной величины. Эти правила основаны на асимптотических результатах теории вероятностей и математической статистики. Использование методов, разработанных в предположении нормальности распределения, может привести к заметно искаженным выводам в ситуации, когда гипотеза нормальности не выполнена. Практические рекомендации таковы: при анализе реальных данных следует использовать непараметрические доверительные границы.

5.2. Методы проверки однородности характеристик двух независимых выборок

В прикладных исследованиях часто возникает необходимость выяснить, различаются ли генеральные совокупности, из которых взяты две независимые выборки. Например, надо выяснить, влияет ли способ упаковки подшипников на их потребительские качества через год после хранения. Или: отличается ли потребительское поведение мужчин и женщин. Если отличается — рекламные ролики и плакаты надо делать отдельно для мужчин и отдельно для женщин. Если нет — рекламная кампания может быть единой.

В математико-статистических терминах постановка задачи такова: имеются две выборки $x_1, x_2, ..., x_m$ и $y_1, y_2, ..., y_n$ (т. е. наборы из m и n действительных чисел), требуется проверить их однородность. Термин «однородность» уточняется ниже.

Противоположным понятием является «различие». Можно переформулировать задачу: требуется проверить, есть ли различие между выборками. Если различия нет, то для дальнейшего изучения две рассматриваемые выборки часто объединяют в одну.

Например, в маркетинге важно выделить сегменты потребительского рынка. Если установлена однородность двух выборок, то возможно объединение сегментов, из которых они взяты, в один. В дальнейшем это позволит осуществлять по отношению к ним одинаковую маркетинговую политику (проводить одни и те же рекламные мероприятия и т.п.). Если же установлено различие, то поведение потребителей в двух сегментах различно, объединять эти сегменты нельзя, и могут понадобиться различные маркетинговые стратегии, своя для каждого из этих сегментов.

Традиционный метод проверки однородности (критерий Стьюдента). Для дальнейшего критического разбора опишем традиционный статистический метод проверки однородности. Он широко использовался в течение всего XX в. Хотя к настоящему времени этот метод устарел (см. ниже), но продолжает встречаться в учебной литературе, и потому и применяться для анализа конкретных данных.

При использовании традиционного метода проверки однородности вычисляют выборочные средние арифметические в каждой выборке:

$$\bar{x} = \frac{1}{m}\sum_{1\le i\le m} x_i, \qquad y = \frac{1}{n}\sum_{1\le i\le n} y_i ,$$

затем выборочные дисперсии:

$$s_x^2 = \frac{1}{m-1}\sum_{1\le i\le m}(x_i - \bar{x})^2 , \; s_y^2 = \frac{1}{n-1}\sum_{1\le i\le n}(y_i - \bar{y})^2$$

и статистику Стьюдента *t*, на основе которой принимают решение:

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{(m-1)s_x^2 + (n-1)s_y^2}}\sqrt{\frac{mn(m+n-2)}{m+n}} . \qquad (1)$$

По заданному уровню значимости α и числу степеней свободы $(m + n - 2)$ из таблиц распределения Стьюдента находят критическое значение $t_{кр}$. Если $|t| > t_{кр}$, то гипотезу однородности (отсутствия различия) отклоняют, если же $|t| \le t_{кр}$, то принимают. (При односторонних альтернативных гипотезах вместо условия $|t|>t_{кр}$ проверяют, что $t>t_{кр}$; эту постановку рассматривать не будем, так как в ней нет принципиальных отличий от обсуждаемой здесь.)

Рассмотрим условия применимости традиционного метода проверки однородности, основанного на использовании статистики t Стьюдента, а также обсудим современные методы проверки однородности двух выборок.

Вероятностная модель порождения данных. Для обоснованного применения эконометрических методов необходимо прежде всего построить и обосновать вероятностную модель порождения данных. При проверке однородности двух выборок общепринята модель, в которой $x_1, x_2, ..., x_m$ рассматриваются как результаты m независимых наблюдений некоторой случайной величины X с функцией распределения $F(x)$, неизвестной статистику, а $y_1, y_2, ..., y_n$ — как результаты n независимых наблюдений, вообще говоря, другой случайной величины Y с функцией распределения $G(x)$, также неизвестной статистику. Предполагается также, что наблюдения в одной выборке не зависят от наблюдений в другой, поэтому выборки и называют независимыми.

Возможность применения модели в конкретной реальной ситуации требует обоснования. Независимость и одинаковая распределенность результатов наблюдений, входящих в выборку, могут быть установлены или исходя из методики проведения конкретных наблюдений, или путем проверки статистических гипотез независимости и одинаковой распределенности с помощью соответствующих критериев [1].

Если проведено $(m + n)$ измерений объемов продаж в $(m + n)$ торговых точках, то описанную выше модель, как правило, можно применять. Если же, например, x_i и y_i — объемы продаж одного и того же товара до и после определенного рекламного воздействия, то рассматриваемую модель применять нельзя. В последнем случае используют модель связанных выборок. В ней обычно строят новую выборку $z_i = x_i - y_i$ и используют статистические методы анализа одной выборки, а не двух. Методы проверка однородности для связанных выборок рассматривается в разделах 5.5 и 5.6.

При дальнейшем изложении принимаем описанную выше вероятностную модель двух выборок.

Уточнения понятия однородности. Понятие «однородность», т. е. «отсутствие различия», может быть формализовано в терминах вероятностной модели различными способами.

Наивысшая степень однородности достигается, если обе выборки взяты из одной и той же генеральной совокупности, т. е. справедлива нулевая гипотеза:

$$H_0 : F(x)=G(x) \text{ } \textit{при всех } x.$$

Отсутствие однородности означает, что верна альтернативная гипотеза, согласно которой:

$$H_1 : F(x_0) \neq G(x_0)$$

хотя бы при одном значении аргумента x_0. Если гипотеза H_0 принята, то выборки можно объединить в одну, если нет — то нельзя.

В некоторых случаях целесообразно проверять не совпадение функций распределения, а совпадение некоторых характеристик случайных величин X и Y — математических ожиданий, медиан, дисперсий, коэффициентов вариации и др. Например, однородность математических ожиданий означает, что справедлива гипотеза:

$$H'_0 : M(X) = M(Y),$$

где $M(X)$ и $M(Y)$ — математические ожидания случайных величин X и Y, результаты наблюдений над которыми составляют первую и вторую выборки соответственно. Доказательство различия между выборками в рассматриваемом случае — это доказательство справедливости альтернативной гипотезы:

$$H'_1 : M(X) \neq M(Y).$$

Если гипотеза H_0 верна, то и гипотеза H'_0 верна, но из справедливости H'_0, вообще говоря, не следует справедливость H_0. Математические ожидания могут совпадать для различающихся между собой функций распределения. В частности, если в результате обработки выборочных данных принята гипотеза H'_0, то отсюда *не следует,* что две выборки можно объединить в одну. Однако в ряде ситуаций целесообразна проверка именно гипотезы H'_0. Например, пусть функция спроса на определенный товар или услугу оценивается путем опроса потребителей (первая выборка) или с помощью данных о продажах (вторая выборка). Тогда маркетологу важно проверить гипотезу об отсутствии систематических расхождений результатов этих двух методов, т.е. гипотезу о равенстве математических ожиданий. Другой пример — из производственного менеджмента. Пусть изучается эффективность управления бригадами рабочих на предприятии с помощью двух организационных схем, результаты наблюдения — объем производства на одного члена бригады (производительность), а показатель эффективности организационной схемы — средний (по предприятию) объем производства на одного рабочего. Тогда для сравнения эффективности препаратов достаточно проверить гипотезу H'_0.

Классические условия применимости критерия Стьюдента. Согласно математико-статистической теории должны быть выполнены два классических условия применимости критерия Стьюдента, основанного на использовании статистики t, заданной формулой (1):

а) результаты наблюдений имеют нормальные распределения:

$$F(x)=N(x;\, m_1,\, \sigma_1^2),\ G(x)=N(x;\, m_2,\, \sigma_2^2)$$

с математическими ожиданиями m_1 и m_2 и дисперсиями σ_1^2 и σ_2^2 в первой и во второй выборках соответственно;

б) дисперсии результатов наблюдений в первой и второй выборках совпадают:

$$D(X)=\sigma_1^2=D(Y)=\sigma_2^2.$$

Если условия а) и б) выполнены, то нормальные распределения $F(x)$ и $G(x)$ отличаются только математическими ожиданиями, а поэтому *обе* гипотезы H_0 и H'_0 сводятся к гипотезе:

$$H''_0 : m_1=m_2,$$

а *обе* альтернативные гипотезы H_1 и H'_1 сводятся к гипотезе

$$H''_1 : m_1 \neq m_2.$$

Если условия а) и б) выполнены, то статистика t при справедливости H''_0 имеет распределение Стьюдента с $(m + n - 2)$ степенями свободы. Только в этом случае описанный выше традиционный метод обоснован безупречно. Если хотя бы одно из условий а) и б) не выполнено, то нет никаких оснований считать, что статистика t имеет распределение Стьюдента, поэтому применение традиционного метода, строго говоря, не обосновано. Обсудим возможность проверки этих условий и последствия их нарушений.

Имеют ли результаты наблюдений нормальное распределение? Как показано в главе 2.1, априори нет оснований предполагать нормальность распределения результатов экономических, технико-экономических, технических, медицинских и иных наблюдений. Следовательно, нормальность надо проверять. Разработано много статистических критериев для проверки нормальности распределения результатов наблюдений [1]. Однако проверка нормальности — более сложная и трудоемкая статистическая процедура, чем проверка однородности (как с помощью статистики t Стьюдента, так и с использованием непараметрических критериев, рассматриваемых ниже).

Для достаточно надежного установления нормальности требуется весьма большое число наблюдений. В главе 2.1 показано, что для того, чтобы гарантировать, что функция распределения результатов наблюдений отличается от некоторой нормальной не более чем на 0,01 (при любом значении аргумента), требуется порядка 2 500 наблюдений. В большинстве технических, экономических, медицинских и иных исследований число наблюдений существенно меньше.

Как уже отмечалось, есть и одна общая причина отклонений от нормальности: любой результат наблюдения записывается конечным (обычно 2–5) количеством цифр, а с математической точки зрения вероятность такого события равна 0. Точнее, для случайной величины с непрерывной плотностью распределения вероятность попадания в счетное множество рациональных чисел равна 0. Следовательно, в статистических методах распределение результатов наблюдений практически всегда более или менее отличается от нормального распределения.

Последствия нарушения условия нормальности. Если условие а) не выполнено, то распределение статистики t не является распределением Стьюдента. Однако при справедливости H'_0 и условия б) распределение статистики t при росте объемов выборок приближается к стандартному нормальному распределению $\Phi(x)=N(x;\,0,\,1)$. К этому же распределению приближается распределение Стьюдента при возрастании числа степеней свободы. Другими словами, несмотря на нарушение условия нормальности традиционный метод (критерий Стьюдента) можно использовать для проверки гипотезы H'_0 при больших объемах выборок. При этом вместо таблиц распределения Стьюдента достаточно пользоваться таблицами стандартного нормального распределения $\Phi(x)$.

Сформулированное в предыдущем абзаце утверждение справедливо для любых функций распределения $F(x)$ и $G(x)$ таких, что $M(X)=M(Y)$, $D(X)=D(Y)$ и выполнены некоторые внутриматематические условия, обычно считающиеся справедливыми в реальных задачах. Если же $M(X)\neq M(Y)$, то нетрудно вычислить, что при больших объемах выборок:

$$P(t\leq x)\approx\Phi(x-a_{mn}), \qquad (2)$$

где

$$a_{mn} = \frac{\sqrt{mn}[M(X) - M(Y)]}{\sqrt{mD(X) + nD(Y)}}. \qquad (3)$$

Формулы (2)–(3) позволяют приближенно вычислять мощность t-критерия (точность возрастает при увеличении объемов выборок m и n).

О проверке условия равенства дисперсий. Иногда условие б) вытекает из методики получения результатов наблюдений, например, когда с помощью одного и того же прибора или методики m раз измеряют характеристику первого объекта и n раз — второго, а параметры распределения погрешностей измерения при этом не меняются. Однако ясно, что в постановках большинства исследовательских и практических задач нет основании априори предполагать равенство дисперсий.

Целесообразно ли проверять равенство дисперсий статистическими методами, например, как это иногда предлагают, с помощью F-критерия Фишера? Этот критерий основан на нормальности распределений результатов наблюдений. А от нормальности неизбежны отклонения (см. выше). Причем хорошо известно, что в отличие от t-критерия распределение F-критерия Фишера сильно меняется при малых отклонениях от нормальности [3]. Кроме того, F-критерий отвергает гипотезу $D(X)=D(Y)$ лишь при большом различии выборочных дисперсий. Так, для данных [1] о двух группах результатов химических анализов отношение выборочных дисперсий равно 1,95, т.е. существенно отличается от 1. Тем не менее, гипотеза о равенстве теоретических дисперсий принимается на 1 %-м уровне значимости. Следовательно, при проверке однородности применение F-критерия для предварительной проверки равенства дисперсий нецелесообразно.

Итак, в большинстве технических, экономических, медицинских и иных задач условие б) нельзя считать выполненным, а проверять его нецелесообразно.

Последствия нарушения условия равенства дисперсий. Если объемы выборок m и n велики, то можно показать, что распределение статистики t описывается с помощью только математических ожиданий $M(X)$ и $M(Y)$, дисперсий $D(X)$, $D(Y)$ и отношения объемов выборок, а именно:

$$P(t \leq x) \approx \Phi(b_{mn}x - a_{mn}), \qquad (4)$$

где a_{mn} определено формулой (3),

$$b_{mn}^{2} = \frac{\lambda D(X) + D(Y)}{D(X) + \lambda D(Y)}, \quad \lambda = \frac{m}{n}. \tag{5}$$

Если $b_{mn} \neq 1$, то распределение статистики t отличается от распределения, заданного формулой (2), полученной в предположении равенства дисперсий. Когда $b_{mn}=1$? В двух случаях — при $m = n$ и при $D(X) = D(Y)$. Таким образом, при больших и равных объемах выборок требовать выполнения условия б) нет необходимости. Кроме того, ясно, что если объемы выборок мало различаются, то b_{mn} близко к 1. Так, для данных [1] о двух группах результатов химических анализов имеем $b^*_{mn}=$ 0,987, где b^*_{mn} — оценка b_{mn} , полученная заменой в формуле (5) теоретических дисперсий на их выборочные оценки.

Область применимости традиционного метода проверки однородности с помощью критерия Стьюдента. Подведем итоги рассмотрения t-критерия. Он позволяет проверять гипотезу H'_0 о равенстве математических ожиданий, но не гипотезу H_0 о том, что обе выборки взяты из одной и той же генеральной совокупности. Классические условия применимости критерия Стьюдента в подавляющем большинстве технических, экономических, медицинских и иных задач не выполнены. Тем не менее, при больших и примерно равных объемах выборок его можно применять. При конечных объемах выборок традиционный метод носит неустранимо приближенный характер.

Критерий Крамера — Уэлча равенства математических ожиданий. Вместо критерия Стьюдента целесообразно для проверки H'_0 использовать критерий Крамера — Уэлча [6], основанный на статистике:

$$T = \frac{\sqrt{mn}(\bar{x} - \bar{y})}{\sqrt{ns_x^2 + ms_y^2}}. \tag{6}$$

Критерий Крамера — Уэлча имеет прозрачный смысл — разность выборочных средних арифметических для двух выборок делится на естественную оценку среднего квадратического отклонения этой разности. Естественность указанной оценки состоит в том, что неизвестные статистику дисперсии заменены их выборочными оценками. Из многомерной центральной предельной теоремы и из теорем о наследовании сходимости [4] вытекает (см. главу 14), что при росте объемов выборок распределение статистики T Крамера — Уэлча сходится к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Итак, при справедливости H'_0 и больших объемах выборок распределение статистики T приближается с помощью стандартного нормального распределения $\Phi(x)$, из таблиц которого следует брать критические значения.

При $m=n$, как следует из формул (1) и (6), $t=T$. При $m\neq n$ этого равенства нет. В частности, при s_x^2 в (1) стоит множитель $(m-1)$, а в (6) — множитель n.

Если $M(X)\neq M(Y)$, то при больших объемах выборок:

$$P(T\leq X)\approx\Phi(x-c_{mn}), \qquad (7)$$

где

$$c_{mn}=\frac{\sqrt{mn}[M(X)-M(Y)]}{\sqrt{nD(X)+mD(Y)}}. \qquad (8)$$

При $m=n$ или $D(X)=D(Y)$, согласно формулам (3) и (8), $a_{mn}=c_{mn}$, в остальных случаях равенства нет.

Из асимптотической нормальности статистики T, формул (7) и (8) следует, что правило принятия решения для критерия Крамера — Уэлча выглядит так:

- если $|T|\leq\Phi^{-1}\left(1-\alpha/2\right)$, то гипотеза однородности (равенства) математических ожиданий принимается на уровне значимости ;
- если же $|T|>\Phi^{-1}\left(1-\alpha/2\right)$, то гипотеза однородности (равенства) математических ожиданий отклоняется на уровне значимости α.

В прикладной статистике наиболее часто применяется уровень значимости $\alpha=0{,}05$. Тогда значение модуля статистики T Крамера — Уэлча надо сравнивать с граничным значением $\Phi^{-1}\left(1-\alpha/2\right)=1{,}96$.

Из сказанного выше следует, что применение критерия Крамера — Уэлча не менее обосновано, чем применение критерия Стьюдента. Дополнительное преимущество — не требуется равенства дисперсий $D(X)=D(Y)$. Распределение статистики T не является распределением Стьюдента, однако и распределение статистики t, как показано выше, не является таковым в реальных ситуациях.

Распределение статистики T при объемах выборок $m=n=6$, 8, 10, 12 и различных функциях распределений выборок $F(x)$ и $G(x)$ изучено нами совместно с Ю.Э. Камнем и Я.Э. Камнем методом статистических испытаний (Монте-Карло). Рассмотрены различные варианты функций распределения $F(x)$ и $G(x)$. Результаты показывают, что даже при таких небольших объемах выборок точность аппроксимации предельным стандартным нормальным распределением вполне удовлетворительна. Поэтому представляется целесообразным во всех тех случаях, когда в настоящее время используется критерий Стьюдента, заменить его на критерий Крамера — Уэлча. Конечно, такая замена потребует переделки ряда нормативно-технических и методических документов, исправления учебников и учебных пособий для вузов.

Пример 1. Пусть объем первой выборки $m = 120, \bar{x} = 13{,}7, s_x = 5{,}3.$ Для второй выборки $n = 541, \bar{y} = 14{,}1, s_y = 8{,}4$. Вычислим величину статистики Крамера — Уэлча:

$$T = \frac{\sqrt{mn}(\bar{x}-\bar{y})}{\sqrt{ns_x^2 + ms_y^2}} = \frac{\sqrt{120\times 541}(13{,}7-14{,}1)}{\sqrt{541\times 5{,}3^2 + 120\times 8{,}4^2}} = \frac{\sqrt{64920}(-0{,}4)}{\sqrt{541\times 28{,}09 + 120\times 141{,}12}} =$$

$$= \frac{254{,}79\times(-0{,}4)}{\sqrt{15196{,}69 + 16934.4}} = \frac{-101{,}916}{\sqrt{32131{,}09}} = \frac{-101{,}916}{179{,}25} = -0{,}57.$$

Поскольку полученное значение по абсолютной величине меньше 1,96, то гипотеза однородности математических ожиданий принимается на уровне значимости 0,05.

Непараметрические методы проверки однородности. В большинстве технических, экономических, медицинских и иных задач представляет интерес не проверка равенства математических ожиданий или иных характеристик распределения, а обнаружение различия генеральных совокупностей, из которых извлечены выборки, т.е. проверка гипотезы H_0. Методы проверки гипотезы H_0 позволяют обнаружить не только изменение математического ожидания, но и любые иные изменения функции распределения результатов наблюдений при переходе от одной выборки к другой (увеличение разброса, появление асимметрии и т. д.). Как установлено выше, методы, основанные на использовании статистик t Стьюдента и T Крамера-Уэлча, не позволяют проверять гипотезу H_0. Априорное предположение о принадлежности функций распределения $F(x)$ и $G(x)$ к какому-либо определенному параметрическому семейству (например, семействам нормальных, логарифмически нормальных, распределений Вейбулла — Гнеденко, гамма-распределений и др.), как также показано выше, обычно нельзя достаточно надежно обосновать. Поэтому для проверки H_0 следует использовать методы, пригодные при любом виде $F(x)$ и $G(x)$, т.е. непараметрические методы. (Напомним, что термин «непараметрический метод» означает, что при использовании этого метода нет необходимости предполагать, что функции распределения результатов наблюдений принадлежат какому-либо определенному параметрическому семейству.)

Для проверки гипотезы H_0 разработано много непараметрических методов — критерии Смирнова, типа омега-квадрат (Лемана — Розенблатта), Вилкоксона (Манна — Уитни), Ван-дер-Вардена, Сэвиджа, хи-квадрат и др. [1, 2, 7]. Распределения статистик всех этих критериев при справедливости H_0 не зависят от конкретного вида совпадающих функций распределения $F(x)\equiv G(x)$. Следовательно, таблицами точных и предельных (при больших объемах выборок) распределений статистик этих критериев и их процентных точек [1, 2] можно пользоваться при любых непрерывных функциях распределения результатов наблюдений.

Каким из непараметрических критериев пользоваться? Как известно из [3], для выбора одного из нескольких критериев необходимо сравнить их мощности, определяемые видом альтернативных гипотез. Сравнению мощностей критериев посвящена обширная литература.

Хорошо изучены свойства критериев при альтернативной гипотезе сдвига:

$$H_{1c} : G(x)=F(x-d),\ d\neq 0.$$

Критерии Вилкоксона, Ван-дер-Вардена и ряд других ориентированы для применения именно в этой ситуации. Если m раз измеряют характеристику одного объекта и n раз — другого, а функция распределения погрешностей измерения произвольна, но не меняется при переходе от объекта к объекту (это более жесткое требование, чем условие равенства дисперсий), то рассмотрение гипотезы H_{1c} оправдано. Однако в большинстве технических, экономических, медицинских и иных исследований нет оснований считать, что функции распределения, соответствующие выборкам, различаются только сдвигом.

5.3. Двухвыборочный критерий Вилкоксона

Покажем (и это — основной теоретический результат настоящего раздела), что двухвыборочный критерий Вилкоксона (в литературе его называют также критерием Манна-Уитни) предназначен для проверки гипотезы:

$$H_0 : P(X < Y) = 1/2,$$

где X — случайная величина, распределенная как элементы первой выборки, а Y — случайная величина, распределенная как элементы второй выборки.

В описанной выше вероятностной модели двух независимых выборок без ограничения общности можно считать, что объем первой из них не превосходит объема второй, $m \le n$, в противном случае выборки можно поменять местами. Обычно предполагается, что функции $F(x)$ и $G(x)$ непрерывны и строго возрастают. Из непрерывности этих функций следует, что с вероятностью 1 все $m + n$ результатов наблюдений различны. При рассмотрении реальных статистических данных иногда наблюдаются совпадения результатов наблюдений, но сам факт их наличия — свидетельство нарушений предпосылок только что описанной базовой математической модели.

Статистика S двухвыборочного критерия Вилкоксона определяется следующим образом. Все элементы объединенной выборки $X_1, X_2, ..., X_m, Y_1, Y_2, ..., Y_n$ упорядочиваются в порядке возрастания. Элементы первой выборки $X_1, X_2, ..., X_m$ занимают в общем вариационном ряду места с номерами $R_1, R_2, ..., R_m$, другими словами, имеют ранги $R_1, R_2, ..., R_m$. Тогда статистика Вилкоксона — это сумма рангов элементов первой выборки:

$$S = R_1 + R_2 + ... + R_m.$$

Статистика U Манна-Уитни определяется как число пар (X_i, Y_j) таких, что $X_i < Y_j$, среди всех mn пар, в которых первый элемент — из первой выборки, а второй — из второй. Как известно [7, с. 160],

$$U = mn + m(m+1)/2 - S.$$

Поскольку S и U линейно связаны, то часто говорят не о двух критериях — Вилкоксона и Манна — Уитни, а об одном — критерии Вилкоксона (Манна — Уитни).

Критерий Вилкоксона — один из самых известных инструментов непараметрической статистики (наряду с критериями на основе статистик типа Колмогорова-Смирнова, омега-квадрат и коэффициентами ранговой корреляции). Свойствам этого критерия и таблицам его критических значений уделяется место во многих монографиях по математической и прикладной статистике (см., например, [1, 2, 7]).

Однако в литературе имеются и неточные утверждения относительно возможностей критерия Вилкоксона. Так, одни полагают, что с его помощью можно обнаружить любое различие между функциями распределения $F(x)$

и $G(x)$. По мнению других, этот критерий нацелен на проверку равенства медиан распределений, соответствующих выборкам. И то, и другое, строго говоря, неверно. Это будет ясно из дальнейшего изложения.

Введем некоторые обозначения. Пусть $F^{-1}(t)$ — функция, обратная к функции распределения $F(x)$. Она определена на отрезке [0;1]. Положим, $L(t) = G(F^{-1}(t))$. Поскольку $F(x)$ непрерывна и строго возрастает, то $F^{-1}(t)$ и $L(t)$ обладают теми же свойствами. Важную роль в дальнейшем изложении будет играть величина $a = P(X < Y)$. Как нетрудно показать,

$$a = P(X < Y) = \int_0^1 t dL(t).$$

Введем также параметры:

$$b^2 = \int_0^1 L^2(t)dt - (1-a)^2, \quad g^2 = \int_0^1 t^2 dL(t) - a^2.$$

Тогда математические ожидания и дисперсии статистик Вилкоксона и Манна — Уитни согласно [7, с. 160] выражаются через введенные величины:

$$M(U) = mna\,,\ M(S) = mn + m(m+1)/2 - M(U) = mn(1-a) + m(m+1)/2,$$
$$D(S) = D(U) = mn\,[(n-1)\,b^2 + (m-1)\,g^2 + a(1-a)]. \qquad (1)$$

Когда объемы обеих выборок безгранично растут, распределения статистик Вилкоксона и Манна — Уитни являются асимптотически нормальными (см., например, [7, гл. 5 и 6]) с параметрами, задаваемыми формулами (1).

Если выборки полностью однородны, т.е. их функции распределения совпадают, справедлива гипотеза:

$$H_0\text{: } F(x) = G(x) \text{ } \textit{при всех } x, \qquad (2)$$

то $L(t) = t$ для t из отрезка [0, 1], $L(t)= 0$ для всех отрицательных t и $L(t)= 1$ для $t > 1$, соответственно $a= 1/2$. Подставляя в формулы (1), получаем, что:

$$M(S) = m(m + n + 1)/2, D(S) = mn(m + n + 1)/12. \qquad (3)$$

Следовательно, распределение нормированной и центрированной статистики Вилкоксона:

$$T = (S - m(m + n + 1)/2)\ (mn(m + n + 1)/12)^{-1/2} \qquad (4)$$

при росте объемов выборок приближается к стандартному нормальному распределению (с математическим ожиданием 0 и дисперсией 1).

Из асимптотической нормальности статистики T следует, что правило принятия решения для критерия Вилкоксона выглядит так:

- если $|T| \le \Phi^{-1}(1-\alpha/2)$, то гипотеза (2) однородности (тождества) функций распределений принимается на уровне значимости α,

- если же $|T| > \Phi^{-1}(1-\alpha/2)$, то гипотеза (2) однородности (тождества) функций распределений отклоняется на уровне значимости α.

В прикладной статистике наиболее часто применяется уровень значимости $\alpha = 0{,}05$. Тогда значение модуля статистики T Вилкоксона надо сравнивать с граничным значением $\Phi^{-1}(1-\alpha/2)=1{,}96$.

Пример 1. Пусть даны две выборки. Первая содержит $m = 12$ элементов 17; 22; 3; 5; 15; 2; 0; 7; 13; 97; 66; 14. Вторая содержит $n=14$ элементов 47; 30; 2; 15; 1; 21; 25; 7; 44; 29; 33; 11; 6; 15. Проведем проверку однородности функций распределения двух выборок с помощью критерия Вилкоксона.

Первым шагом является построение общего вариационного ряда для элементов двух выборок (табл. 1).

Таблица 1

Общий вариационный ряд для элементов двух выборок

Ранги	1	2	3,5	3,5	5	6	7	8,5	8,5	10	11	12	14
Элементы выборок	0	1	2	2	3	5	6	7	7	11	13	14	15
Номера выборок	1	2	1	2	1	1	2	1	2	2	1	1	1
Ранги	14	14	16	17	18	19	20	21	22	23	24	25	26
Элементы выборок	15	15	17	21	22	25	29	30	33	44	47	66	97
Номера выборок	2	2	1	2	1	2	2	2	2	2	2	1	1

Хотя с точки зрения теории математической статистики вероятность совпадения двух элементов выборок равна 0, в реальных выборках статистических данных совпадения встречаются. Так, в рассматриваемых выборках, как видно из табл. 1, два раза повторяется величина 2, два раза — величина 7 и три раза — величина 15. В таких случаях говорят о наличии «связанных рангов», а соответствующим совпадающим величинам приписывают среднее арифметическое тех рангов, которые они занимают. Так, величины 2 и 2 занимают в объединенной выборке места 3 и 4, поэтому им приписывается ранг (3 + 4)/2 = 3,5. Величины 7 и 7 занимают в объединенной выборке места 8 и 9, поэтому им приписывается ранг (8 + 9)/2 = 8,5. Величины 15, 15 и 15 занимают в объединенной выборке места 13, 14 и 15, поэтому им приписывается ранг (13 + 14 + 15)/3 = 14.

Следующий шаг — подсчет значения статистики Вилкоксона, т.е. суммы рангов элементов первой выборки:

$$S = R_1 + R_2 + ... + R_m = 1 + 3,5 + 5 + 6 + 8,5 + 11 + 12 + 14 + 16 + 18 + 25 + 26 = 146.$$

Подсчитаем также сумму рангов элементов второй выборки:

$$S_1 = 2 + 3,5 + 7 + 8,5 + 10 + 14 + 14 + 17 + 19 + 20 + 21 + 22 + 23 + 24 = 205.$$

Величина S_1 может быть использована для контроля вычислений. Дело в том, что суммы рангов элементов первой выборки S и второй выборки S_1 вместе составляют сумму рангов объединенной выборки, т.е. сумму всех натуральных чисел от 1 до $m+n$. Следовательно,

$$S + S_1 = (m + n)\,(m + n + 1)/2 = (12 + 14)\,(12 + 14 + 1)/2 = 351.$$

В соответствии с ранее проведенными расчетами $S + S_1 = 146 + 205 = 351$. Необходимое условие правильности расчетов выполнено. Ясно, что справедливость этого условия не гарантирует правильности расчетов.

Перейдем к расчету статистики T. Согласно формуле (3):

$$M(S) = 12\,(12 + 14 + 1)/\,2 = 162,\ D(S) = 12 \cdot 14(12 + 14 + 1)/12 = 378.$$

Следовательно,

$$T = (S - 162)\,(378)^{-1/2} = (146 - 162)/19,44 = -\,0.82.$$

Поскольку $|T|\leq 1,96$, то гипотеза однородности принимается на уровне значимости 0,05.

Что будет, если поменять выборки местами, вторую назвать первой? Тогда вместо S надо рассматривать S_1. Имеем:

$$M(S_1) = 14\,(12 + 14 + 1)/2 = 189,\ D(S) = D(S_1) = 378,$$
$$T_1 = (S_1 - 189)\,(378)^{-1/2} = (205 - 162)\,/\,19,44 = 0.82.$$

Таким образом, значения статистики критерия отличаются только знаком (можно показать, что это утверждение верно всегда). Поскольку в правиле принятия решения используется только абсолютная величина статистики, то принимаемое решение не зависит от того, какую выборку считаем первой, а какую второй. Для уменьшения объема таблиц критических значений принято считать первой выборку меньшего объема.

Продолжим обсуждение критерия Вилкоксона. Правила принятия решений и таблицы критических значений для критерия Вилкоксона строятся в предположении справедливости гипотезы полной однородности, описываемой формулой (2). А что будет, если эта гипотеза неверна? Другими словами, какова мощность критерия Вилкоксона?

Пусть объемы выборок достаточно велики, так что можно пользоваться асимптотической нормальностью статистики Вилкоксона. Тогда в соответствии с формулами (1) статистика T будет асимптотически нормальна с параметрами:

$$M(T) = (12mn)^{1/2}\,(1/2 - a)\,(m + n + 1)^{-1/2},$$
$$D(T) = 12\,[(n-1)\,b^2 + (m-1)\,g^2 + a(1-a)]\,(m + n + 1)^{-1}. \qquad (5)$$

Из формул (5) видно большое значение гипотезы:

$$H_{01}\colon a = P(X < Y) = 1/2. \qquad (6)$$

Если эта гипотеза неверна, то, поскольку $m \leq n$, справедлива оценка

$$|M(T)| \geq (12m\,n\,(2n + 1)^{-1})^{1/2}\,|1/2 - a|,$$

а потому $|M(T)|$ безгранично растет при росте объемов выборок. В то же время, поскольку

$$b^2 \leq \int_0^1 L^2(t)dt \leq 1,\quad g^2 \leq \int_0^1 t^2 dL(t) \leq 1,\quad \alpha(1-\alpha) \leq 1/4,$$

то

$$D(T) \leq 12\,[(n-1) + (m-1) + 1/4]\,(m + n + 1)^{-1} \leq 12. \qquad (7)$$

Следовательно, вероятность отклонения гипотезы H_{01}, когда она неверна, т.е. мощность критерия Вилкоксона как критерия проверки гипотезы (6), стремится к 1 при возрастании объемов выборок, т.е. критерий Вилкоксона является состоятельным для этой гипотезы при альтернативе:

$$AH_{01}: a = P(X < Y) \neq 1/2. \tag{8}$$

Если же гипотеза (6) верна, то статистика T асимптотически нормальна с математическим ожиданием 0 и дисперсией, определяемой формулой:

$$D(T) = 12[(n-1)b^2 + (m-1)g^2 + 1/4](m+n+1)^{-1}. \tag{9}$$

Гипотеза (6) является сложной, дисперсия (9), как показывают приводимые ниже примеры, в зависимости от значений b^2 и g^2 может быть как больше 1, так и меньше 1, но согласно неравенству (7) никогда не превосходит 12.

Приведем пример двух функций распределения $F(x)$ и $G(x)$ таких, что гипотеза (6) выполнена, а гипотеза (2) — нет. Поскольку

$$a = P(X < Y) = \int_{-\infty}^{+\infty} F(x)dG(x),\ 1 - a = P(Y < X) = \int_{-\infty}^{+\infty} G(x)dF(x) \tag{10}$$

и $a = 1/2$ в случае справедливости гипотезы (2), то для выполнения условия (6) необходимо и достаточно, чтобы

$$\int_{-\infty}^{+\infty} (F(x) - G(x))dF(x) = 0, \tag{11}$$

а потому естественно в качестве $F(x)$ рассмотреть функцию равномерного распределения на интервале (– 1; 1). Тогда формула (11) переходит в условие:

$$\int_{-\infty}^{+\infty} (F(x) - G(x))dF(x) = -\frac{1}{2}\int_{-1}^{+1}\left(G(x) - \frac{(x+1)}{2}\right)dx = 0. \tag{11}$$

Это условие выполняется, если функция $(G(x) - (x+1)/2)$ является нечетной.

Пример 2. Пусть функции распределения $F(x)$ и $G(x)$ сосредоточены на интервале (– 1; 1), на котором

$$F(x) = (x+1)/2,\ G(x) = (x + 1 + 1/\pi\ \sin\pi x)/2.$$

Тогда

$X = F^{-1}(t) = 2t - 1$, $L(t) = G(F^{-1}(t)) = (2t + 1/\pi \sin\pi(2t - 1))/2 = t + 1/2\pi \sin\pi(2t - 1)$.

Условие (11) выполнено, поскольку функция $(G(x) - (x + 1)/2)$ является нечетной. Следовательно, $a = 1/2$. Начнем с вычисления:

$$g^2 = \int_0^1 t^2 dL(t) - 1/4 = \int_0^1 t^2 d(t + \frac{1}{2\pi}\sin\pi(2t-1)) - \frac{1}{4}.$$

Поскольку

$$d(t + \frac{1}{2\pi}\sin\pi(2t-1)) = (1 + \cos\pi(2t-1))dt,$$

то

$$g^2 = \int_0^1 t^2(1 + \cos\pi(2t-1))dt - \frac{1}{4} = \frac{1}{12} + \int_0^1 t^2 \cos\pi(2t-1)dt.$$

С помощью замены переменных $t = (x + 1) / 2$ получаем:

$$\int_0^1 t^2 \cos\pi(2t-1)dt = \frac{1}{8}\left(\int_{-1}^1 x^2 \cos\pi x dx + 2\int_{-1}^1 x\cos\pi x dx + \int_{-1}^1 \cos\pi x dx\right).$$

В правой части последнего равенства стоят табличные интегралы (см., например, справочник [8, с. 71]). Проведя соответствующие вычисления, получаем, что в правой части стоит $1/8 \cdot (-4/\pi^2) = -1/(2\pi^2)$. Следовательно,

$$g^2 = 1/12 - 1/(2\pi^2) = 0{,}032672733...$$

Перейдем к вычислению b^2. Поскольку

$$b^2 = \int_0^1 L^2(t)dt - \frac{1}{4} = \int_0^1 \left(t + \frac{1}{2}\pi\sin\pi(2t-1)\right)^2 dt - \frac{1}{4},$$

то

$$b^2 = \frac{1}{12} + \frac{1}{\pi}\int_0^1 (t \sin \pi(2t-1))dt + \left(\frac{\pi}{2}\right)^2 \int_0^1 \sin^2 \pi(2t-1)dt.$$

С помощью замены переменных $t = (x+1)/2$ переходим к табличным интегралам (см., например, справочник [8, с. 65]):

$$b^2 = \frac{1}{12} + \frac{1}{4\pi}\int_{-1}^1 x \sin \pi x dx + \frac{1}{4\pi}\int_{-1}^1 \sin \pi x dx + \frac{1}{8\pi^2}\int_{-1}^1 \sin^2 \pi x dx.$$

Проведя необходимые вычисления, получим:

$$b^2 = \frac{1}{12} + \frac{1}{4\pi}\left(-\frac{2}{\pi}\right) + 0 + \frac{1}{8\pi^2} = \frac{1}{12} - \frac{3}{8\pi^2} = 0{,}045337893\ldots$$

Следовательно, для рассматриваемых функций распределения нормированная и центрированная статистика Вилкоксона (см. формулу (4)) асимптотически нормальна с математическим ожиданием 0 и дисперсией (см. формулу (9)):

$$D(T) = (0{,}544\, n + 0{,}392\, m + 2{,}064)\,(m+n+1)^{-1}.$$

Как легко видеть, дисперсия всегда меньше 1. Это значит, что в рассматриваемом случае гипотеза полной однородности (2) при проверке с помощью критерия Вилкоксона будет приниматься чаще, чем если она на самом деле верна.

На наш взгляд, это означает, что критерий Вилкоксона нельзя считать критерием для проверки гипотезы (2) при альтернативе общего вида. Он не всегда позволяет проверить однородность — не при всех альтернативах. Точно так же критерии типа хи-квадрат нельзя считать критериями проверки гипотез согласия и однородности — они позволяют обнаружить не все различия, поскольку некоторые из них «скрадывает» группировка.

Обсудим теперь, действительно ли критерий Вилкоксона нацелен на проверку равенства медиан распределений, соответствующих выборкам.

Пример 3. Построим семейство пар функций распределения $F(x)$ и $G(x)$ таких, что их медианы различны, но для $F(x)$ и $G(x)$ выполнена гипотеза (6).

Пусть распределения сосредоточены на интервале (0; 1), и на нем $G(x) = x$, а $F(x)$ имеет кусочно-линейный график с вершинами в точках (0; 0), (λ, 1/2), (δ, 3/4), (1; 1). Следовательно,

$F(x) = 0$ при $x < 0$;
$F(x) = x/(2\lambda)$ на $[0;\ \lambda)$;
$F(x) = 1/2 + (x - \lambda)/(4\delta - 4\lambda)$ на $[\lambda;\ \delta)$;
$F(x) = 3/4 + (x - \delta)/(4 - 4\delta)$ на $[\delta;\ 1]$;
$F(x) = 1$ при $x > 1$.

Очевидно, что медиана $F(x)$ равна λ, а медиана $G(x)$ равна 1/2.

Согласно соотношению (9) для выполнения гипотезы (6) достаточно определить δ как функцию λ, т.е. $\delta = \delta(\lambda)$, из условия:

$$\int_0^1 F(x)dx = \frac{1}{2}.$$

Вычисления дают

$$\delta = \delta(\lambda) = 3(1 - \lambda)/2.$$

Учитывая, что δ лежит между λ и 1, не совпадая ни с тем, ни с другим, получаем ограничения на λ, а именно, $1/3 < \lambda < 3/5$. Итак, построено искомое семейство пар функций распределения.

Пример 4. Пусть, как и в примере 3, распределения сосредоточены на интервале (0; 1), и на нем $F(x) = x$. А $G(x)$ — функция распределения, сосредоточенного в двух точках — β и 1. Т.е. $G(x) = 0$ при x, не превосходящем β; $G(x) = h$ на $(\beta;\ 1]$; $G(x) = 1$ при $x > 1$. С такой функцией $G(x)$ легко проводить расчеты. Однако она не удовлетворяет принятым выше условиям непрерывности и строгого возрастания. Вместе с тем легко видеть, что она является предельной (сходимость в каждой точке отрезка [0; 1]) для последовательности функций распределения, удовлетворяющих этим условиям. А распределение статистики Вилкоксона для пары функций распределения примера 4 является предельным для последовательности соответствующих распределений статистики Вилкоксона, полученных в рассматриваемых условиях непрерывности и строгого возрастания.

Условие $P(X < Y) = 1/2$ выполнено, если $h = (1 - \beta)^{-1}/2$ (при β из отрезка [0; 1/2]). Поскольку $h > 1/2$ при положительном β, то очевидно, что медиана

$G(x)$ равна β, в то время как медиана $F(x)$ равна 1/2 . Значит, при $\beta = 1/2$ медианы совпадают, при всех иных положительных β — различны. При $\beta = 0$ медианой $G(x)$ является любая точка из отрезка [0; 1].

Легко подсчитать, что в условиях примера 4 параметры предельного распределения имеют вид:

$$b^2 = \beta(1 - \beta)^{-1}/4 \, , \; g^2 = (1 - 2\beta)/4.$$

Следовательно, распределение нормированной и центрированной статистики Вилкоксона будет асимптотически нормальным с математическим ожиданием 0 и дисперсией:

$$D(T) = 3\,[(n - 1)\,\beta(1 - \beta)^{-1} + (m - 1)\,(1 - 2\beta) + 1]\,(m + n + 1)^{-1}.$$

Проанализируем величину $D(T)$ в зависимости от параметра β и объемов выборок m и n. При достаточно больших m и n:

$$D(T) = 3w\beta(1 - \beta)^{-1} + 3(1 - w)\,(1 - 2\beta)$$

с точностью до величин порядка $(m+n)^{-1}$, где $w= n/(m+n)$. Значит, $D(T)$ — линейная функция от w, а потому достигает экстремальных значений на границах интервала изменения w, т.е. при $w = 0$ и $w = 1$. Легко видеть, что при $\beta(1 - \beta)^{-1} < 1 - 2\beta$ минимум равен $3\beta(1 - \beta)^{-1}$ (при $w = 1$), а максимум равен $3(1 - 2\beta)$ (при $w = 0$). В случае $\beta(1 - \beta)^{-1} > 1 - 2\beta$ максимум равен $3\beta(1 - \beta)^{-1}$ (при $w = 1$), а минимум равен $3(1 - 2\beta)$ (при $w = 0$). Если же $\beta(1 - \beta)^{-1} = 1 - 2\beta$ (это равенство справедливо при $\beta = \beta_0 = 1 - 2^{-1/2} = 0{,}293$), то $D(T) = 3(2^{1/2} - 1) = 1{,}2426...$ при всех w из отрезка [0; 1].

Первый из описанных выше случаев имеет быть при $\beta < \beta_0$. При этом минимум $D(T)$ возрастает от 0 (при β=0, w=1 — предельный случай) до $3(2^{1/2} - 1)$ (при $\beta = \beta_0$, w — любом), а максимум уменьшается от 3 (при β=0, w=0 — предельный случай) до $3\,(2^{1/2} - 1)$ (при $\beta = \beta_0$, w — любом). Второй случай относится к β из интервала (β_0; 1/2]. При этом минимум убывает от приведенного выше значения для $\beta = \beta_0$ до 0 (при β=1/2, w=0 — предельный случай), а максимум возрастает от того же значения при $\beta = \beta_0$ до 3 (при β=1/2, w=0).

Таким образом, $D(T)$ может принимать все значения из интервала (0; 3) в зависимости от значений β и w. Если $D(T) < 1$, то при применении критерия Вилкоксона к выборкам с рассматриваемыми функциями распределения гипотеза однородности (2) будет приниматься чаще (при соответствующих значениях β и w — с вероятностью, сколь угодно близкой к 1), чем если бы она самом деле была верна. Если $1 < D(T) < 3$, то гипотеза (2) также принимается достаточно часто. Так, если уровень значимости критерия Вилкоксона равен 0,05, то (асимптотическая) критическая область этого критерия, как показано выше, имеет вид $\{T\colon |T| \geq 1{,}96\}$. Если — самый плохой случай — $D(T)=3$, то гипотеза (2) принимается с вероятностью 0,7422.

Гипотеза сдвига. При проверке гипотезы однородности мы рассмотрели различные виды нулевых и альтернативных гипотез — гипотезу (2) и ее отрицание в качестве альтернативы, гипотезу (6) и ее отрицание, гипотезы о равенстве или различии медиан. В теоретических работах по математической статистике часто рассматривают гипотезу сдвига, в которой альтернативой гипотезе (2) является гипотеза:

$$H_1\colon F(x) = G(x + r) \tag{12}$$

при всех x и некотором сдвиге r, отличным от 0. Если верна альтернативная гипотеза H_1, то вероятность $P(X < Y)$ отлична от 1/2, а потому при альтернативе (12) критерий Вилкоксона является состоятельным.

В некоторых прикладных постановках гипотеза (12) представляется естественной. Например, если одним и тем же прибором проводятся две серии измерений двух значений некоторой величины (физической, химической и т.п.). При этом функция распределения $G(x)$ описывает погрешности измерения одного значения, а $G(x + r)$ — другого. Вопреки распространенному заблуждению, хорошо известно, что распределение погрешностей измерений, как правило, не является нормальным (см. об этом главу 2.1). Однако при анализе конкретных статистических данных, как правило, нет никаких оснований считать, что отсутствие однородности всегда выражается столь однозначным образом, как следует из формулы (12). Поэтому для проверки однородности необходимо использовать статистические критерии, состоятельные против любого отклонения от гипотезы однородности (2).

Почему же математики так любят гипотезу сдвига (12)? Да потому, что она дает возможность доказывать глубокие математические результаты, например, об асимптотической оптимальности критериев. К сожалению, с точки зрения статистики это напоминает поиск ключей под фонарем, где светло, а не в кустах, где они потеряны.

Отметим еще одно обстоятельство. Часто говорят (в соответствии с классическим подходом математической статистики), что нельзя проверять нулевые гипотезы без рассмотрения альтернативных. Однако при анализе данных, полученных в ходе технических, экономических, медицинских или иных исследований, зачастую полностью ясна формулировка той гипотезы, которую желательно проверить (например, гипотезы полной однородности — см. формулу (2)), в то время как формулировка альтернативной гипотезы не очевидна. То ли это гипотеза о неверности равенства (2) хотя бы для одного значения x, то ли это альтернатива (8), то ли — альтернатива сдвига (12), и т. д. В таких случаях целесообразно «обернуть» задачу — исходя из статистического критерия найти альтернативы, относительно которых он состоятелен. Именно это и проделано в настоящем разделе для критерия Вилкоксона.

Подведем итоги рассмотрения критерия Вилкоксона.

1. Критерий Вилкоксона (Манна — Уитни) является одним из самых распространенных непараметрических ранговых критериев, используемых для проверки однородности двух выборок. Его значение не меняется при любом монотонном преобразовании шкалы измерения (т.е. он пригоден для статистического анализа данных, измеренных в порядковой шкале).

2. Распределение статистики критерия Вилкоксона определяется функциями распределения $F(x)$ и $G(x)$ и объемами m и n двух выборок. При больших объемах выборок распределение статистики Вилкоксона является асимптотически нормальным с параметрами, выписанными выше (см. формулы (1), (3) и (5)).

3. При альтернативной гипотезе, когда функции распределения выборок $F(x)$ и $G(x)$ не совпадают, распределение статистики Вилкоксона зависит от величины $a = P(X < Y)$. Если a отличается от 1/2, то мощность критерия Вилкоксона стремится к 1, и он отличает нулевую гипотезу $F = G$ от альтернативной. Если же $a = 1/2$, то это не всегда имеет место. В примере 2 приведены две *различные* функции распределения выборок $F(x)$ и $G(x)$ такие, что гипотеза однородности $F = G$ при проверке с помощью критерия Вилкоксона будет приниматься *чаще*, чем если бы она на самом деле была верна.

4. Следовательно, в случае общей альтернативы критерий Вилкоксона не является состоятельным, т.е. не всегда позволяет обнаружить различие функций распределения. Однако это не лишает его практической ценности, точно так же, как несостоятельность критериев типа хи-квадрат при проверке согла-

сия, независимости или однородности не мешает отклонять нулевую гипотезу во многих практически важных случаях. Однако принятие нулевой гипотезы с помощью критерия Вилкоксона может означать не совпадение F и G, а лишь выполнение равенства $a = 1/2$.

5. Иногда утверждают, что с помощью критерия Вилкоксона можно проверять равенство медиан функций распределения F и G. Это не так. В примерах 3 и 4 указаны функции распределения F и G с $a = 1/2$, но с различными медианами. Во многих случаях это различие нельзя обнаружить с помощью критерия Вилкоксона, как это показано при численном анализе асимптотической дисперсии в примере 4.

6. Указанные выше недостатки критерия Вилкоксона исчезают для специального вида альтернативы — «альтернативы сдвига» H_1: $F(x) = G(x + r)$. В этом частном случае при справедливости альтернативной гипотезы мощность стремится к 1, различие медиан также всегда обнаруживается. Однако альтернатива сдвига не всегда естественна. Ее целесообразно принять, если одним и тем же прибором проводятся две серии измерений двух значений некоторой величины (физической, химической и т.п.). При этом функция распределения $G(x)$ описывает результаты измерений (с погрешностями) одного значения, а $F(x) = G(x + r)$ — другого. Другими словами, меняется лишь измеряемое значение, а собственно распределение погрешностей — одно и то же, присущее используемому средству измерения (и обычно описанное в его техническом паспорте). Однако в большинстве прикладных статистических исследований нет никаких оснований считать, что при альтернативе функция распределения второй выборки лишь сдвигается, но не меняется каким-либо иным образом.

7. При всех своих недостатках критерий Вилкоксона прост в применении и часто позволяет обнаруживать различие групп (поскольку оно часто сводится к отличию $a = P(X < Y)$ от 1/2). Приведенные здесь критические замечания не следует понимать как призыв к полному отказу от использования критерия Вилкоксона. Однако для проверки гипотезы однородности в случае альтернативы общего вида можно порекомендовать состоятельные критерии, в частности, рассматриваемые в следующем разделе критерии Смирнова и типа омега-квадрат (Лемана — Розенблатта).

8. В литературе по прикладным статистическим методам соседствуют два стиля изложения. Один из них исходит из формулировок нулевой и альтернативных гипотез (или описания набора гипотез, из которого надо выбрать наиболее адекватную), для проверки которых строятся те или иные критерии. При другом стиле изложения упор делается на алгоритмическое описание кри-

териев для проверки тех или иных гипотез, а об альтернативах даже не упоминается.

Например, в литературе по математической статистике часто говорится, что для проверки нормальности используются критерии асимметрии и эксцесса (они описаны, например, в главе 2.3 и в лучшем справочнике 1960–1980-х гг. в [1, табл. 4.7]). Однако эти критерии позволяют проверять некоторые соотношения между моментами распределения, но отнюдь не являются состоятельными критериями нормальности (не все отклонения от нормальности обнаруживают). Впрочем, для прикладной статистики эти критерии большого практического значения не имеют, поскольку заранее известно, что распределения конкретных технических, экономических, медицинских и иных статистических данных скорее всего отличны от нормальных.

Так что недостатки критерия Вилкоксона не является исключением, мощность ряда иных популярных в математической статистике критериев заслуживает тщательного изучения, при этом заранее можно сказать, что зачастую они не позволяют проверять те гипотезы, с которыми традиционно связаны. При применении подобных критериев к анализу реальных данных необходимо тщательно взвешивать их достоинства и недостатки.

5.4. Состоятельные критерии проверки однородности независимых выборок

В соответствии с методологией статистических методов естественно потребовать, чтобы рекомендуемый для массового использования в технических, экономических, медицинских и иных исследованиях критерий однородности был состоятельным. Напомним: это значит, что для любых отличных друг от друга функций распределения $F(x)$ и $G(x)$ (другими словами, при справедливости альтернативной гипотезы H_1) вероятность отклонения гипотезы H_0 должна стремиться к 1 при увеличении объемов выборок m и n. Из перечисленных выше (в конце раздела 5.2) критериев однородности состоятельными являются только критерии Смирнова и типа омега-квадрат.

Проведенное в Институте высоких статистических технологий и эконометрики исследование мощности (методом статистических испытаний) первых четырех из перечисленных выше критериев (при различных вариантах функций распределения $F(x)$ и $G(x)$) подтвердило преимущество критериев Смирнова и омега-квадрат и при объемах выборок 6–12. Рассмотрим эти критерии подробнее.

Критерий Смирнова однородности двух независимых выборок. Он был предложен членом-корреспондентом АН СССР Н.В. Смирновым в 1939 г. (см. справочник [1]). Единственное ограничение — функции распределения $F(x)$ и $G(x)$ должны быть непрерывными. Напомним, что согласно Л.Н. Большеву и Н.В. Смирнову [1] значение эмпирической функции распределения в точке x равно доле результатов наблюдений в выборке, меньших x. Критерий Смирнова основан на использовании эмпирических функций распределения $F_m(x)$ и $G_n(x)$, построенных по первой и второй выборкам соответственно. Значение статистики Смирнова:

$$D_{m,n} = \sup_x | F_m(x) - G_n(x) | ,$$

сравнивают с соответствующим критическим значением (см., например, [1]) и по результатам сравнения принимают или отклоняют гипотезу H_0 о совпадении (однородности) функций распределения. Практически значение статистики $D_{m,n}$ рекомендуется согласно монографии [1] вычислять по формулам:

$$D^+_{m,n} = \max_{1 \le r \le n}\left[\frac{r}{n} - F_m(y'_r)\right] = \max_{1 \le r \le m}\left[G_n(x'_s) - \frac{s-1}{m}\right],$$

$$D^-_{m,n} = \max_{1 \le r \le n}\left[F_m(y'_r) - \frac{r-1}{n}\right] = \max_{1 \le r \le m}\left[\frac{s}{m} - G_n(x'_s)\right],$$

$$D_{m,n} = \max(D^+_{m,n}, D^-_{m,n}),$$

где $x'_1 < x'_2 < \ldots < x'_m$ — элементы первой выборки $x_1, x_2, \ldots, x_m$, переставленные в порядке возрастания, а $y'_1 < y'_2 < \ldots < y'_n$ — элементы второй выборки $y_1, y_2, \ldots, y_n$, также переставленные в порядке возрастания. Поскольку функции распределения $F(x)$ и $G(x)$ предполагаются непрерывными, то вероятность совпадения каких-либо выборочных значений равна 0.

Разработаны алгоритмы и компьютерные программы, позволяющие рассчитывать точные распределения, процентные точки и достигаемый уровень значимости для двухвыборочной статистики Смирнова $D_{m,n}$, разработаны подробные таблицы (см., например, методику [9], содержащую описание алгоритмов, тексты программ и подробные таблицы).

Однако у критерия Смирнова есть и недостатки. Его распределение сосредоточено в сравнительно небольшом числе точек, поэтому функция распределения растет большими скачками. В результате не удается выдержать заданный уровень значимости. Реальный уровень значимости может в несколько раз отличаться от номинального (подробному обсуждению неклассического феномена существенного отличия реального уровня значимости от номинального посвящена работа [10] и раздел 5.7).

Критерий типа омега-квадрат (Лемана — Розенблатта). Статистика критерия типа омега-квадрат для проверки однородности двух независимых выборок имеет вид:

$$A = \frac{mn}{m+n}\int_{-\infty}^{\infty}(F_m(x) - G_n(x))^2\, dH_{m+n}(x)\,,$$

где $H_{m+n}(x)$ — эмпирическая функция распределения, построенная по объединенной выборке. Легко видеть, что:

$$H_{m+n}(x) = \frac{m}{m+n}F_m(x) + \frac{n}{m+n}G_n(x).$$

Статистика A типа омега-квадрат была предложена Э. Леманом в 1951 г., изучена М. Розенблаттом в 1952 г., а затем и другими исследователями. Она зависит лишь от рангов элементов двух выборок в объединенной выборке. Пусть $x_1, x_2,...,x_m$ — первая выборка, $x_1' < x_2' <..x_m'$ — соответствующий вариационный ряд, $y_1, y_2,...,y_n$ — вторая выборка, $y_1' < y_2' <...< y_n'$ — вариационный ряд, соответствующий второй выборке. Поскольку функции распределения независимых выборок непрерывны, то с вероятностью 1 все выборочные значения различны, совпадения отсутствуют. Статистика A представляется в виде (см., например, [1]):

$$A = \frac{1}{mn(m+n)}[m\sum_{i=1}^{m}(r_i - i)^2 + n\sum_{j=1}^{n}(s_j - j)^2] - \frac{4mn-1}{6(m+n)},$$

где r_i — ранг x'_i и s_j — ранг y'_j в общем вариационном ряду, построенном по объединенной выборке.

Правила принятия решений при проверке однородности двух выборок на основе статистик Смирнова и типа омега-квадрат, т.е. таблицы критических значений в зависимости от уровней значимости и объемов значимости приведены, например, в таблицах [1].

Рекомендации по выбору критерия однородности. Для критерия типа омега-квадрат нет выраженного эффекта различия между номинальными и реальными уровнями значимости. Поэтому мы *рекомендуем для проверки однородности функций распределения (гипотеза H_0) применять статистику А типа омега-квадрат. Если методическое, табличное или программное обеспечение для статистики Лемана — Розенблатта отсутствует, рекомендуем использовать критерий Смирнова. Для проверки однородности математических ожиданий (гипотеза H'_0) целесообразно применять критерий Крамера-Уэлча.* По нашему мнению, статистики Стьюдента, Вилкоксона и др. допустимо использовать лишь в отдельных частных случаях, рассмотренных выше.

Некоторые соображения о внедрении современных методов прикладной статистики в практику технических, экономических, медицинских и иных исследований. Даже из проведенного выше разбора лишь одной из типичных статистических задач — задачи проверки однородности двух независимых выборок — можно сделать вывод о целесообразности широкого развертывания работ по критическому анализу сложившейся практики статистической обработки данных и по внедрению накопленного арсенала современных методов прикладной статистики. По нашему мнению, широкого внедрения заслуживают, в частности, методы многомерного статистического анализа, планирования эксперимента, статистики объектов нечисловой природы. Очевидно, рассматриваемые работы должны быть плановыми, организационно оформленными, проводиться мощными самостоятельными организациями и подразделениями. Целесообразно создание службы статистических консультаций в системе научно-исследовательских учреждений и вузов технического, экономического, медицинского профиля.

5.5. Методы проверки однородности связанных выборок

Начнем с практического примера. Приведем письмо главного инженера подмосковного химического комбината (некоторые названия изменены).

«Директору Института высоких статистических технологий
и эконометрики (Фамилия, имя, отчество)

Наш комбинат выпускает мастику по ГОСТ (следует номер) и является разработчиком указанного стандарта.

В результате исследовательских работ по подбору стандартного метода определения вязкости мастики на комбинате накоплен большой опыт сравнительных данных определения вязкости по двум методам:

- неразбавленной мастики — на нестандартном приборе фабрики им. А.А. Петрова;
- раствора мастики — на стандартном вискозиметре ВЗ–4.

Учитывая высокую компетентность сотрудников Вашего института, прошу Вас, в порядке оказания технической помощи нашему предприятию, поручить соответствующей лаборатории провести обработку представленных данных современными статистическими методами и выдать заключение о наличии (или отсутствии) зависимости между указанными выше методами определения вязкости мастики. Ваше заключение необходимо для решения спорного вопроса о целесообразности вновь ввести в ГОСТ (следует номер) метода определения вязкости мастики по вискозиметру ВЗ–4, который, по мнению некоторых потребителей, был необоснованно исключен из этого ГОСТ по изменению № 1.

Заранее благодарю Вас за оказанную помощь.

Приложение: статистика на 3 листах.

Главный инженер (Подпись) (Фамилия, имя, отчество)»

Комментарий. Вязкость мастики — один из показателей качества мастики. Измерять этот показатель можно по-разному. И, как оказалось, разные способы измерения дают разные результаты. Ничего необычного в этом нет. Однако поставщику и потребителю следует согласовать способы измерения показателей качества. Иначе достаточно часто поставщик (производитель) будет утверждать, что он выполнил условия контракта, а потребитель заявлять, что нет. Такая конфликтная ситуация иногда называется арбитражной, поскольку для ее решения стороны могут обращаться в арбитражный суд. Простейший метод согласования способов измерения показателей состоит в том, чтобы выбрать один из них и внести в государственный стандарт, который тем самым будет содержать не только описание продукции, перечень ее показателей качества и требований к ним, но и способы измерения этих показателей.

Заключение по статистическим данным, представленным химическим комбинатом. Для каждой из 213 партий мастики представлены два числа — результат измерения вязкости на нестандартном приборе фабрики им. А.А. Петрова и результат измерения вязкости на стандартном вискозиметре ВЗ-4. Требуется установить, дают ли два указанных метода сходные результаты. Если они дают сходные результаты, то нет необходимости вводить в соответствующий ГОСТ требование об обязательном использовании определенного метода определения вязкости. Если же методы дают существенно различные результаты, то подобное требование ввести необходимо.

Для применения статистических методов в рассматриваемой задаче необходимо описать вероятностную модель. Считаем, что статистические данные имеют вид $(x_i, y_i), i = 1,2,\ldots,213,$ где x_i-результат измерения на нестандартном приборе фабрики им. А.А. Петрова в i-й партии, а y_i — результат измерения вязкости на стандартном вискозиметре ВЗ–4 в той же i-й партии. Пусть a_i — истинное значение показателя качества в i-ой партии. Естественно считать, что указанные выше случайные вектора независимы в совокупности. При этом они не являются одинаково распределенными, поскольку отличаются истинными значениями показателей качества a_i.

Принимаем, что *при каждом i случайные величины* $x_i - a_i$ *и* $y_i - a_i$ *независимы и одинаково распределены*. Это условие и означает *однородность в связанных выборках.* Параметры связи — величины a_i . Их наличие не позволяет объединить первые координаты в одну выборку, вторую — во вторую, как делалось в случае проверки однородности двух независимых выборок.

В предположении непрерывности функций распределения из условия однородности в связанных выборках вытекает, что:

$$P(x_i < y_i) = P(x_i \geq y_i) = \frac{1}{2}.$$

Рассмотрим случайные величины $Z_i = x_i - y_i, \quad i = 1,2,...,213$. Из последнего соотношения вытекает, что при справедливости гипотезы однородности для связанных выборок эти случайные величины имеют нулевые медианы. Другими словами, проверка того, что метода измерения вязкости дают схожие результаты, эквивалентна проверке равенства 0 медиан величин Z_i.

Для проверки гипотезы о том, что медианы величин Z_i нулевые, применим широко известный критерий знаков (см., например, справочник [1, с. 89–91]). Согласно этому критерию, необходимо подсчитать, в скольких партиях $x_i < y_i$ и в скольких $x_i \geq y_i$. Для представленных химическим комбинатом данных $x_i < y_i$ в 187 случаях из 213 и $x_i \geq y_i$ в 26 случаях из 213.

Если рассматриваемая гипотеза верна, то число W осуществлений события $\{x_i < y_i\}$ имеет биномиальное распределение с параметрами p = ½ и n = 213. Математическое ожидание $M(W)$ = 106,5, а среднее квадратическое отклонение $\sigma = \sqrt{np(1-p)} = 7{,}3$. Следовательно, интервал $M(W) \pm 3\sigma$ — это интервал $84 \leq W \leq 129$. Найденное по данным химического комбината значение $W = 187$ лежит далеко вне этого интервала. Поэтому рассматриваемую гипотезу необходимо отвергнуть (на любом используемом в прикладных работах уровне значимости, в частности, на уровне значимости 1 %).

Таким образом, статистический анализ показывает, что два метода дают существенно различные результаты — по прибору фабрики им. А.А. Петрова результаты измерений, как правило, меньше, чем по вискозиметру ВЗ-4. Это означает, что в соответствующий ГОСТ целесообразно ввести указание на метод определения вязкости.

Система вероятностных моделей при проверке гипотезы однородности связанных выборок. Как и в случае проверки однородности независимых выборок, система вероятностных моделей состоит из трех уровней. Наиболее простая модель — на уровне однородности альтернативного признака — уже рассмотрена. Она сводится к проверке гипотезы о значении параметра биномиального распределения:

$$H_0 : p = \frac{1}{2}.$$

Речь идет о «критерии знаков». При справедливости гипотезы однородности число W осуществлений события $\{x_i < y_i\}$ имеет биномиальное распределение с вероятностью успеха p = 1/2 и числом испытаний n. Альтернативная гипотеза состоит в том, что вероятность успеха отличается от 1/2:

$$H_1 : p \neq \frac{1}{2}.$$

Гипотезу p = 1/2 можно проверять как непосредственно с помощью биномиального распределения (используя таблицы или программное обеспечение), так и опираясь на теорему Муавра-Лапласа. Согласно этой теореме,

$$\lim_{n\to\infty} P\{\frac{2W - n}{\sqrt{n}} \le x\} = \Phi(x)$$

при всех x, где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

Из теоремы Муавра-Лапласа вытекает правило принятия решений на уровне значимости 5 %: если:

$$\left|\frac{2W - n}{\sqrt{n}}\right| \le 1{,}96,$$

то гипотезу однородности связанных выборок принимают, в противном случае отклоняют.

Как обычно, при желании использовать другой уровень значимости применяют в качестве критического значения иную квантиль нормального распределения. Использование предельных теорем допустимо при достаточно больших объемах выборки. По поводу придания точного смысла термину «достаточно большой» продолжаются дискуссии. Обычно считается, что несколько десятков (два-три десятка) — это уже «достаточно много». Более правильно сказать, что ответ зависит от задачи, от ее сложности и практической значимости.

Второй уровень моделей проверки однородности связанных выборок — это уровень проверки однородности характеристик, прежде всего однородности математических ожиданий. Исходные данные — количественные результаты измерений (наблюдений, испытаний, анализов, опытов) двух признаков x_j и y_j, $j = 1,2, \ldots, n$, а непосредственно анализируются их разности $Z_j = x_j - y_j$, $j = 1,2, \ldots, n$. Предполагается, что эти разности независимы в совокупности и одинаково распределены, однако функция распределения неизвестна статистику. Необходимо проверить непараметрическую гипотезу:

$$H_{01} : M(Z_j) = 0.$$

Альтернативная гипотеза также является непараметрической и имеет вид:

$$H_{11} : M(Z_j) \neq 0.$$

Как и в случае проверки гипотезы согласованности для независимых выборок с помощью критерия Крамера — Уэлча, в рассматриваемой ситуации используем статистику:

$$Q = \sqrt{n}\frac{\overline{Z}}{s(Z)},$$

где

$$\overline{Z} = \frac{Z_1 + Z_2 + \ldots + Z_n}{n} -$$

среднее арифметическое разностей, а

$$s(Z) = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(Z_j - \overline{Z})^2} -$$

выборочное среднее квадратическое отклонение. Из Центральной Предельной Теоремы теории вероятностей и теорем о наследовании сходимости, полученных в монографии [4] и описанных в главе 14, вытекает, что:

$$\lim_{n\to\infty} P\{Q \le x\} = \Phi(x)$$

при всех x, где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Отсюда вытекает правило принятия решений на уровне значимости 5 %: если:

$$|Q| \le 1{,}96,$$

то гипотезу однородности математических ожиданий связанных выборок принимают, в противном случае отклоняют. Как обычно, при желании использовать другой уровень значимости применяют в качестве критического значения иную квантиль нормального распределения. Повторим, что использование предельных теорем допустимо при достаточно больших объемах выборки.

Третий уровень моделей проверки однородности связанных выборок — это уровень проверки однородности (совпадения) функций распределения. Необходимо проверить непараметрическую гипотезу наиболее общего вида:

$$H_{03}: F(x) = G(x), x \in R^1,$$

где

$$F(x) = P(x_i \le x), G(x) = P(y_i \le x).$$

При этом предполагается, что все участвующие в вероятностной модели случайные величины независимы (в совокупности) между собой.

Отметим одно важное свойство функции распределения случайной величины Z. Если случайные величины X и Y независимы и одинаково распределены, то для функции распределения $H(x) = P(Z \le x)$ случайной величины $Z = X - Y$ выполнено, как нетрудно видеть, соотношение:

$$H(-x) = 1 - H(x).$$

Это соотношение означает симметрию функции распределения относительно 0. Плотность такой функции распределения является четной функцией, ее значения в точках x и $(-x)$ совпадают.

Проверка гипотезы однородности связанных выборок в наиболее общем случае сводится к проверке симметрии функции распределения разности $Z = X - Y$ относительно 0.

5.6. Проверка гипотезы симметрии

Рассмотрим методы проверки гипотезы симметрию функции распределения относительно 0. Сначала обсудим, какого типа отклонения от гипотезы симметрии можно ожидать при альтернативных гипотезах?

Как и в случае проверки однородности независимых выборок, в зависимости от вида альтернативной гипотезы выделяют два подуровня моделей. Рассмотрим сначала альтернативу сдвига:

$$H_{13} : G(x) = F(x + a).$$

В этом случае распределение Z при альтернативе отличается сдвигом от симметричного относительно 0. Для проверки гипотезы однородности может быть использован критерий знаковых рангов, разработанный Вилкоксоном (см., например, справочник [2, с. 46–53]).

Он строится следующим образом. Пусть $R(Z_j)$ является рангом $|Z_j|$ в ранжировке от меньшего к большему абсолютных значений разностей $|Z_1|$, $|Z_2|$, ..., $|Z_n|$, $j=1, 2, \ldots, n$. Положим для $j=1, 2, \ldots, n$:

$$Q(Z_j) = \begin{cases} 1, Z_j > 0, \\ 0, Z_j < 0. \end{cases}$$

Статистика критерия знаковых рангов имеет вид:

$$W^+ = \sum_{j=1}^{n} R(Z_j) Q(Z_j).$$

Таким образом, нужно просуммировать ранги положительных разностей в вариационном ряду, построенном стандартным образом по абсолютным величинам всех разностей.

Для практического использования статистики критерия знаковых рангов Вилкоксона либо обращаются к соответствующим таблицам и программному обеспечению, либо применяют асимптотические соотношения. При выполнении нулевой гипотезы статистика:

$$W^{++} = \frac{W^{+} - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

имеет асимптотическое (при $n \to \infty$) стандартное нормальное распределение с математическим ожиданием 0 и дисперсией 1. Следовательно, правило принятия решений на уровне значимости 5 %: имеет обычный вид: если

$$\left| W^{++} \right| \le 1{,}96,$$

то гипотезу однородности связанных выборок по критерию знаковых рангов Вилкоксона принимают, в противном случае отклоняют. Как обычно, при желании использовать другой уровень значимости применяют в качестве критического значения иную квантиль нормального распределения. Повторим еще раз, что использование предельных теорем допустимо при достаточно больших объемах выборки.

Альтернативная гипотеза общего вида записывается как:

$$H_{14} : H(-x_0) \ne 1 - H(x_0)$$

при некотором x_0. Таким образом, проверке подлежит гипотеза симметрии относительно 0, которую можно переписать в виде:

$$H(x) + H(-x) - 1 = 0.$$

Для построенной по выборке $Z_j = x_j - y_j$, j = 1, 2, …, n, эмпирической функции распределения $H_n(x)$ последнее соотношение выполнено лишь приближенно:

$$H_n(x) + H_n(-x) - 1 \approx 0.$$

Как измерять отличие от 0? По тем же соображениям, что и в предыдущем пункте, целесообразно использовать статистику типа омега-квадрат. Соответствующий критерий был предложен в работе [17]. Он имеет вид:

$$\omega_n^2 = \sum_{j=1}^{n} (H_n(Z_j) + H_n(-Z_j) - 1)^2.$$

В настоящем разделе удобно считать, что значение эмпирической функции распределения $H_n(x)$ в точке x равно доле результатов наблюдений в выборке, меньших *или равных* x.

В работе [11] найдено предельное распределение только что введенной статистики:

$$\lim_{n\to\infty} P(\omega_n^2 < x) = S_0(x).$$

Это предельное соотношение выполнено при любом определении эмпирической функции распределения $H_n(x)$. Как тогда, когда значение функции $H_n(x)$ равно доле результатов наблюдений в выборке, меньших *или равных* x. Другими словами, когда $H_n(x)$ строится как непрерывная справа функция — в тех точках, куда попадают результаты наблюдений. Так и тогда, когда значение функции $H_n(x)$ равно доле результатов наблюдений в выборке, меньших x. В этом варианте эмпирическая функция распределения $H_n(x)$ строится как непрерывная слева функция — в тех точках, куда попадают результаты наблюдений.

В табл. 1 приведены критические значения статистики типа омега-квадрат для проверки симметрии распределения (и тем самым для проверки однородности связанных выборок), соответствующие наиболее распространенным значениям уровней значимости. Расчеты проведены Г.В. Мартыновым из МГУ им. М.В. Ломоносова.

Таблица 1

Критические значения статистики ω_n^2, предназначенной для проверки симметрии распределения

Значение функции распределения $S_0(x)$	**Уровень значимости** $\alpha = 1 - S_0(x)$	**Критическое значение** x **статистики** ω_n^2
0,90	0,10	1,20
0,95	0,05	1,66
0,99	0,01	2,80

Как следует из табл. 1, правило принятия решений при проверке однородности связанных выборок в наиболее общей постановке и при уровне значимости 5 % формулируется так. Вычислить статистику ω_n^2. Если $\omega_n^2 \leq 1{,}66$, то принять гипотезу однородности. В противном случае — отвергнуть.

Пример. Пусть величины Z_j, j=1, 2, …, 20, таковы:

20, 18, (– 2), 34, 25, (– 17), 24, 42, 16, 26,
13, (– 23), 35, 21, 19, 8, 27, 11, (– 5), 7.

Соответствующий вариационный ряд $Z(1) < Z(2) < ... < Z(20)$ имеет вид:

(– 23) < (– 17) < (– 5) < (– 2) < 7 < 8 < 11 < 13 < 16 < 18 <
< 19 < 20 < 21 < 24 < 25 < 26 < 27 < 34 < 35 < 42.

Для расчета значения статистики ω_n^2 построим табл. 2 из 7 столбцов и 20 строк, не считая заголовков столбцов (сказуемого таблицы). В первом столбце указаны номера (ранги) членов вариационного ряда, во втором — сами эти члены, в третьем — значения эмпирической функции распределения при значениях аргумента, совпадающих с членами вариационного ряда. В следующем столбце приведены члены вариационного ряда с обратным знаком, а затем указываются соответствующие значения эмпирической функции распределения. Например, поскольку минимальное наблюдаемое значение равно (– 23), то $H_n(x) = 0$ при $x < – 23$, а потому для членов вариационного ряда с 14-го по 20-й в пятом столбце стоит 0. В качестве другого примера рассмотрим минимальный член вариационного ряда, т.е. (– 23). Меняя знак, получаем 23. Это число стоит между 13-м и 14-м членами вариационного ряда, 21 < 23 < 24. На этом интервале эмпирическая функция распределения совпадает со своим значением в левом конце, поэтому следует записать в пятом столбце значение 0,65. Остальные ячейки пятого столбца заполняются аналогично. На основе третьего и пятого столбцов элементарно заполняется шестой столбец, а затем и седьмой. Остается найти сумму значений, стоящих в седьмом столбце. Подобная таблица удобна как для ручного счета, так и при использовании электронных таблиц типа *Excel*.

Таблица 2

Расчет значения статистики ω_n^2
для проверки симметрии распределения

j	$Z(j)$	$H_n(Z(j))$	$– Z(j)$	$H_n(– Z(j))$	$H_n(Z(j)) + H_n(– Z(j)) – 1$	$(H_n(Z(j)) + H_n(– Z(j)) – 1)^2$
1	– 23	0,05	23	0,65	– 0,30	0,09
2	– 17	0,10	17	0,45	– 0,45	0,2025
3	– 5	0,15	5	0,20	– 0,65	0,4225

j	$Z(j)$	$H_n(Z(j))$	$-Z(j)$	$H_n(-Z(j))$	$H_n(Z(j)) + H_n(-Z(j)) - 1$	$(H_n(Z(j)) + H_n(-Z(j)) - 1)^2$
4	– 2	0,20	2	0,20	– 0,60	0,36
5	7	0,25	– 7	0,10	– 0,65	0,4225
6	8	0,30	– 8	0,10	– 0,60	0,36
7	11	0.35	– 11	0,10	– 0,55	0,3025
8	13	0,40	– 13	0,10	– 0,50	0,25
9	16	0,45	– 16	0,10	– 0,45	0,2025
10	18	0,50	– 18	0,05	– 0,45	0,2025
11	19	0,55	– 19	0,05	– 0,40	0,16
12	20	0,60	– 20	0,05	– 0,35	0,1225
13	21	0,65	– 21	0,05	– 0,30	0,09
14	24	0,70	– 24	0	– 0,30	0,09
15	25	0,75	– 25	0	– 0,25	0,0625
16	26	0,80	– 26	0	– 0,20	0,04
17	27	0,85	– 27	0	– 0,15	0,0225
18	34	0,90	– 34	0	– 0,10	0,01
19	35	0,95	– 35	0	– 0,05	0,0025
20	42	1,00	– 42	0	0	0

Результаты расчетов (суммирование значений по седьмому столбцу табл. 2) показывают, что значение статистики ω_n^2=3,055. В соответствии с табл. 1 это означает, что на любом используемом в прикладных статистических исследованиях уровнях значимости отклоняется гипотеза симметрии распределения относительно 0 (а потому и гипотеза однородности в связанных выборках).

5.7. Реальные и номинальные уровни значимости в задачах проверки статистических гипотез

Во многих монографиях и справочных таблицах (например, [12–14]) при проверке статистических гипотез критические значения статистик указаны для априорно фиксированных (номинальных в терминологии [1]) уровней значимости α_f . В качестве таковых обычно используются значения из тройки чисел 0,01, 0,05, 0,1, к которым иногда добавляют еще несколько: 0,001, 0,005 и др.

Однако ясно, что для дискретных статистик (т.е. статистик с дискретными функциями распределения) к которым, в частности, относятся все непараметрические статистические критерии [1, 2], реальные уровни значимости $\alpha_Đ$ могут и совпадать с номинальными. Под $\alpha_Đ$ понимается максимально возможный уровень значимости дискретной статистики, не превосходящий заданный номинальный $\alpha_{\text{í}}$ (т.е. при переходе к следующему по величине возможному значению дискретной статистики соответствующий уровень значимости оказывается больше заданного номинального). Поэтому в лучших таблицах [1, 2] для ограниченных объемов выборок (2–100) табулируются точные распределения дискретных статистик. Для каждой конкретной статистики реальный уровень значимости $\alpha_{\text{Ð}}$ — функция от объемов выборок $n = (n_1, \ldots, n_t)$, т.е. $\alpha_{\text{Ð}} = \alpha_{\text{Ð}}(n)$.

В одних таблицах приведены $\alpha_{\text{Ð}}$ [1, 2], в других — нет [12–14]. Возникает естественный вопрос: с чем это связано? Либо в работах [12–14] нарушена культура табулирования, либо реальные $\alpha_{\text{Ð}}$ и номинальные $\alpha_{\text{í}}$ уровни значимости практически совпадают для всех n. Продемонстрируем, что по крайней мере для некоторых статистик выполнено первое из этих двух утверждений.

В качестве примера рассмотрим критерий серий (Вольфовица) проверки однородности двух независимых выборок. Статистика этого критерия V — это число серий, т.е. частей общего вариационного ряда двух выборок, каждая из которых состоит из элементов одной выборки. При справедливости нулевой гипотезы о тождестве функций распределения, соответствующих двум независимым выборкам объемов n_1 и n_2, известно точное распределение [1, табл. 6.7]:

$$P(V = r \mid n_1, n_2) = \begin{cases} \dfrac{2C_{n_1-1}^{k-1}C_{n_2-1}^{k-1}}{C_{n_1+n_2}^{n_1}}, & \text{åñëè } r = 2k, \\ \dfrac{C_{n_1-1}^{k}C_{n_2-1}^{k-1} + C_{n_1-1}^{k-1}C_{n_2-1}^{k}}{C_{n_1+n_2}^{n_1}}, & \text{åñëè } r = 2k+1, \end{cases}$$

где $r = 2, 3, \ldots, 2n_1$ при $n_1 = n_2$ и $r = 2, 3, \ldots, 2n_1 + 1$ при $n_1 < n_2$ (без ограничения общности можно принять, что объем первой выборки не превосходит объема второй выборки, т.е. $n_1 \leq n_2$).

Несложный расчет для номинального уровня значимости $\alpha_{\text{í}} = 0{,}05$ показывается, что:

- при $n_1 = n_2 = 6$ реальный уровень значимости $\alpha_{\text{Ð}} = 0{,}0260$;
- при $n_1 = n_2 = 8$ реальный уровень значимости $\alpha_{\text{Ð}} = 0{,}0178$;
- при $n_1 = n_2 = 10$ реальный уровень значимости $\alpha_{\text{Ð}} = 0{,}0370$;
- при $n_1 = n_2 = 12$ реальный уровень значимости $\alpha_{\text{Ð}} = 0{,}0190$.

Таким образом, для рассматриваемых объемов выборок реальный уровень значимости в 2–3 раза меньше, чем номинальный. Это, очевидно, необходимо учитывать при интерпретации результатов анализа реальных статистических данных.

В табл. 1, построенным по данным [1, 2, 15], для ряда непараметрических критериев проверки однородности двух независимых выборок приведены реальные уровни значимости $\alpha_D(n)$ для номинального уровня значимости α_f = 0,05 и объемов выборок $n_1 = n_2 = 6, 8, 10, 12$. Проанализированы пять критериев.

1. Двухвыборочный критерий Вилкоксона, являющийся линейной функцией от критерия Манна — Уитни и подробно рассмотренный в разделе 5.3. Напомним, что статистика Вилкоксона S — это сумма рангов элементов первой выборки:

$$S = R_1 + R_2 + \ldots + R_{n_1}$$

в общем вариационном ряду, построенном по объединенной выборке, включающей в себя все элементы обеих выборок (без ограничения общности можно принять, что объем первой выборки не превосходит объема второй выборки, т.е. $n_1 \le n_2$).

2. Критерий Ван-дер-Вардена [1, 15], представляющий собой дальнейшее развитие критерия Вилкоксона и предназначенный для анализа выборок, распределение которых близко к нормальному. Статистика X критерия Ван-дер-Вардена имеет вид:

$$X = \Phi^{-1}\left\{\frac{R_1}{n_1+n_2+1}\right\} + \Phi^{-1}\left\{\frac{R_2}{n_1+n_2+1}\right\} + \ldots + \Phi^{-1}\left\{\frac{R_{n_1}}{n_1+n_2+1}\right\},$$

где $\Phi^{-1}(p)$ есть квантиль порядка p стандартного нормального распределения $\Phi(x)$ с математическим ожиданием 0 и дисперсией 1, т.е. $\Phi^{-1}(p)$ — обратная функция к $\Phi(x)$.

3. Двухвыборочный двухсторонний критерий Смирнова однородности двух независимых выборок рассмотрен в разделе 5.4. Он основан на использовании разности эмпирических функций распределения $F_{n_1}(x)$ и $G_{n_2}(x)$ построенных по первой и второй выборкам соответственно. Термин «двухсторонний» означает, что берется супремум модуля этой разности. Статистика двухвыборочного двухстороннего критерия Смирнова:

$$D = D(n_1, n_2) = \sup_x \left| F_{n_1}(x) - G_{n_2}(x) \right|$$

в случае равенства объемов выборок $n_1 = n_2$ принимает значения кратные $1/n_1$, поскольку только такие значения принимают эмпирические функции распределения $F_{n_1}(x)$ и $G_{n_2}(x)$, а потому имеет $(n_1 + 1)$ возможных значений.

4. Критерий знаков Z используют в случае равенства объемов выборок $n_1 = n_2$. Статистика этого критерия равна числу положительных разностей X_k — Y_k элементов двух выборок с одинаковыми номерами. При справедливости нулевой гипотезы статистика Z имеет биномиальное распределение $B(1/2; n_1)$, а потому имеет (n_1+1) возможных значений.

5. Критерий серий (Вольфовица) V, о котором шла речь выше в начале настоящего значения. Число его возможных значений не превосходит $2n_1$.

Таблица 1

Реальные уровни значимости $\alpha_Р(n)$ для $\alpha_Н = 0{,}05$

Наименование и обозначение критерия	Объемы выборок $n_1 = n_2$				Примечания и ссылки
	6	8	10	12	
Вилкоксона S	0,0320	0,0400	0,0480	0,0420	[2, с. 280–281], [15, с. 418]
Ван-дер-Вардена X	0,0498	0,0498	0,0500	0,0500	Рассчитано по методике [15, с. 249–250]
Смирнова D	0,0044	0,0372	0,0246	0,0158	[1, с. 350], [2, с. 406–427]
Знаков Z	0,0312	0,0078	0,0214	0,0386	[2, с. 273–274]
Вольфовица (серий) V	0,0260	0,0178	0,0370	0,0190	Рассчитано по методике [1, с. 91–92]

Анализ содержания табл. 1 подтверждает предположение о существенности отличия реальных уровней значимости $\alpha_Р(n)$ от номинальных уровней значимости $\alpha_Н$.

Предположим теперь, что, несмотря на установленные отличия, мы используем при проверке гипотезы однородности таблицы [12–14], в которых указаны $\alpha_Н > \alpha_Р$, а не $\alpha_Р$. Это приводит к снижению мощности критерия по сравнению с соответствующим рандомизированным критерием, обеспечивающим равенство $\alpha_Р$ и $\alpha_Н$.

Разъяснение. Поясним, что такое рандомизированный критерий. Пусть Y — статистика некоторого статистического критерия, принимающая соседние значения, числа a и b, где $a < b$, — два соседних значения этой статистики, такие, что $P(Y > b) < \alpha_{í}$ и $P(Y > a) > \alpha_{í}$ (вероятности взяты в предположении справедливости нулевой гипотезы). Если критическое значение критерия равно b, т.е. нулевая гипотеза принимается при $Y \leq b$, то $\alpha_{Ð} = P(Y > b) < \alpha_{í}$. Если же критическое значение равно следующему возможному (при движении в сторону уменьшения) значению a, т.е. нулевая гипотеза принимается при $Y \leq a$, то $\alpha_{Ð} = P(Y > a) > \alpha_{í}$. Рандомизированный критерий получим, если при $Y = b$ в некоторой доле p случаев будем принимать нулевую гипотезу, а в остальных случаях — альтернативную. Поскольку:

$$P(Y = b) = P(Y > a) - P(Y > b),$$

то (реальный) уровень значимости рандомизированного критерия равен

$$(1 - p)\, P(Y = b) + P(Y > b) = (1 - p)\, P(Y > a) + p\, P(Y > b).$$

Ясно, что при соответствующем выборе параметра рандомизации p уровень значимости рандомизированного критерия совпадет с заданным номинальным уровнем $\alpha_{í}$.

Для малых объемов выборок (2–20 элементов) понижение мощности из-за того, что $\alpha_{í} > \alpha_{Ð}$, может быть существенным. Для иллюстрации этого в табл. 2 приведены результаты моделирования наиболее употребительных (согласно [1]) критериев проверки однородности двух независимых выборок.

Моделируются выборки одинакового объема из нормального закона распределения с математическими ожиданиями m_1 и m_2 и дисперсиями σ_1^2 и σ_2^2. Номинальный уровень значимости, определяющий конкретные критические значения для критериев, принят равным $\alpha_{í}$ = 0,05. Мощность критерия определяется моделированием N = 5 000 пар выборок. При использовании N = 5 000 моделируемых пар выборок среднее квадратическое отклонение оценок мощности $\sigma_M \leq 0{,}0223$ (при $M \geq 0{,}95$ имеем $\sigma_M \leq 0{,}01$).

Изучены критерии Вилкоксона S, Вольфовица V, Ван-дер-Вардена X, Смирнова D. Критерий Стьюдента t (см. например, [1]), как равномерно наиболее мощный в классе нормальных законов распределения, приведен для сравнительной оценки мощности рассматриваемых непараметрических критериев. (Моделирование и расчеты, приведенные в настоящем разделе, выполнены Ю.Э. Камнем и Я.Э. Камнем [10].)

Таблица 2

Мощности статистических критериев при α_1 = 0,05

Номер эксперимента	Объем выборок $n_1 = n_2$	Параметры				Мощность M статистического критерия				
		m_1	m_2	σ_1^2	σ_2^2	S	V	X	D	T
1	6	0	1	1	1	0,318	0,006	0,298	0,238	0,396
2	8	0	1	1	1	0,452	0,104	0,426	0,068	0,484
3	10	0	1	1	1	0,520	0,180	0,534	0,116	0,598
4	12	0	1	1	1	0,632	0,076	0,618	0,462	0,682
5	6	0	2	1	1	0,828	0,308	0,808	0,716	0,904
6	8	0	2	1	1	0,958	0,510	0,954	0,458	0,976
7	10	0	2	1	1	0,984	0,704	0,990	0,632	0,988
8	12	0	2	1	1	0,996	0,568	0,996	0,978	0,998

Замечание. Приведенные в табл. 2 значения мощностей критериев интересны нам с точки зрения обсуждения их зависимости от различия реальных и номинальных уровней значимости. При этом необходимо подчеркнуть, что эти значения зависят от предположений, принятых при моделировании. Так, критерии Вилкоксона и Ван-дер-Вардена «настроены» на использование в случае распределений, близких к нормальному семейству. При проверке гипотезы о совпадении функций распределения двух независимых выборок из логистического распределения с альтернативой сдвига критерий Вилкоксона является асимптотически оптимальным. А в случае выборок из нормального распределения аналогичным свойством обладает критерий Ван-дер-Вардена, причем известно, что семейства нормальных и логистических распределений весьма близки — расстояние Колмогорова между ними не превышает 0,01 (см. по вопросам асимптотической оптимальности непараметрических критериев [16–18]). Поэтому нет ничего удивительного в том, что мощности критериев Вилкоксона и Ван-дер-Вардена близки к оптимуму в случае нормального распределения — к мощности критерия Стьюдента. При этом мощности критериев Смирнова и особенно критерия Вольфовица заметно меньше. Однако для выборок из других распределений (например, распределений Вейбулла — Гнеденко или гамма-распределений) ситуация иная — критерий Смирнова, как показывает компьютерное моделирование, оказывается более мощным, чем критерии Вилкоксона и Ван-дер-Вардена. Более того, критерий Смирнова — состоятельный, т.е. позволяет отклонить любую конкретную альтернативу (при соответствующих объемах выборок), а критерии Вилкоксона и Ван-дер-Вардена не являются состоятельными, некоторых альтернатив они «не чувствуют» (см. раздел 5.3). Поэтому вполне обоснованной является рекомендация о широком использовании состоятельных критериев Смирнова и типа омега-квадрат (Лемана — Розенблатта), данная в разделе 5.4. Что же касается критерия серий (Вольфовица), то из-за его отрицательных свойств (выраженной дискретности, низкой мощности) он в настоящее время выходит из употребления при анализе реальных данных, несмотря на прозрачность определения.

Рассмотрения настоящего раздела позволяют сделать следующие выводы [10].

1. При создании методик и таблиц необходимо соблюдать определенную культуру табулирования. В качестве положительных примеров можно указать работы [1, 2].

2. При малых объемах выборок использовать номинальные уровни значимости $\alpha_{н}$ вместо реальных уровней значимости $\alpha_{р}$ для дискретных статистик недопустимо.

3. При конечных объемах выборок выбор того или иного критерия с дискретной статистикой должен сопровождаться исследованием влияния варьирования уровня значимости на качественную интерпретацию результатов проверки гипотез. В частности, выбор одного из двух конкурирующих непараметрических критериев K_1 и K_2 прежде всего должен зависеть от априорного выбора исследователем реального уровня значимости $\alpha_{р1}$ или $\alpha_{р2}$, соответствующего первому критерию K_1 или второму K_2, в качестве номинального уровня значимости $\alpha_{н}$.

Последний вывод демонстрирует сложность сравнения критериев с дискретными статистиками между собой, поскольку точки скачков распределений их статистик не совпадают. Следовательно, в отличие от критериев с непрерывными статистиками нельзя выбрать единый фиксированный уровень значимости и сравнить свойства критериев при этом уровне значимости.

В заключение отметим, что для любого критерия проверки статистических гипотез реальный уровень значимости приближается к номинальному при безграничном возрастании объемов выборок, т.е. $\alpha_{р}(n) \to \alpha_{н}$ при $\min(n_1, \ldots, n_t) \to \infty$. Поэтому для прикладных исследований значительный интерес представляет определение верхней оценки скорости сходимости $\alpha_{р}(n)$ к $\alpha_{н}$. Соответствующие теоретические результаты для критериев проверки однородности двух независимых или связанных выборок можно получить, основываясь на оценках скорости сходимости в принципе инвариантности (см. раздел 14.5). Некоторые оценки приведены в [4, гл. 2]. Скорость сходимости также может быть оценена методом статистических испытаний (Монте-Карло). Пример подобного исследования подробно рассмотрен в разделе 8.5 в ходе обсуждения проблем вероятностно-статистического моделирования помех, создаваемых электровозами.

В настоящей главе затронута лишь небольшая часть непараметрических методов анализа числовых статистических данных. В частности, обратим внимание на непараметрические оценки плотности, которые используются для описания данных, проверки однородности, в задачах восстановления зависимостей и других областях эконометрики. Непараметрические оценки плотности рассмотрены в главе 2.

Литература

1. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1983. — 416 с.

2. *Холлендер, М.* Непараметрические методы статистики / М. Холлендер, Д. Вульф. — Москва : Финансы и статистика, 1983. — 518 с.

3. *Боровков, А.А.* Математическая статистика / А.А. Боровков. — Москва : Наука, 1984. — 472 с.

4. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

5. *Орлов, А.И.* Непараметрическое оценивание коэффициентов вариации технических характеристик и показателей качества / А.И. Орлов, Г.Б. Друянова // Надежность и контроль качества. — 1987. — № 7. — С. 10–16.

6. *Крамер, Г.* Математические методы статистики / Г. Крамер ; перевод с английского. — 2-е изд. — Москва : Мир, 1975. — 648 с.

7. *Гаек, Я.* Теория ранговых критериев / Я. Гаек, 3. Шидак : перевод с английского. — Москва : Наука, 1971. — 376 с.

8. *Смолянский, М.Л.* Таблицы неопределенных интегралов / М.Л. Смолянский. — Москва : ГИФМЛ, 1961. — 108 с.

9. Методика. Проверка однородности двух выборок параметров продукции при оценке ее технического уровня и качества. — Москва : ВНИИ стандартизации, 1987. — 116 с.

10. *Камень, Ю.Э.* Реальные и номинальные уровни значимости в задачах проверки статистических гипотез / Ю.Э. Камень, Я.Э. Камень, А.И. Орлов // Заводская лаборатория. — 1986. — Т. 52. — № 12. — С. 55–57.

11. *Орлов, А.И.* О проверке симметрии распределения / А.И. Орлов // Теория вероятностей и ее применения. — 1972. — Т. 17. — № 2. — С. 372–377.

12. *Афифи, А.* Статистический анализ. Подход с использованием ЭВМ / А. Афифи, С. Эйзен. — Москва : Мир, 1982. — 488 с.

13. *Гублер, Е.В.* Применение критериев непараметрической статистики в медико-биологических исследованиях / Е.В. Гублер, А.А. Генкин. — Ленинград : Медицина, 1973. — 144 с.

14. *Ивченко, Г.И.* Математическая статистика / Г.И. Ивченко, Ю.И. Медведев. — Москва : Высшая школа, 1984. — 248 с.

15. *Ван-дер-Варден, Б.Л.* Математическая статистика / Б.Л. Ван-дер-Варден. — Москва : ИЛ, 1960. — 434 с.

16. *Кендалл, М.Дж.* Статистические выводы и связи / М.Дж. Кендалл, А. Стьюарт. — Москва : Наука, 1973. — 900 с.

17. *Кокс, Д.Р.* Теоретическая статистика / Д.Р. Кокс, Д.В. Хинкли. — Москва : Мир, 1978. — 560 с.

18. *Никитин, Я.Ю.* Асимптотическая эффективность непараметрических критериев / Я.Ю. Никитин. — Москва : Наука, 1995. — 240 с.

19. *Орлов, А.И.* Непараметрическое оценивание характеристик распределений вероятностей / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 112. — С. 1–20.

20. *Орлов, А.И.* Система моделей и методов проверки однородности двух независимых выборок / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 157. — С. 145–169.

21. *Орлов, А.И.* Проверка статистической гипотезы однородности математических ожиданий двух независимых выборок: критерий Крамера-Уэлча вместо критерия Стьюдента / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 110. — С. 197–218.

22. *Орлов, А.И.* Двухвыборочный критерий Вилкоксона — анализ двух мифов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 104. — С. 91–111.

23. *Орлов, А.И.* Состоятельные критерии проверки абсолютной однородности независимых выборок / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2012. — Т. 78. — № 11. — С. 66–70.

24. *Орлов, А.И.* О проверке однородности связанных выборок / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 123. — С. 708–726.

25. *Орлов, А.И.* Компьютерно-статистические методы: состояние и перспективы / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 103. — С. 163–195.

26. *Орлов, А.И.* Реальные и номинальные уровни значимости при проверке статистических гипотез / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 114. — С. 42–54.

Контрольные вопросы и задачи

1. Почему непараметрические методы анализа числовых данных предпочтительнее параметрических?

2. Указать доверительные границы для математических ожиданий (с доверительной вероятностью 0,95) и проверить гипотезу о равенстве математических ожиданий с помощью критерия Крамера-Уэлча (уровень значимости α=0.05):

Номер варианта	n_1	$\overline{X}$	s_x	n_2	$\overline{Y}$	s_y
1	100	13,7	7,3	200	12,1	2,5
2	200	10	5,3	400	12	1,7

3. Проверить гипотезу об однородности функций распределения с помощью критерия Вилкоксона (на уровне значимости α=0.05):

Первая выборка	33	27	12	27	39	42	47	48	50	32
Вторая выборка	11	20	30	31	22	18	17	25	28	29

4. Для каждого из N = 20 объектов даны значения X_j и Y_j , j = 1, 2, …, N, результатов измерений (наблюдений, испытаний, анализов, опытов) двух признаков. Необходимо проверить, есть ли значимое различие между значениями двух признаков или же это различие может быть объяснено случайными отклонениям значений признаков. Другими словами, требуется проверить однородность (т.е. отсутствие различия) связанных выборок.

Таблица 3

Исходные данные для задачи 4

j	1	2	3	4	5	6	7	8	9	10
X_j	74	79	65	69	71	66	71	73	72	68
Y_j	73	65	71	69	70	69	78	70	60	62
j	11	12	13	14	15	16	17	18	19	20
X_j	70	69	76	74	72	69	74	72	77	75
Y_j	61	67	73	67	73	64	67	65	63	70

Проверку однородности на уровне значимости 0,05 проведите с помощью трех критериев:

а) критерия знаков (основанного на проверке гипотезы $p = 0{,}5$ для биномиального распределения с использованием теоремы Муавра-Лапласа);

б) критерия для проверки равенства 0 математического ожидания (критерий основан на асимптотической нормальности выборочного среднего арифметического, деленного на выборочное среднее квадратическое отклонение);

в) критерия типа омега-квадрат для проверки гипотезы симметрии функции распределения (разности результатов измерений, наблюдений, испытаний, анализов, опытов для двух признаков) относительно 0.

5. Какова роль альтернативных гипотез, в частности, гипотезы сдвига, в выборе критерия для проверки нулевой гипотезы?

6. Чем реальный уровень значимости отличается от номинального?

7. Как выбрать параметр рандомизации p, чтобы уровень значимости рандомизированного критерия совпал с заданным номинальным уровнем α_f ?

8. Почему трудно сравнивать между собой статистические критерии проверки гипотез, статистики которых имеют дискретные распределения?

Темы докладов, рефератов, исследовательских работ

1. Асимптотическая нормальность оценок параметров как основа для проверки гипотез о параметрах.

2. Сравнение двухвыборочных критериев Крамера — Уэлча и Стьюдента.

3. Достоинства и недостатки двухвыборочного критерия Вилкоксона по сравнению с другими непараметрическими критериями однородности.

4. Для данных задачи 3 рассчитайте значения статистик Смирнова и типа омега-квадрат (Лемана — Розенблатта) и проверьте однородность двух выборок.

Примечание. В соответствии с [1] для уровня значимости 0,05 критическим значением для критерия Смирнова является 0,7 (т.е. гипотеза однородности отклоняется, если значение статистики Смирнова не менее 0,7). Для того же уровня значимости критическим значением для критерия типа омега-квадрат (Лемана — Розенблатта) является 0,461.

5. Многообразие непараметрических критериев проверки гипотез (по монографиям [1, 2, 7]).

6. Сравнение мощностей непараметрических критериев однородности.

7. Рандомизированные критерии.

8. Подходы к определению асимптотической эффективности непараметрических критериев.

ГЛАВА 6. МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ

В многомерном статистическом анализе выборка состоит из элементов многомерного пространства. Отсюда и название этого раздела статистических методов. Из многих задач многомерного статистического анализа рассмотрим основные — корреляцию, восстановление зависимости, классификацию, уменьшение размерности, индексы.

6.1. Коэффициенты корреляции

Термин «корреляция» означает «связь». В области статистических методов этот термин обычно используется в сочетании «коэффициенты корреляции». Рассмотрим линейный и непараметрические парные коэффициенты корреляции.

Обсудим способы измерения связи между двумя случайными переменными. Пусть исходными данными является набор случайных векторов $(x_i, y_i) = (x_i(\omega), y_i(\omega)),\ i = 1,2,..n.$ Выборочным коэффициентом корреляции, более подробно, выборочным линейным парным коэффициентом корреляции К. Пирсона, как известно, называется число:

$$r_n = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}.$$

Если $r_n = 1$, то $y_i = ax_i + b$, причем $a>0$. Если же $r_n = -1$, то $y_i = ax_i + b$, причем $a<0$. Таким образом, близость коэффициента корреляции к 1 (по абсолютной величине) говорит о достаточно тесной линейной связи.

Если случайные вектора $(x_i, y_i) = (x_i(\omega), y_i(\omega)),\ i = 1,2,..n,$ независимы и одинаково распределены, то выборочный коэффициент корреляции сходится к теоретическому при безграничном возрастании объема выборки:

$$r_n \to \rho = \frac{M(x_1 - M(x_1))(y_1 - M(y_1))}{\sqrt{D(x_1)}\sqrt{D(y_1)}}$$

(сходимость по вероятности).

Более того, выборочный коэффициент корреляции является асимптотически нормальным. Это означает, что:

$$\lim_{n\to\infty} P\left(\frac{r_n-\rho}{\sqrt{D_0(r_n)}}<x\right)=\Phi(x),$$

где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1, а $D_0(r_n)$ — асимптотическая дисперсия выборочного коэффициента корреляции. Она имеет довольно сложное выражение, приведенное в монографии [1, с. 393]:

$$D_0(r_n)=\frac{\rho^2}{4n}\left(\frac{\mu_{40}}{\mu_{20}^2}+\frac{\mu_{04}}{\mu_{02}^2}+\frac{2\mu_{22}}{\mu_{20}\mu_{02}}+\frac{4\mu_{22}}{\mu_{11}^2}-\frac{4\mu_{31}}{\mu_{11}\mu_{20}}-\frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right).$$

Здесь под μ_{km} понимаются теоретические центральные моменты порядка k и m, а именно,

$$\mu_{km}=M(x_1-M(x_1))^k(y_1-M(y_1))^m.$$

Коэффициенты корреляции типа r_n используются во многих алгоритмах многомерного статистического анализа. В теоретических рассмотрениях часто считают, что случайные вектора $(x_i,y_i)=(x_i(\omega),y_i(\omega))$, $i=1,2,..n$, имеют двумерное нормальное распределение. Распределения реальных данных, как правило, отличны от нормальных (см. главу 2.1). Почему же распространено представление о двумерном нормальном распределении? Дело в том, что теория в этом случае проще. В частности, равенство 0 теоретического коэффициента корреляции эквивалентно независимости случайных величин. Поэтому проверка независимости сводится к проверке статистической гипотезы о равенстве 0 теоретического коэффициента корреляции. Эта гипотеза принимается, если $|r_n|<C(n,\alpha)$, где $C(n,\alpha)$ — некоторое граничное значение, зависящее от объема выборки n и уровня значимости α.

Если предположение о двумерной нормальности не выполнено, то из равенства 0 теоретического коэффициента корреляции не вытекает независимость случайных величин. Нетрудно построить пример случайного вектора, для которого коэффициент корреляции равен 0, но координаты зависимы. Кроме того, для проверки гипотез о коэффициенте корреляции нельзя пользоваться таблицами, рассчитанными в предположении нормальности. Можно построить правила принятия решений на основе асимптотической нормальности выборочного коэффициента корреляции. Но есть и другой путь — перейти к непараметрическим коэффициентам корреляции, одинаково пригодным при любом непрерывном распределении случайного вектора.

Для расчета непараметрического *коэффициента ранговой корреляции Спирмена* необходимо сделать следующее. Для каждого x_i рассчитать его ранг r_i в вариационном ряду, построенном по выборке $x_1, x_2, \dots x_n$. Для каждого y_i рассчитать его ранг q_i в вариационном ряду, построенном по выборке $y_1, y_2, \dots y_n$. Для набора из n пар (r_i, q_i), $i = 1, 2, \dots, n,$ вычислить линейный коэффициент корреляции. Он называется коэффициентом ранговой корреляции, поскольку определяется через ранги. В качестве примера рассмотрим данные из табл. 1 (см. монографию [2]).

Таблица 1

Данные для расчета коэффициентов корреляции

i	1	2	3	4	5
x_i	5	10	15	20	25
y_i	6	7	30	81	300
r_i	1	2	3	4	5
q_i	1	2	3	4	5

Для данных табл. 1 коэффициент линейной корреляции равен 0,83, непосредственной линейной связи нет. А вот коэффициент ранговой корреляции равен 1, поскольку увеличение одной переменной однозначно соответствует увеличению другой переменной. Во многих экономических задачах, например, при выборе инвестиционных проектов, достаточно именно монотонной зависимости одной переменной от другой.

Поскольку суммы рангов и их квадратов нетрудно подсчитать, то *коэффициент ранговой корреляции Спирмена* равен:

$$\rho_n = 1 - \frac{6\sum_{i=1}^{n}(r_i - q_i)^2}{n^3 - n}.$$

Отметим, что *коэффициент ранговой корреляции Спирмена* остается постоянным при любом строго возрастающем преобразовании шкалы измерения результатов наблюдений. Другими словами, он является адекватным в порядковой шкале (см. главу 11.3), как и другие ранговые статистики, например, статистики Вилкоксона, Смирнова, типа омега-квадрат для проверки однородности независимых выборок (глава 5).

Широко используется также коэффициент ранговой корреляции τ Кендалла, коэффициент ранговой конкордации Кендалла и Б. Смита (речь идет о сумме попарных коэффициентов ранговой корреляции Кендалла в случае более чем двух переменных) и др. Наиболее подробное обсуждение этой тематики содержится в монографии [3], необходимые для практических расчетов таблицы имеются в справочнике [4]. Дискуссия о наиболее адекватном выборе вида коэффициентов корреляции представлена в [2, 24].

6.2. Восстановление линейной зависимости между двумя переменными

Начнем с задачи точечного и доверительного оценивания линейной функции одной переменной.

Исходные данные — набор n пар чисел (t_k, x_k), $k = 1,2, ..., n$, где t_k — независимая переменная (например, время), а x_k — зависимая (например, индекс инфляции, курс доллара США, объем месячного производства или размер дневной выручки торговой точки). Предполагается, что переменные связаны зависимостью

$$x_k = a\,(t_k - t_{cp}) + b + e_k,\ k = 1,2,\ ...,\ n,$$

где a и b — параметры, неизвестные статистику и подлежащие оцениванию, а e_k — погрешности, искажающие зависимость. Среднее арифметическое моментов времени:

$$t_{cp} = (t_1 + t_2 + ... + t_n)/n$$

введено в модель для облегчения дальнейших выкладок.

Обычно оценивают параметры a и b линейной зависимости методом наименьших квадратов. Затем восстановленную зависимость используют, например, для точечного и интервального прогнозирования.

Как известно, метод наименьших квадратов был разработан великим немецким математиком К. Гауссом в 1794 г. Согласно этому методу для расчета наилучшей функции, приближающей линейным образом зависимость x от t, следует рассмотреть функцию двух переменных:

$$f(a,b) = \sum_{i=1}^{n} (x_i - a(t_i - t_{cp}) - b)^2.$$

Оценки метода наименьших квадратов — это такие значения a^* и b^*, при которых функция $f(a,b)$ достигает минимума по всем значениям аргументов.

Чтобы найти эти оценки, надо вычислить частные производные от функции $f(a,b)$ по аргументам a и b, приравнять их 0, затем из полученных уравнений найти оценки. Имеем:

$$\frac{\partial f(a,b)}{\partial a} = \sum_{i=1}^{n} 2(x_i - a(t_i - t_{cp}) - b)(-(t_i - t_{cp})),$$
$$\frac{\partial f(a,b)}{\partial b} = \sum_{i=1}^{n} 2(x_i - a(t_i - t_{cp}) - b)(-1).$$

Преобразуем правые части полученных соотношений. Вынесем за знак суммы общие множители 2 и (–1). Затем рассмотрим слагаемые. Раскроем скобки в первом выражении, получим, что каждое слагаемое разбивается на три. Во втором выражении также каждое слагаемое есть сумма трех. Значит, каждая из сумм разбивается на три суммы. Имеем:

$$\frac{\partial f(a,b)}{\partial a} = (-2)(\sum_{i=1}^{n} x_i(t_i - t_{cp}) - a\sum_{i=1}^{n}(t_i - t_{cp})^2 - b\sum_{i=1}^{n}(t_i - t_{cp})),$$
$$\frac{\partial f(a,b)}{\partial b} = (-2)(\sum_{i=1}^{n} x_i - a\sum_{i=1}^{n}(t_i - t_{cp}) \ - bn).$$

Приравняем частные производные 0. Тогда в полученных уравнениях можно сократить множитель (–2). Поскольку:

$$\sum_{i=1}^{n}(t_i - t_{cp}) = 0, \qquad (1)$$

уравнения приобретают вид:

$$\sum_{i=1}^{n} x_i(t_i - t_{cp}) - a\sum_{i=1}^{n}(t_i - t_{cp})^2 = 0,$$
$$\sum_{i=1}^{n} x_i - bn = 0.$$

Следовательно, оценки метода наименьших квадратов имеют вид:

$$a^* = \frac{\sum_{i=1}^{n} x_i (t_i - t_{cp})}{\sum_{i=1}^{n} (t_i - t_{cp})^2}, \quad b^* = x_{cp} = \frac{x_1 + x_2 + \ldots + x_n}{n} \quad . \tag{2}$$

В силу соотношения (1) оценку a* можно записать в более симметричном виде:

$$a^* = \frac{\sum_{i=1}^{n} (x_i - x_{cp})(t_i - t_{cp})}{\sum_{i=1}^{n} (t_i - t_{cp})^2}. \tag{3}$$

Эту оценку нетрудно преобразовать и к виду:

$$a^* = \frac{\sum_{i=1}^{n} x_i t_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} t_i^{\,2} - \frac{1}{n} \left(\sum_{i=1}^{n} t_i \right)^2}. \tag{4}$$

Следовательно, восстановленная функция, с помощью которой можно прогнозировать и интерполировать, имеет вид:

$$x^*(t) = a^*(t - t_{cp}) + b^*.$$

Обратим внимание на то, что использование t_{cp} в последней формуле ничуть не ограничивает ее общность. Сравним с моделью вида

$$x_k = c\, t_k + d + e_k,\ k = 1, 2, \ldots, n.$$

Ясно, что:

$$c = a, \ \ d = b - at_{cp}.$$

Аналогичным образом связаны оценки параметров:

$$c^* = a^*, \ \ d^* = b^* - a^* t_{cp}.$$

Для получения оценок параметров и прогностической формулы нет необходимости обращаться к какой-либо вероятностной модели. Однако для того, чтобы изучать погрешности оценок параметров и восстановленной функции, т.е. строить доверительные интервалы для a^*, b^* и $x^*(t)$, подобная модель необходима.

Непараметрическая вероятностная модель. Пусть значения независимой переменной t детерминированы, а погрешности e_k, $k = 1, 2, \ldots, n$, — независимые одинаково распределенные случайные величины с нулевым математическим ожиданием и дисперсией σ^2, неизвестной статистику.

В дальнейшем неоднократно будем использовать Центральную Предельную Теорему (ЦПТ) теории вероятностей для величин e_k, $k = 1, 2, \ldots, n$ (с весами), поэтому для выполнения ее условий необходимо предположить, например, что погрешности e_k, $k = 1, 2, \ldots, n$, финитны или имеют конечный третий абсолютный момент. Однако заострять внимание на этих внутриматематических «условиях регулярности» нет необходимости.

Асимптотические распределения оценок параметров. Из формулы (2) следует, что:

$$b^* = \frac{a}{n}\sum_{i=1}^{n}(t_i - t_{cp}) + b + \frac{1}{n}\sum_{i=1}^{n} e_i = b + \frac{1}{n}\sum_{i=1}^{n} e_i \quad . \tag{5}$$

Согласно ЦПТ, оценка b^* имеет асимптотически нормальное распределение с математическим ожиданием b и дисперсией σ^2/n, оценка которой приводится ниже.

Из формул (2) и (5) вытекает, что:

$$x_i - x_{cp} = a(t_i - t_{cp}) + b + e_i - b - \frac{1}{n}\sum_{i=1}^{n} e_i,$$

$$(x_i - x_{cp})(t_i - t_{cp}) = a(t_i - t_{cp})^2 + e_i(t_i - t_{cp}) - \frac{(t_i - t_{cp})}{n}\sum_{i=1}^{n} e_i.$$

Последнее слагаемое во втором соотношении при суммировании по i обращается в 0, поэтому из формул (2–4) следует, что:

$$a^* = a + \sum_{i=1}^{n} c_i e_i, \quad c_i = \frac{(t_i - t_{cp})}{\sum_{i=1}^{n}(t_i - t_{cp})^2}. \tag{6}$$

Формула (6) показывает, что оценка a^* является асимптотически нормальной с математическим ожиданием a и дисперсией:

$$D(a^*) = \sum_{i=1}^{n} c_i^2 D(e_i) = \frac{\sigma^2}{\sum_{i=1}^{n}(t_i - t_{cp})^2} \quad .$$

Отметим, что многомерная нормальность имеет быть, когда каждое слагаемое в формуле (6) мало сравнительно со всей суммой, т.е.:

$$\lim_{n\to\infty} \max |t_i - t_{cp}| / \{\sum_{i=1}^{n}(t_i - t_{cp})^2\}^{1/2} = 0 \quad .$$

Из формул (5) и (6) и исходных предположений о погрешностях вытекает также несмещенность оценок параметров.

Несмещенность и асимптотическая нормальность оценок метода наименьших квадратов позволяют легко указывать для них асимптотические доверительные границы (аналогично границам в предыдущей главе) и проверять статистические гипотезы, например, о равенстве определенным значениям, прежде всего 0. Предоставляем читателю возможность выписать формулы для расчета доверительных границ и сформулировать правила проверки упомянутых гипотез.

Асимптотическое распределение прогностической функции. Из формул (5) и (6) следует, что:

$$M(x^*(t)) = M\{a^*(t - t_{cp}) + b^*\} = M(a^*)(t - t_{cp}) + M(b^*) = a(t - t_{cp}) + b = x(t),$$

т.е. рассматриваемая оценка прогностической функции является несмещенной. Поэтому:

$$D(x^*(t)) = D(a^*)(t - t_{cp})^2 + 2M\{(a^* - a)(b^* - b)(t - t_{cp})\} + D(b^*).$$

При этом, поскольку погрешности независимы в совокупности и $M(e_i) = 0$, то:

$$M\{(a^* - a)(b^* - b)(t - t_{cp})\} = \frac{1}{n}\sum_{i=1}^{n} c_i (t - t_{cp}) M(e_i^2) = \frac{1}{n}(t - t_{cp})\sigma^2 \sum_{i=1}^{n} c_i = 0 \quad .$$

Таким образом:

$$D(x^*(t)) = \sigma^2\{\frac{1}{n} + \frac{(t-t_{cp})^2}{\sum_{i=1}^{n}(t_i - t_{cp})^2}\} \ .$$

Итак, оценка $x^*(t)$ является несмещенной и асимптотически нормальной. Для ее практического использования необходимо уметь оценивать остаточную дисперсию $M(e_i^2) = \sigma^2$.

Оценивание остаточной дисперсии. В точках t_k, $k = 1, 2, ..., n$, имеются исходные значения зависимой переменной x_k и восстановленные значения $x^*(t_k)$. Рассмотрим остаточную сумму квадратов:

$$SS = \sum_{i=1}^{n}(x^*(t_i) - x(t_i))^2 = \sum_{i=1}^{n}\{(a^* - a)(t_i - t_{cp}) + (b^* - b) - e_i)^2.$$

В соответствии с формулами (5) и (6)

$$SS = \sum_{i=1}^{n}\{(t_i - t_{cp})\sum_{j=1}^{n}c_j e_j + \tfrac{1}{n}\sum_{j=1}^{n}e_j - e_i\}^2 = \sum_{i=1}^{n}\{\sum_{j=1}^{n}\{c_j(t_i - t_{cp}) + \frac{1}{n}\}e_j - e_i\}^2 = \sum_{i=1}^{n}SS_i.$$

Найдем математическое ожидание каждого из слагаемых:

$$M(SS_i) = \sum_{j=1}^{n}\{c_j(t_i - t_{cp}) + \frac{1}{n}\}^2\sigma^2 - 2\{c_i(t_i - t_{cp}) + \frac{1}{n})\sigma^2 + \sigma^2.$$

Из сделанных ранее предположений вытекает, что при $n \to \infty$ имеем $M(SS_i) \to \sigma^2, i = 1,2,...,n,$ следовательно, по закону больших чисел статистика SS/n является состоятельной оценкой остаточной дисперсии σ^2.

Получением состоятельной оценкой остаточной дисперсии завершается последовательность задач, связанных с рассматриваемым простейшим вариантом метода наименьших квадратов. Не представляет труда выписывание верхней и нижней границ для прогностической функции:

$$x_{верх}(t) = a^*(t - t_{cp}) + b^* + \delta(t), \quad x_{нижн}(t) = a^*(t - t_{cp}) + b^* - \delta(t),$$

где погрешность $\delta(t)$ имеет вид:

$$\delta(t) = U(p)\sigma^* \left\{ \frac{1}{n} + \frac{(t - t_{ñð})^2}{\sum_{i=1}^{n}(t_i - t_{ñð})^2} \right\}^{1/2}, \quad \sigma^* = \left(\frac{SS}{n} \right)^{1/2}.$$

Здесь p — доверительная вероятность, $U(p)$, как и в главе 3.1 — квантиль нормального распределения порядка $(1+p)/2$, т.е.:

$$\Phi(U(p)) = \frac{1+p}{2}.$$

При p = 0,95 (наиболее применяемое значение) имеем $U(p)$ = 1,96. Для других доверительных вероятностей соответствующие значения квантилей можно найти в статистических таблицах (см., например, наилучшее в этой сфере издание [4]).

Сравнение параметрического и непараметрического подходов. Во многих литературных источниках рассматривается параметрическая вероятностная модель метода наименьших квадратов. В ней предполагается, что погрешности имеют нормальное распределение. Это предположение позволяет математически строго получить ряд выводов. Так, распределения статистик вычисляются точно, а не в асимптотике, соответственно вместо квантилей нормального распределения используются квантили распределения Стьюдента, а остаточная сумма квадратов SS делится не на n, а на $(n - 2)$. Ясно, что при росте объема данных различия стираются.

Рассмотренный выше непараметрический подход не использует нереалистическое предположение о нормальности погрешностей (см. раздел 2.1). Платой за это является асимптотический характер результатов. В случае простейшей модели метода наименьших квадратов оба подхода дают практически совпадающие рекомендации. Это не всегда так, не всегда два подхода дают близкие результаты. Напомним, что в задаче обнаружения выбросов методы, опирающиеся на нормальное распределение, нельзя считать обоснованными, и обнаружено было это их свойство с помощью непараметрического подхода (см. раздел 2.3).

Общие принципы. Кратко сформулируем несколько общих принципов построения, описания и использования методов прикладной статистики. Во-первых, должны быть четко сформулированы исходные предпосылки, т.е. полностью описана используемая вероятностно-статистическая модель. Во-вторых, не следует принимать предпосылки, которые редко выполняются на практике. В-третьих, алгоритмы расчетов должны быть корректны с точки зрения математико-статистической теории. В-четвертых, алгоритмы должны давать полезные для практики выводы.

Применительно к задаче восстановления зависимостей это означает, что целесообразно применять непараметрический подход, что и сделано выше. Однако предположение нормальности, хотя и очень сильно сужает возможности применения, с чисто математической точки зрения позволяет продвинуться дальше. Поэтому для первоначального изучения ситуации, так сказать, «в лабораторных условиях», нормальная модель может оказаться полезной.

Пример оценивания по методу наименьших квадратов. Пусть даны $n = 6$ пар чисел (t_k, x_k), $k = 1, 2, \ldots, 6$, представленных во втором и третьем столбцах табл. 2. В соответствии с формулами (2) и (4) выше для вычисления оценок метода наименьших квадратов достаточно найти суммы выражений, представленных во втором, третьем, четвертом и пятом столбцах табл. 2.

Таблица 2

Расчет по методу наименьших квадратов при восстановлении линейной функции одной переменной

i	t_i	x_i	t_i^2	$t_i x_i$	$a^* t_i$	$\hat{x}_i$	$x_i - \hat{x}_i$	$(x_i - \hat{x}_i)^2$
1	1	12	1	12	3,14	12,17	– 0,17	0,03
2	3	20	9	60	9,42	18,45	1,55	2,40
3	4	20	16	80	12,56	21,59	– 1,59	2,53
4	7	32	49	224	21,98	31,01	0,99	0,98
5	9	35	81	315	28,26	37,29	– 2,29	5,24
6	10	42	100	420	31,40	40,43	1,57	2,46
Σ	34	161	256	1111	–	–	0,06	13,64
$\frac{\Sigma}{n}$	5,67	26,83	42,67	185,17	–	–	–	–

В соответствии с формулой (2) $b^* = 26{,}83$, а согласно формуле (4):

$$a^* = \frac{1111 - \frac{1}{6}161 \times 34}{256 - \frac{1}{6}(34)^2} = \frac{1111 - 912{,}33}{256 - 192{,}67} = \frac{198{,}67}{63{,}33} = 3{,}14.$$

Следовательно, прогностическая формула имеет вид:

$$x^*(t) = 3{,}14(t - 5{,}67) + 26{,}83 = 3{,}14t - 3{,}14 \times 5{,}67 + 26{,}83 =$$
$$= 3{,}14t - 17{,}80 + 26{,}83 = 3{,}14t + 9{,}03.$$

Следующий этап анализа данных — оценка точности приближения функции методом наименьших квадратов. Сначала рассматриваются т.н. восстановленные значения:

$$\hat{x}_i = x^*(t_i), \quad i = 1,2,\ldots,n.$$

Это те значения, которые полученная в результате расчетов прогностическая функция принимает в тех точках, в которых известны истинные значения зависимой переменной x_i.

Сравним восстановленные и истинные значения. Это и сделано в шестом — восьмом столбцах табл. 2. Для простоты расчетов в шестом столбце представлены произведения $a \cdot t_i$, седьмой отличается от шестого добавлением константы 9,03 и содержит восстановленные значения. Восьмой столбец — это разность третьего и седьмого.

Непосредственный анализ восьмого столбца табл. 2 показывает, что содержащиеся в нем числа сравнительно невелики по величине по сравнению с третьим столбцом (на порядок меньше по величине). Кроме того, знаки «+» и «-» чередуются. Эти два признака свидетельствуют о правильности расчетов. При использовании метода наименьших квадратов знаки не всегда чередуются. Однако если сначала идут только плюсы, а потом только минусы (или наоборот, сначала только минусы, а потом только плюсы), то это верный показатель того, что в вычислениях допущена ошибка.

Верно следующее утверждение.

Теорема.

$$\sum_{i=1}^{n}(x_i - \hat{x}_i) = 0.$$

Доказательство этой теоремы оставляем читателю в качестве упражнения.

Однако сумма по восьмому столбцу дает 0,06, а не 0. Незначительное отличие от 0 связано с ошибками округления при вычислениях. Близость суммы значений зависимой переменной и суммы восстановленных значений — практический критерий правильности расчетов.

В последнем девятом столбце табл. 2 приведены квадраты значений из восьмого столбца. Их сумма — это остаточная сумма квадратов SS = 13,64. В соответствии со сказанным выше оценками дисперсии погрешностей и их среднего квадратического отклонения являются:

$$(\sigma^2)^* = \frac{SS}{n} = \frac{13{,}64}{6} = 2{,}27; \quad \sigma^* = \sqrt{\frac{SS}{n}} = \sqrt{\frac{13{,}64}{6}} = 1{,}49.$$

Рассмотрим распределения оценок параметров. Оценка b^* имеет асимптотически нормальное распределение с математическим ожиданием b и дисперсией, которая оценивается как 2,27/6=0,38 (здесь считаем, что 6 — «достаточно большое» число, что, конечно, можно оспаривать). Оценкой среднего квадратического отклонения является 0,615. Следовательно, при доверительной вероятности 0,95 доверительный интервал для параметра b имеет вид (26,83 – 1,96·0,615; 26,83 + 1,96·0,615) = (25,625; 28,035).

В формулах для дисперсий участвует величина:

$$\sum_{i=1}^{n}(t_i - t_{cp})^2 = \sum_{i=1}^{n}(t_i^2 - 2t_i t_{cp} + t_{cp}^2) = \sum_{i=1}^{n} t_i^2 - 2t_{cp}\sum_{ш=1}^{n} t_i + nt_{cp}^2 = \sum_{i=1}^{n} t_i^2 - nt_{cp}^2.$$

Подставив численные значения, получаем, что:

$$\sum_{i=1}^{n} t_i^2 - nt_{cp}^2 = 256 - 6(5{,}67)^2 = 63{,}1.$$

Дисперсия для оценки a^* коэффициента при линейном члене прогностической функции оценивается как 2,27/63,1 = 0,036, а среднее квадратическое отклонение — как 0,19. Следовательно, при доверительной вероятности 0,95 доверительный интервал для параметра a имеет вид (3,14 – 1,96·0,19; 3,14 + 1,96·0,19) = (2,77; 3,51).

Прогностическая формула с учетом погрешности имеет вид (при доверительной вероятности 0,95):

$$x^*(t) = 3{,}14t + 9{,}03 \pm 1{,}96 \times 1{,}49\sqrt{\frac{1}{6} + \frac{(t - 5{,}67)^2}{63{,}1}}.$$

В этой записи сохранено происхождение различных составляющих. Упростим:

$$x^*(t) = 3{,}14t + 9{,}03 \pm 2{,}92\sqrt{\frac{1}{6} + \frac{(t-5{,}67)^2}{63{,}1}}.$$

Например, при $t = 12$ эта формула дает:

$$x^*(12) = 46{,}71 \pm 2{,}615.$$

Следовательно, нижняя доверительная граница — это 44,095, а верхняя доверительная граница — это 49,325.

Насколько далеко можно прогнозировать? Обычный ответ таков — до тех пор, пока сохраняется тот стабильный комплекс условий, при котором справедлива рассматриваемая зависимость. Изобретатель метода наименьших квадратов Карл Гаусс исходил из задачи восстановления орбиты астероида (малой планеты) Церера. Движение подобных небесных тел может быть рассчитано на сотни лет. А вот параметры комет (например, срок возвращения) не поддаются столь точному расчету, поскольку за время пребывания в окрестности Солнца сильно меняется масса кометы. В социально-экономической области горизонты надежного прогнозирования еще менее определены. В частности, они сильно зависят от решений центральной власти.

Чтобы выявить роль погрешностей в прогностической формуле, рассмотрим формальный предельный переход $t \to \infty$. Тогда слагаемые 9,03; 1/6; 5,67 становятся бесконечно малыми, и:

$$x^*(t) \approx 3{,}14t \pm \frac{2{,}92}{\sqrt{63{,}1}}t = (3{,}14 \pm 0{,}37)t.$$

Таким образом, погрешности составляют около

$$\frac{100 \cdot 0{,}37}{3{,}14}\,\% = 11{,}8\,\%$$

от тренда (математического ожидания) прогностической функции. В социально-экономических исследованиях подобные погрешности считаются вполне приемлемыми.

6.3. Основы линейного регрессионного анализа

Метод наименьших квадратов, рассмотренный в простейшем случае, допускает различные обобщения. Например, метод наименьших квадратов дает алгоритм расчетов, если исходные данные — по-прежнему набор n пар чисел (t_k , x_k), $k = 1,2,...,n$, где t_k — независимая переменная (например, время), а x_k — зависимая (например, индекс инфляции), а восстанавливать надо не линейную зависимость, а квадратическую:

$$x(t) = at^2 + bt + c.$$

Следует рассмотреть функцию трех переменных:

$$f(a,b,c) = \sum_{k=1}^{n}(x_k - at_k^2 - bt_k - c)^2.$$

Оценки метода наименьших квадратов — это такие значения параметров a^*, b^* и c^*, при которых функция $f(a,b,c)$ достигает минимума по всем значениям аргументов. Чтобы найти эти оценки, надо вычислить частные производные от функции $f(a,b,c)$ по аргументам a, b и c, приравнять их 0, затем из полученных уравнений найти оценки. Имеем:

$$\frac{\partial f(a,b,c)}{\partial a} = \sum_{k=1}^{n}\frac{\partial}{\partial a}(x_k - at_k^2 - bt_k - c)^2 = \sum_{k=1}^{n}2(-t_k^2)(x_k - at_k^2 - bt_k - c)^2.$$

Приравнивая частную производную к 0, получаем линейное уравнение относительно трех неизвестных параметров a, b, c:

$$a\sum_{k=1}^{n}t_k^4 + b\sum_{k=1}^{n}t_k^3 + c\sum_{k=1}^{n}t_k^2 = \sum_{k=1}^{n}t_k^2 x_k.$$

Приравнивая частную производную по параметру b к 0, аналогичным образом получаем уравнение:

$$a\sum_{k=1}^{n}t_k^3 + b\sum_{k=1}^{n}t_k^2 + c\sum_{k=1}^{n}t_k = \sum_{k=1}^{n}t_k x_k.$$

Наконец, приравнивая частную производную по параметру c к 0, получаем уравнение:

$$a\sum_{k=1}^{n} t_k^2 + b\sum_{k=1}^{n} t_k + cn = \sum_{k=1}^{n} x_k.$$

Решая систему трех уравнений с тремя неизвестными, находим оценки метода наименьших квадратов.

Другие задачи, рассмотренные в предыдущем подразделе (доверительные границы для параметров и прогностической функции и др.), также могут быть решены. Соответствующие алгоритмы более громоздки. Для их записи полезен аппарат матричной алгебры (см., например, одну из лучших в этой области монографий [5]). Для реальных расчетов используют соответствующие компьютерные программы.

Раздел прикладной статистики, посвященный восстановлению зависимостей, называется регрессионным анализом. Термин «линейный регрессионный анализ» используют, когда рассматриваемая функция линейно зависит от оцениваемых параметров (от независимых переменных зависимость может быть произвольной). Теория оценивания неизвестных параметров хорошо развита именно в случае линейного регрессионного анализа. Если же линейности нет и нельзя перейти к линейной задаче, то, как правило, хороших свойств от оценок ожидать не приходится.

Продемонстрируем подходы в случае зависимостей различного вида. Если зависимость имеет вид многочлена (полинома):

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \ldots + a_m t^m,$$

то коэффициенты многочлена могут быть найдены путем минимизации функции:

$$f(a_0, a_1, a_2, a_3, \ldots, a_m) = \sum_{k=1}^{n} (x_k - a_0 - a_1 t_k - a_2 t_k^2 - a_3 t_k^3 - \ldots - a_m t_k^m)^2.$$

Функция от t не обязательно должна быть многочленом. Можно, например, добавить периодическую составляющую, соответствующую сезонным колебаниям. Хорошо известно, например, что инфляция (рост потребительских цен) имеет четко выраженный годовой цикл. А именно, в среднем цены быстрее всего растут зимой, в декабре — январе, а медленнее всего (иногда в среднем даже падают) летом, в июле — августе. Пусть для определенности:

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \ldots + a_m t^m + A \sin Bt,$$

тогда неизвестные параметры могут быть найдены путем минимизации функции:

$$f(a_0, a_1, a_2, a_3, \ldots, a_m, A, B) = \sum_{k=1}^{n} (x_k - a_0 - a_1 t_k - a_2 t_k^2 - a_3 t_k^3 - \ldots - a_m t_k^m - A \sin Bt_k)^2.$$

Пусть $I(t)$ -индекс инфляции в момент t. Принцип стабильности условий приводит к гипотезе о постоянстве темпов роста средних цен, т.е. индекса инфляции. Таким образом, естественная модель для индекса инфляции — это:

$$I(t) = Ae^{Bt}.$$

Эта модель не является линейной, метод наименьших квадратов непосредственно применять нельзя. Однако если прологарифмировать обе части предыдущего равенства:

$$\ln I(t) = \ln A + Bt,$$

то получим линейную зависимость, рассмотренную выше.

Независимых переменных может быть не одна, а несколько. Пусть, например, по исходным данным $(x_k, y_k, z_k), k = 1,2,\ldots,n,$ требуется оценить неизвестные параметры a и b в зависимости:

$$z = ax + by + \varepsilon,$$

где ε — погрешность. Это можно сделать, минимизируя функцию:

$$f(a,b) = \sum_{k=1}^{n} (z_k - ax_k - by_k)^2.$$

Зависимость от x и y не обязательно должна быть линейной. Предположим, что из каких-то соображений известно, что зависимость должна иметь вид:

$$z = ax + by + cx^2 y + dxy + ey^3 + \varepsilon,$$

тогда для оценки пяти параметров необходимо минимизировать функцию:

$$f(a,b,c,d,e) = \sum_{k=1}^{n} (z_k - ax_k - by_k - cx_k^2 y_k - dx_k y_k - ey_k^3)^2.$$

Более подробно рассмотрим пример из микроэкономики. В одной из оптимизационных моделей поведения фирмы используется т.н. производственная функция $f(K,L)$, задающая объем выпуска в зависимости от затрат капитала K и труда L. В качестве конкретного вида производственной функции часто используется так называемая функция Кобба — Дугласа:

$$f(K,L) = K^{\alpha} L^{\beta}.$$

Однако откуда взять значения параметров α и β? Естественно предположить, что они — одни и те же для предприятий отрасли. Поэтому целесообразно собрать информацию (f_k, K_k, L_k), $k = 1,2,\ldots,n$, где f_k — объем выпуска на k-м предприятии, K_k — объем затрат капитала на k-ом предприятии, L_k — объем затрат труда на k-м предприятии (в кратком изложении не пытаемся дать точных определений используемым понятиям из экономики предприятия). По собранной информации естественно попытаться оценить параметры α и β. Но они входят в зависимость нелинейно, поэтому сразу применить метод наименьших квадратов нельзя. Помогает логарифмирование:

$$\ln f(K,L) = \alpha \ln K + \beta \ln L.$$

Следовательно, целесообразно сделать *замену переменных*:

$$x_k = \ln K_k, y_k = \ln L_k, z_k = \ln f_k, k = 1,2,3,\ldots,n,$$

а затем находить оценки параметров α и β, минимизируя функцию

$$g(\alpha,\beta) = \sum_{k=1}^{n} (z_k - \alpha x_k - \beta y_k)^2.$$

Найдем частные производные:

$$\frac{\partial g(\alpha,\beta)}{\partial\alpha}=\sum_{k=1}^{n}2(z_k-\alpha x_k-\beta y_k)(-x_k),$$
$$\frac{\partial g(\alpha,\beta)}{\partial\beta}=\sum_{k=1}^{n}2(z_k-\alpha x_k-\beta y_k)(-y_k).$$

Приравняем частные производные к 0, сократим на 2, раскроем скобки, перенесем свободные члены вправо. Получим систему двух линейных уравнений с двумя неизвестными:

$$\alpha\sum_{k=1}^{n}x_k^2+\beta\sum_{k=1}^{n}x_k y_k=\sum_{k=1}^{n}x_k z_k,$$

$$\alpha\sum_{k=1}^{n}x_k y_k+\beta\sum_{k=1}^{n}y_л^2=\sum_{k=1}^{n}y_k z_k.$$

Таким образом, для вычисления оценок метода наименьших квадратов необходимо найти пять сумм:

$$\sum_{k=1}^{n}x_k^2,\quad \sum_{k=1}^{n}x_k y_k,\quad \sum_{k=1}^{n}y_k^2,\quad \sum_{k=1}^{n}x_k z_k,\quad \sum_{k=1}^{n}y_k z_k.$$

Для упорядочения расчета этих сумм может быть использована таблица типа той, что применялась выше. Отметим, что рассмотренная в предыдущем подразделе постановка переходит в разбираемую сейчас при $y_k=1,\ k=1,2,\ldots,n$.

Подходящая замена переменных во многих случаях позволяет перейти к линейной зависимости. Например, если

$$y=\frac{1}{a+bx},$$

то замена $z=1/y$ приводит к линейной зависимости $z = a + bx$. Если $y = (a + bx)^2$, то замена $z=\sqrt{y}$ приводит к линейной зависимости $z = a + bx$.

Основной показатель качества регрессионной модели. Одни и те же данные можно обрабатывать различными способами. На первый взгляд, показателем отклонений данных от модели может служить остаточная сумма квадратов SS. Чем этот показатель меньше, тем приближение лучше, значит, и модель лучше описывает реальные данные. Однако это рассуждение годится только для моделей с одинаковым числом параметров. Ведь если добавляется новый параметр, по которому можно минимизировать, то и минимум, как правило, оказывается меньше.

В качестве основного показателя качества регрессионной модели используют оценку остаточной дисперсии:

$$\hat{\sigma}^2(m) = \frac{SS}{n-m},$$

скорректированную на число m параметров, оцениваемых по наблюдаемым данным. В случае задачи восстановления линейной функции одной переменной, рассмотренной в предыдущем подразделе, оценка остаточной дисперсии имеет вид:

$$\hat{\sigma}^2 = \frac{SS}{n-2},$$

поскольку число оцениваемых параметров m=2.

Почему эта формула отличается от приведенной в предыдущем подразделе? Там в знаменателе n, а здесь — $(n-2)$. Дело в том, что там была рассмотрена непараметрическая теория при большом объеме данных (при $n \to \infty$). А при безграничном возрастании n разница между n и $(n-2)$ сходит на нет.

Однако *при подборе вида модели* знаменатель дроби, оценивающей остаточную дисперсию, приходится корректировать на число параметров. Если этого не делать, то придется заключить, что всегда многочлен второй степени лучше соответствует данным, чем линейная функция, многочлен третьей степени лучше приближает исходные данные, чем многочлен второй степени, и т.д. В конце концов доходим до многочлена степени $(n-1)$ с n коэффициентами, который проходит через все заданные точки. Но его прогностические возможности, скорее всего, существенно меньше, чем у линейной функции. *Излишнее усложнение статистических моделей вредно.*

Типовое поведение скорректированной оценки остаточной дисперсии:

$$v(m) = \hat{\sigma}^2(m)$$

в зависимости от параметра m в случае расширяющейся системы моделей выглядит так. Сначала наблюдаем заметное убывание. Затем оценка остаточной дисперсии колеблется около некоторой константы (теоретического значения дисперсии погрешности).

Поясним ситуацию на примере модели восстановления зависимости, выраженной многочленом:

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \ldots + a_m t^m .$$

Пусть эта модель справедлива при $m = m_0$. При $m < m_0$ в скорректированной оценке остаточной дисперсии учитываются не только погрешности измерений, но и соответствующие (старшие) члены многочлена (предполагаем, что коэффициенты при них отличны от 0). При $m \geq m_0$ имеем:

$$\lim_{n \to \infty} v(m) = \sigma^2 .$$

Следовательно, скорректированная оценка остаточной дисперсии будет колебаться около указанного предела. Поэтому в качестве оценки неизвестной статистику степени многочлена (полинома) можно использовать первый локальный минимум скорректированной оценки остаточной дисперсии:

$$m^* = \min\{m : v(m-1) > v(m), \quad v(m) \leq v(m+1)\} .$$

В работе [6] найдено предельное распределение этой оценки степени многочлена.

Теорема. При справедливости некоторых условий регулярности:

$$\lim_{n \to \infty} P(m^* < m_0) = 0, \quad \lim_{n \to \infty} P(m^* = m_0 + u) = \lambda(1-\lambda)^u , \ \ u = 0,1,2,\ldots.,$$

где

$$\lambda = \Phi(1) - \Phi(-1) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} \exp\left\{-\frac{x^2}{2}\right\} dx \approx 0{,}68268.$$

Таким образом, предельное распределение оценки m^* степени многочлена (полинома) является геометрическим. Это означает, в частности, что оценка не является состоятельной. При этом вероятность получить меньшее значение, чем истинное, исчезающе мала. Далее имеем:

$$P(m^* = m_0) \to 0{,}68268, \quad P(m^* = m_0 + 1) \to 0{,}68268(1 - 0{,}68268) = 0{,}21663,$$

$$P(m^* = m_0 + 2) \to 0{,}68268(1 - 0{,}68268)^2 = 0{,}068744,$$

$$P(m^* = m_0 + 3) \to 0{,}68268(1 - 0{,}68268)^3 = 0{,}021814\ldots$$

Разработаны и иные методы оценивания неизвестной степени многочлена, например, путем многократного применения процедуры проверки адекватности регрессионной зависимости с помощью статистики Фишера (см. работу [7]). Предельное поведение оценок — таково же, как в приведенной выше теореме, только значение параметра λ иное.

Пример практического использования линейного регрессионного анализа. Руководитель маркетинговой службы завода ГАРО (г. Великий Новгород) А.А. Пивень применил его для построения математической модели рынка легковых подъемников. Требуется выявить факторы (показатели), оказывающие наибольшее влияние на объем продаж подъемников, найти зависимость объема продаж от этих факторов и использовать эту зависимость для прогнозирования объема продаж.

Зависимая переменная — объем продаж V, независимые переменные:

- грузоподъемность (X1);
- цена (X2);
- наличие напольной рамы (X3);
- наличие синхронизации (X4);
- количество двигателей (X5);
- суммарная мощность двигателей (X6);
- высота подхвата в нижнем положении (X7);
- максимальная высота подъема (X8);
- скорость подъема (X9);
- гарантийный срок (X10);
- срок службы (X11);
- время на рынке (X12);
- внешний вид (X13);
- срок поставки (X14);
- уровень сервисного обслуживания (X15);
- наличие системы смазки (X16);
- масса (X17).

Для восстановления зависимости использовалась линейная регрессионная модель. По результатам пошагового анализа из рассмотрения последовательно исключались независимые переменные (параметры подъемника), имеющие (в линейной модели) коэффициенты, незначимо отличающиеся от нуля, иными словами, мало отличающиеся в сравнении с их дисперсией.

В результате расчетов получена зависимость объема продаж подъемника П3-Т от 12 факторов:

$$V = -1\,769.77 - 65.09X1 - 0.03X2 + 68.79X3 + 147.54X4 \\ + 156.28X5 + 2.53X7 + 1.06X8 + 25.75X12 \\ - 132.26X13 - 12.41X14 + 107.78X15 + 397X16.$$

Влияние остальных пяти факторов оказалось незначимым.

Исходя из расчетов, прогнозное значение продаж подъемников на второй год продаж составит ориентировочно 1 010 шт. С вероятностью 95 % можно утверждать, что объем продаж будет лежать в границах [695, 1 332] шт.

Оценивание условного математического ожидания. Рассмотрим общее понятие регрессии как условного математического ожидания. Пусть случайный вектор $(x(\omega), y(\omega))$ имеет плотность $p(x,y)$. Как известно из любого курса теории вероятностей, плотность условного распределения $y(\omega)$ при условии $x(\omega) = x_0$ имеет вид:

$$p(y \mid x) = p(y \mid x(\omega) = x_0) = \frac{p(x,y)}{\int\limits_{-\infty}^{+\infty} p(x,y)dy}.$$

Условное математическое ожидание, т.е. регрессионная зависимость y от x, имеет вид:

$$f(x) = \int\limits_{-\infty}^{+\infty} yp(y \mid x)dy = \frac{\int\limits_{-\infty}^{+\infty} yp(x,y)dy}{\int\limits_{-\infty}^{+\infty} p(x,y)dy}.$$

Таким образом, для нахождения оценок регрессионной зависимости достаточно найти оценки совместной плотности распределения вероятности $p_n(x,y)$ такие, что:

$$p_n(x,y) \to p(x,y)$$

при $n \to \infty$. Тогда непараметрическая оценка регрессионной зависимости:

$$f_n(x) = \frac{\int_{-\infty}^{+\infty} y p_n(x,y)dy}{\int_{-\infty}^{+\infty} p_n(x,y)dy}$$

при $n \to \infty$ является состоятельной оценкой регрессии как условного математического ожидания:

$$f_n(x) \to f(x).$$

Общий подход к построению непараметрических оценок плотности распределения вероятностей развит в главе 2.1 выше.

Регрессионному анализу (т.е. методам восстановления зависимостей) посвящена огромная литература. Он хорошо представлен в программных продуктах по анализу данных, особенно та его часть, которая связана с методом наименьших квадратов. Обзор современных методов и моделей дан в учебнике [6].

Когда догоним? Рассмотрим практически важную задачу об оценивании точки пересечения двух регрессионных прямых. Пусть зависимость от времени t некоторого показателя качества $x_1(t)$ продукции нашего предприятия описывается линейной функцией:

$$x_1(t) = a_1 t + d_1.$$

Пусть аналогичный показатель у конкурента также описывается линейной функцией, но с другими параметрами:

$$x_2(t) = a_2 t + d_2.$$

Предположим, что мы находимся в положении догоняющей стороны. Это значит, что в рассматриваемый момент времени t_0 (например, «сегодня») значение показателя качества у нашей продукции ниже:

$$x_1(t_0) < x_2(t_0),$$

но темп роста у нас выше, чем у конкурента:

$$a_1 > a_2.$$

Возникает естественный вопрос — когда мы догоним конкурента? Другими словами, в какой момент времени $x_1(t) = x_2(t)$? Решая относительно t уравнение

$$a_1t + d_1 = a_2t + d_2,$$

получаем, что встреча произойдет в момент:

$$t_B = \frac{d_2 - d_1}{a_1 - a_2}.$$

Представляют интерес еще две величины. Во-первых, уровень качества, при котором мы сравняемся с конкурентом, т.е. общий уровень качества в момент встречи:

$$x = x_1(t_B) = x_2(t_B) = \frac{a_1d_2 - a_2d_1}{a_1 - a_2}.$$

Во-вторых, временной лаг, т.е. величина нашего отставания в рассматриваемый момент времени t_0. В какой (более ранний) момент времени t_k конкурент имел тот уровень качества, которого мы достигли сейчас? Для ответа на этот вопрос надо решить уравнение:

$$x_2(t) = x_1(t_0).$$

Решением является:

$$t_k = \frac{x_1(t_0) - d_2}{a_2}.$$

Следовательно, мы отстаем на:

$$L = t_0 - t_k = \frac{(a_2 - a_1)t_0 + d_2 - d_1}{a_2} = \frac{x_2(t_0) - x_1(t_0)}{a_2}$$

единиц времени (лет).

Вероятностно-статистическая модель в задаче о точке встречи. В реальных ситуациях линейные зависимости неизвестны. Однако известны исходные данные $(t_{i1}; x_{i1}), i = 1,2,...,n(1)$, для нашего предприятия и $(t_{j1}; x_{j1}), j = 1,2,...,n(2)$, для предприятия-конкурента. При этом значения показателя качества $x_1(t_{i1}) = x_{i1}$ у нашего предприятия в моменты времени t_{i1} представляются в виде

$$x_1(t_{i1}) = x_{i1} = a_1 t_{i1} + d_1 + e_{i1},\ i = 1, 2, \ldots, n, \tag{1}$$

где коэффициенты a_1 и d_1 неизвестны статистику, а e_{i1} — погрешности измерения (невязки). Как и ранее в разделе 6.2, будем считать, что e_{i1}, $i = 1, 2, \ldots, n(1)$, — совокупность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием и дисперсией $D(e_{i1}) = \sigma_1^2$, неизвестной статистику.

Для предприятия-конкурента справедливо аналогичное представление:

$$x_2(t_{j2}) = x_{j2} = a_2 t_{j2} + d_2 + e_{j2},\ j = 1, 2, \ldots, n, \tag{2}$$

где коэффициенты a_2 и d_2 неизвестны статистику, а e_{j2} — погрешности измерения (невязки). Примем, что e_{j2}, $j = 1, 2, \ldots, n$ (2), — совокупность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием и дисперсией $D(e_{j2}) = \sigma_2^2$, неизвестной статистику.

Примем, что две совокупности случайных величин e_{i1}, $i = 1, 2, \ldots, n$ (1), и e_{j2}, $j = 1, 2, \ldots, n$ (2), независимы между собой. В каждой совокупности случайные величины одинаково распределены, но функции распределения, соответствующие разным совокупностям (т.е. нашему предприятию и предприятию-конкуренту), могут различаться между собой. Как и в разделе 6.2, не предполагаем, что эти функции распределения входят в какое-либо параметрическое семейство распределений (в частности, не предполагаем, что невязки имеют нормальное распределение). Это и значит, что рассматривается непараметрическая постановка. Однако считаем, что объемы данных n (1) и n (2) достаточно велики, так что можно применять центральную предельную теорему и приближать совместное распределение оценок метода наименьших квадратов с помощью многомерного нормального распределения.

Итак, решение задачи о точке встречи получим в рамках непараметрической вероятностно-статистической модели. Весьма частный случай, когда невязки e_{i1}, $i = 1, 2, \ldots, n$ (1), и e_{j2}, $j = 1, 2, \ldots, n$ (2), имеют нормальное распределение, рассмотрен в [22].

Метод решения задачи о встрече. Вместо неизвестных статистику зависимостей $x_1(t)$ и $x_2(t)$ будем использовать их оценки $x_1^*(t)$ и $x_2^*(t)$, полученные методом наименьших квадратов. Для этого необходимо оценить коэффициенты по правилам, полученным в разделе 6.2, а затем рассчитать оценки момента встречи t_B^*, уровня качества в момент встречи $x^* = x_1^*(t_B^*) = x_2^*(t_B^*)$ и временного лага (величины отставания):

$$L^* = \frac{x_2^*(t_0) - x_1^*(t_0)}{a_2^*},$$

используя оценки коэффициентов зависимостей $x_1(t)$ и $x_2(t)$ вместо неизвестных истинных коэффициентов.

Полезным является, как и в разделе 6.2, использование центрирования средними значениями независимой переменной при параметризации зависимостей:

$$x_1(t) = a_1(t - t_{\text{ср}}(1)) + b_1 = a_1 t + d_1,$$
$$x_2(t) = a_2(t - t_{\text{ср}}(2)) + b_2 = a_2 t + d_2,$$

где

$$t_{\text{ср}}(1) = \frac{t_{11} + t_{21} + \ldots + t_{n(1)1}}{n(1)}, \quad t_{\text{ср}}(2) = \frac{t_{12} + t_{22} + \ldots + t_{n(2)2}}{n(2)},$$
$$d_k = b_k - a_k t_{\text{ср}}(k), \ \ k = 1,2.$$

Связано это с тем, что асимптотическое описание совместного распределения коэффициентов линейной зависимости проще в случае центрированной зависимости, в частности, оценки коэффициентов $a_k^*, b_k^*, k = 1,2,$ асимптотически независимы (см. раздел 6.2).

Как показано выше, оценки метода наименьших квадратов имеют вид

$$a_k^* = \frac{\sum_{i=1}^{n(k)} x_{ik}(t_{ik} - t_{\text{ср}}(k))}{\sum_{i=1}^{n(k)} (t_{ik} - t_{\text{ср}}(k))^2}, \quad b_k^* = x_{\text{ср}}(k) = \frac{x_{1k} + x_{2k} + \ldots + x_{n(k)k}}{n(k)}, \ \ k = 1,2.$$

Точечные оценки момента встречи t_B^*, уровня качества в момент встречи x^* и временного лага L^* выражаются через оценки коэффициентов линейных зависимостей так:

$$t_B^* = \frac{d_2^* - d_1^*}{a_1^* - a_2^*} = \frac{b_2^* - b_1^* + a_1^* t_{\text{ñð}}(1) - a_2^* t_{\text{ñð}}(2)}{a_1^* - a_2^*},$$

$$x^* = \frac{a_1^* d_2^* - a_2^* d_1^*}{a_1^* - a_2^*} = \frac{a_1^* b_2^* - a_2^* b_1^* + a_1^* a_2^* (t_{\text{ñð}}(1) - t_{\text{ñð}}(2))}{a_1^* - a_2^*},$$

$$L^* = \frac{(a_2^* - a_1^*) t_0 + d_2^* - d_1^*}{a_2^*} = \frac{(a_2^* - a_1^*) t_0 + b_2^* - b_1^* + a_1^* t_{\text{ñð}}(1) - a_2^* t_{\text{ñð}}(2)}{a_2^*}.$$

Из приведенных формул вытекает:

$$t_B^* = f_1(a_1^*, a_2^*, b_1^*, b_2^*),\ \ x^* = f_2(a_1^*, a_2^*, b_1^*, b_2^*),\ \ L^* = f_1(a_1^*, a_2^*, b_1^*, b_2^*),$$

где

$$f_1(z_1, z_2, z_3, z_4) = \frac{z_4 - z_3 + z_1 t_{\text{ñð}}(1) - z_2 t_{\text{ñð}}(2)}{z_1 - z_2},$$

$$f_2(z_1, z_2, z_3, z_4) = \frac{z_1 z_4 - z_2 z_3 + z_1 z_2 (t_{\text{ñð}}(1) - t_{\text{ñð}}(2))}{z_1 - z_2},$$

$$f_3(z_1, z_2, z_3, z_4) = \frac{(z_2 - z_1) t_0 + z_4 - z_3 + z_1 t_{\text{ñð}}(1) - z_2 t_{\text{ñð}}(2)}{z_2}.$$

Поскольку все входящие в полученные формулы моменты времени предполагаются заданными (детерминированными), то интересующие нас оценки задаются гладкими функциями от четырехмерного вектора $(a_1^*, a_2^*, b_1^*, b_2^*)$ оценок метода наименьших квадратов коэффициентов в линейных зависимостях.

Асимптотическое распределение вектора оценок МНК. Как показано в разделе 6.2, $a_1^*, a_2^*, b_1^*, b_2^*$ являются несмещенными, их дисперсии таковы:

$$D(a_k^*) = \frac{\sigma_k^2}{\sum_{i=1}^{n(k)} (t_{ik} - t_{\text{ñð}}(k))^2}, \quad D(b_k^2) = \frac{\sigma_k^2}{n(k)}, \quad k = 1{,}2.$$

Отметим, что в соответствии с принятыми предположениями все четыре дисперсии стремятся к 0 при безграничном росте n (1) и n (2).

Все ковариации вектора $(a_1^*, a_2^*, b_1^*, b_2^*)$ равны 0. Для пар координат с различающимися нижними индексами это вытекает из предположения о независимости между собой совокупностей невязок, соответствующим измерениям значений двух разных линейных функций. Для пар координат с одинаковыми нижними индексами, т.е. для пар (a_1^*, b_1^*) и (a_2^*, b_2^*), это установлено в подразделе «Асимптотическое распределение прогностической функции» раздела 6.2. Таким образом, в ковариационной матрице вектора $(a_1^*, a_2^*, b_1^*, b_2^*)$ отличны от 0 только элементы, стоящие на главной диагонали, т.е. дисперсии.

Каждый из векторов (a_1^*, b_1^*) и (a_2^*, b_2^*) является суммой n (1) и n (2) слагаемых соответственно. Если каждое из слагаемых мало по сравнению со всей суммой, т.е. если

$$\lim_{n(k)\to\infty} \max_{1\le i\le n(k)} | t_{ik} - t_{\tilde{n}\delta}(k) | \Big/ \left\{ \sum_{i=1}^{n(k)} (t_{ik} - t_{\tilde{n}\delta}(k))^2 \right\}^{1/2} = 0, \quad k = 1,2,$$

то при больших $n(1)$ и $n(2)$ распределение вектора $(a_1^*, a_2^*, b_1^*, b_2^*)$ приближается нормально распределенным случайным вектором с независимыми координатами. Математические ожидания и дисперсии координат приближающего вектора совпадают с одноименными характеристиками вектора $(a_1^*, a_2^*, b_1^*, b_2^*)$. Другими словами, вектор $(a_1^*, a_2^*, b_1^*, b_2^*)$ является асимптотически нормальным с указанными выше параметрами.

Распределение функции от вектора оценок МНК. Если функция $f(z_1, z_2, z_3, z_4)$ достаточно гладкая, то согласно методу линеаризации (см. раздел 14.4):

$$f(a_1^*, a_2^*, b_1^*, b_2^*) - f(a_1, a_2, b_1, b_2) = \frac{\partial f}{\partial z_1}\left(a_1^* - a_1\right) +$$
$$+ \frac{\partial f}{\partial z_2}\left(a_2^* - a_2\right) + \frac{\partial f}{\partial z_3}\left(b_1^* - b_1\right) + \frac{\partial f}{\partial z_4}\left(b_2^* - b_2\right)$$

с точностью до бесконечно малых величин более высокого порядка.

Как показано выше, правая часть последней формулы приближается суммой четырех независимых нормально распределенных величин с нулевыми математическими ожиданиями. Следовательно, функция $f(a_1^*, a_2^*, b_1^*, b_2^*)$ от вектора оценок МНК является асимптотически нормальной случайной величиной с математическим ожиданием $f(a_1, a_2, b_1, b_2)$, совпадающим с теоретическим значением, и дисперсией:

$$Df(a_1^*, a_2^*, b_1^*, b_2^*) = \left(\frac{\partial f}{\partial z_1}\right)^2 D\left(a_1^*\right) + \left(\frac{\partial f}{\partial z_2}\right)^2 D\left(a_2^*\right) + \left(\frac{\partial f}{\partial z_2}\right)^2 D\left(b_1^*\right) + \left(\frac{\partial f}{\partial z_4}\right)^2 D\left(b_2^*\right).$$

Подставив приведенные выше значения дисперсий, получаем:

$$Df(a_1^*, a_2^*, b_1^*, b_2^*) = \left(\frac{\partial f}{\partial z_1}\right)^2 \frac{\sigma_1^2}{\sum\limits_{i=1}^{n(1)} (t_{i1} - t_{\text{ñð}}(1))^2} +$$

$$+\left(\frac{\partial f}{\partial z_2}\right)^2 \frac{\sigma_2^2}{\sum\limits_{i=1}^{n(2)} (t_{i2} - t_{\text{ñð}}(2))^2} + \left(\frac{\partial f}{\partial z_2}\right)^2 \frac{\sigma_1^2}{n(1)} + \left(\frac{\partial f}{\partial z_4}\right)^2 \frac{\sigma_2^2}{n(2)}.$$

Асимптотическое распределение момента встречи. Начнем с функции $t_B^* = f_1(a_1^*, a_2^*, b_1^*, b_2^*)$. Имеем:

$$\frac{\partial f_1}{\partial z_1} = \frac{z_3 - z_4 + z_2(t_{\text{ñð}}(2) - t_{\text{ñð}}(1))}{(z_1 - z_2)^2},$$

$$\frac{\partial f_1}{\partial z_2} = \frac{z_4 - z_3 + z_1(t_{\text{ñð}}(1) - t_{\text{ñð}}(2))}{(z_1 - z_2)^2},$$

$$\frac{\partial f_1}{\partial z_3} = -\frac{1}{z_1 - z_2},$$

$$\frac{\partial f_1}{\partial z_4} = \frac{1}{z_1 - z_2}.$$

В приведенных выше формулах частные производные можно брать как в точке (a_1, a_2, b_1, b_2), так и в точке $(a_1^*, a_2^*, b_1^*, b_2^*)$. Различение — бесконечно малые величины более высокого порядка. Поскольку истинные значения коэффициентов линейных зависимостей неизвестны, частные производные будем брать в точке $(a_1^*, a_2^*, b_1^*, b_2^*)$.

Из последних формул с помощью несложных преобразований получаем:

$$D(t_B^*) = \left(\frac{b_1^* - b_2^* + a_2^*(t_{\text{ñð}}(2) - t_{\text{ñð}}(1))}{(a_1^* - a_2^*)^2}\right)^2 \frac{\sigma_1^2}{\sum\limits_{i=1}^{n(1)} (t_{i1} - t_{\text{ñð}}(1))^2} +$$

$$+\left(\frac{b_1^* - b_2^* + a_1^*(t_{\text{ñð}}(2) - t_{\text{ñð}}(1))}{(a_1^* - a_2^*)^2}\right)^2 \frac{\sigma_2^2}{\sum\limits_{i=1}^{n(2)} (t_{i2} - t_{\text{ñð}}(2))^2} + \frac{1}{(a_1^* - a_2^*)^2}\left(\frac{\sigma_1^2}{n(1)} + \frac{\sigma_2^2}{n(2)}\right).$$

Для практического применения полученных результатов остается заменить неизвестные дисперсии невязок σ_1^2 и σ_2^2 на их состоятельные оценки. Как показано в разделе 6.2, при больших объемах данных n (1) и n (2) используют оценки дисперсий невязок:

$$(\sigma_1^2)^* = \frac{SS(1)}{n(1)}, \quad (\sigma_2^2)^* = \frac{SS(2)}{n(2)},$$

где $SS(1)$ и $SS(2)$ — соответствующие остаточные суммы квадратов. В соответствии с рассуждениями начала настоящего раздела можно порекомендовать применение несмещенных оценок дисперсий невязок:

$$(\sigma_1^2)^{**} = \frac{SS(1)}{n(1)-2}, \quad (\sigma_2^2)^{**} = \frac{SS(2)}{n(2)-2}.$$

Ясно, что с ростом объемов данных n (1) и n (2) различие между двумя последними формулами исчезает.

Доверительное оценивание и проверка гипотез в задаче о точке встречи. На основе полученных результатов легко указать методы доверительного оценивания и проверки гипотез для момента встречи t_B. Так, асимптотический доверительный интервал, соответствующий доверительной вероятности p, имеет вид:

$$\left[t_B^* - U(p)\{D^*(t_B^*)\}^{1/2};\ t_B^* + U(p)\{D^*(t_B^*)\}^{1/2}\right].$$

Здесь $D^*(t_B^*)$ — только что описанная оценка дисперсии случайной величины t_B (с использованием той или иной оценки дисперсий невязок), $U(p)$ — квантиль нормального распределения порядка $(1+p)/2$, т.е.:

$$\Phi(U(p)) = \frac{1+p}{2},$$

где $\Phi(w)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

Рассмотрим проверку гипотезы H_0 о том, что встреча произойдет не позднее чем в заданный момент T, т.е.:

$$H_0\colon t_B \le T,$$

при альтернативной гипотезе:

$$H_1\colon t_B > T.$$

Ясно, что правило принятия решений должно выглядеть так:

$$t_B^* \le K \Rightarrow H_0,\ t_B^* > K \Rightarrow H_1.$$

Пороговое значение K выбирают, задавая уровень значимости α: при справедливости простой гипотезы $t_B = T$ должно быть справедливо равенство:

$$P(t_B^* > K) = \alpha.$$

Ясно, что следует принять:

$$K = T + u(1-\alpha)\{D^*(t_B^*)\}^{1/2},$$

где $u(1-\alpha)$ — квантиль порядка $(1-\alpha)$ стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

Отметим, что рассмотренная схема проверки гипотез позволяет заключить, что в случае $t_B^* > K$ обгон в намеченные сроки практически наверняка (с вероятностью $(1-\alpha)$, близкой к 1) не произойдет. Если же мы хотим выяснить условия гарантированного обгона в заданные сроки, то следует поменять местами гипотезы, в качестве нулевой рассмотреть гипотезу:

$$H_0\colon t_B > T,$$

а в качестве альтернативной — гипотезу:

$$H_1\colon t_B \le T.$$

Тогда правило принятия решений будет иметь вид:

$$t_B^* > K(1) \Rightarrow H_0,\ t_B^* \le K(1) \Rightarrow H_1,$$

где

$$K(1) = T - u(1-\alpha)\{D^*(t_B^*)\}^{1/2}$$

в прежних обозначениях.

Итак, можно сказать, что при $t_B^* \leq K(1)$ обгон практически наверняка произойдет в заданные сроки, а при $t_B^* > K$ — практически наверняка не произойдет, в то время как при $K(1) < t_B^* \leq K$ ситуация является неопределенной

По аналогичной схеме можно оценить асимптотические дисперсии оценок уровня качества в момент встречи x^* и временного лага L* и использовать полученные результаты для доверительного оценивания и проверки гипотез. Поскольку новые идеи при этом не используются, не будем приводить соответствующие выкладки.

О практическом применении статистических оценок точки встречи. Приведенные выше результаты, касающиеся оценок момента встречи t_B^*, уровня качества в момент встречи x^* и временного лага L*, были получены в результате решения прикладных проблем, возникших при разработке системы автоматического проектирования (САПР) стандартов на продукцию [23]. В этой системе реализуются функции информационного обеспечения и анализа данных о характеристиках качества группы отечественных и зарубежных образцов (марок, моделей) аналогичной продукции, требований нормативно-технической документации на эту продукцию, а также поддерживаются функции человек0-машинного принятия решений по стандартизации.

Статистические методы в САПР стандартов используются для анализа распределений показателей качества продукции, исследования взаимосвязей показателей, выявления группировок продукции по уровню качества, анализа временных рядов и прогнозирования качества продукции. Основные проблемы программной реализации этих методов связаны с обеспечением человеко-машинного решения задач пользователями (инженерами по стандартизации и техническому регулированию), не имеющими специальной подготовки по статистическим методам. Кроме того, возникли специфические задачи, требующие совершенствования известного статистического аппарата, в частности, задача сравнительного анализа тенденций развития отечественной и зарубежной групп продукции, решению которой и посвящены предыдущие страницы. Система САПР стандартов реализована в виде пакета прикладных программ, включающего программы статистического анализа уровня качества продукции. Отдельные программы переданы в государственный фонд алгоритмов и программ. Разработанное математическое и программное обеспечение применялось для анализа данных о характеристиках качества изделий электронной техники. Госстандарт отметил их высокую практическую ценность.

6.4. Статистические методы классификации

При внедрении современных статистических методов в практику фундаментальных и прикладных научно-технических, социально-экономических, медицинских и иных исследований, при разработке соответствующих программных продуктов невозможно обойтись без классификации самих этих методов. Логично исходить из вида обрабатываемых данных. В соответствии с современными воззрениями делим статистические методы и, прежде всего, прикладную статистику на четыре области: — статистика случайных величин (одномерная статистика); многомерный статистический анализ; статистика временных рядов и случайных величин; статистика объектов нечисловой природы. В первой области элемент выборки — число, во второй — вектор, в третьей — функция, в четвертой — объект нечисловой природы.

Как известно, математический аппарат статистики объектов нечисловой природы базируется на использовании расстояний (мер близости, показателей различия) в пространствах таких объектов. Это вызвано отсутствием в таких пространствах операций суммирования, на которых основано большинство методов других областей статистики. Любые методы, использующие только расстояния (меры близости, показатели различия) между объектами, следует относить к статистике объектов нечисловой природы, поскольку такие методы могут работать с объектами произвольного пространства, если в нем задана метрика или ее аналоги. Таким образом, весьма многие методы прикладной статистики следует включать в статистику объектов нечисловой природы.

В настоящем разделе рассматривается важное направление статистических методов — математические методы классификации. Значительную их часть следовало бы отнести к статистике объектов нечисловой природы, а именно, методы классификации, основанные на расстояниях между объектами. Однако исторически теория классификации рассматривается в основном в рамках многомерного статистического анализа, поскольку многие ее методы используют специфику конечномерного евклидова пространства.

Основные направления в математической теории классификации. Какие научные исследования относить к этой теории? Исходя из потребностей специалиста, применяющего математические методы классификации, целесообразно принять, что сюда входят исследования, во-первых, отнесенные самими авторами к этой теории; во-вторых, связанные с ней общностью тематики, хотя бы их авторы и не упоминали термин «классификация». Это предполагает ее сложную внутреннюю структуру.

В литературных источниках наряду с термином «классификация» в близких смыслах используются термины «группировка», «распознавание образов», «диагностика», «дискриминация», «сортировка» и др. Терминологический разнобой связан прежде всего с традициями научных кланов, к которым относятся авторы публикаций, а также с внутренним делением самой теории классификации.

В научных исследованиях по современной теории классификации можно выделить два относительно самостоятельных направления. Одно из них опирается на опыт таких наук, как биология, география, геология, и таких прикладных областей, как ведение классификаторов продукции и библиотечное дело. Типичные объекты рассмотрения — классификация химических элементов (таблица Д.И. Менделеева), биологическая систематика, универсальная десятичная классификация публикаций (УДК), классификатор товаров на основе штрих-кодов.

Другое направление опирается на опыт технических исследований, экономики, маркетинговых исследований, социологии, медицины. Типичные задачи — техническая и медицинская диагностика, а также, например, разбиение на группы отраслей промышленности, тесно связанных между собой, выделение групп однородной продукции. Обычно используются такие термины, как «распознавание образов» или «дискриминантный анализ». Это направление обычно опирается на математические модели; для проведения расчетов интенсивно используются компьютеры. Однако относить его к математике столь же нецелесообразно, как астрономию или квантовую механику. Рассматриваемые математические модели можно и нужно изучать на формальном уровне, и такие исследования проводятся. Но направление в целом сконцентрировано на решении конкретных задач прикладных областей и вносит вклад в технические или экономические науки, медицину, социологию, но, как правило, не в математику. Использование математических методов как инструмента исследования нельзя относить к чистой математике.

В 60-х гг. XX в. внутри прикладной статистики достаточно четко оформилась область, посвященная методам классификации. Несколько модифицируя формулировки М. Дж. Кендалла и А. Стьюарта 1966 г. (см. русский перевод [8, с. 437]), в теории классификации выделим три подобласти: дискриминация (дискриминантный анализ), кластеризация (кластер-анализ), группировка. Опишем эти подобласти.

В дискриминантном анализе классы предполагаются заданными — плотностями вероятностей или обучающими выборками. Задача состоит в том, чтобы вновь поступающий объект отнести в один из этих классов. У понятия «дискриминация» имеется много синонимов: диагностика, распознавание образов с учителем, автоматическая классификация с учителем, статистическая классификация и т.д.

При кластеризации и группировке целью является выявление и выделение классов. Синонимы — построение классификации, распознавание образов без учителя, автоматическая классификация без учителя, типология, таксономия и др. Задача кластер-анализа состоит в выяснении по эмпирическим данным, насколько элементы «группируются» или распадаются на изолированные «скопления», «кластеры». Название происходит от *cluster* (англ.) — гроздь, скопление. Иными словами, задача — выявление естественного разбиения на классы, свободного от субъективизма исследователя, а цель — выделение групп однородных объектов, сходных между собой, при резком отличии этих групп друг от друга.

При группировке, наоборот, «мы хотим разбить элементы на группы независимо от того, естественны ли границы разбиения или нет» [8, с. 437]. Цель по-прежнему состоит в выявлении групп однородных объектов, сходных между собой (как в кластер-анализе), однако «соседние» группы могут не иметь резких различий (в отличие от кластер-анализа). Границы между группами условны, не являются естественными, зависят от субъективизма исследователя. Аналогично при лесоустройстве проведение просек (границ участков) зависит от специалистов лесного ведомства, а не от свойств леса.

Задачи кластеризации и группировки принципиально различны, хотя для их решения могут применяться одни и те же алгоритмы. Важная для практической деятельности проблема состоит в том, чтобы понять, разрешима ли задача кластер-анализа для конкретных данных или возможна только их группировка, поскольку совокупность объектов достаточно однородна и не разбивается на резко разделяющиеся между собой кластеры.

Как правило, в математических задачах кластеризации и группировки основное — выбор метрики, расстояния между объектами, меры близости, сходства, различия. Хорошо известно, что для любого заданного разбиения объектов на группы и любого $\varepsilon > 0$ можно указать метрику такую, что расстояния между объектами из одной группы будут меньше ε, а между объектами из разных групп — больше $1/\varepsilon$. Тогда любой разумный алгоритм кластеризации даст именно заданное разбиение.

Понимание и обсуждение постановок задач осложняется использованием одного и того же термина в разных смыслах. Термином «классификация» (и термином «диагностика») обозначают, по крайней мере, три разные вещи: процедуру построения классификации (и выделение классов, используемых при диагностике), построенную классификацию (систему выделенных классов) и процедуру ее использования (правила отнесения вновь поступающего объекта к одному из ранее выделенных классов). Другими словами, имеем естественную триаду: построение — изучение — использование классификации.

Как уже отмечалось, для построения системы диагностических классов используют разнообразные методы кластерного анализа и группировки объектов. Наименее известен второй член триады (отсутствующий у Кендалла и Стьюарта [8]) — изучение отношений эквивалентности, полученных в результате построения системы диагностических классов. Статистический анализ полученных, в частности экспертами, отношений эквивалентности — часть статистики бинарных отношений и тем самым — статистики объектов нечисловой природы (см. раздел 3.4).

Диагностика в узком смысле слова (процедура использования классификации, т.е. отнесения вновь поступающего объекта к одному из выделенных ранее классов) — предмет дискриминантного анализа. Отметим, что с точки зрения статистики объектов нечисловой природы дискриминантный анализ является частным случаем общей схемы регрессионного анализа, соответствующим ситуации, когда зависимая переменная принимает конечное число значений, а именно — номера классов, а вместо квадрата разности стоит функция потерь от неправильной классификации. Однако есть ряд специфических постановок, выделяющих задачи диагностики среди всех регрессионных задач.

О построении диагностических правил. Начнем с краткого обсуждения одного распространенного заблуждения. Иногда рекомендуют сначала построить систему диагностических классов, а потом в каждом диагностическом классе отдельно проводить регрессионный анализ (в классическом смысле) или применять иные методы многомерного статистического анализа. Однако обычно забывают, что при этом нельзя опираться на вероятностную модель многомерного нормального распределения, так как распределение результатов наблюдений, попавших в определенный кластер, будет отнюдь не нормальным, а усеченным нормальным (усечение определяется границами кластера).

Процедуры построения диагностических правил делятся на вероятностные и детерминированные. К первым относятся так называемые задачи расщепления смесей. В них предполагается, что распределение вновь поступающего случайного элемента является смесью вероятностных законов, соответствующих диагностическим классам. Как и при выборе степени полинома в регрессии (см. предыдущий подраздел), при анализе реальных социально-экономических данных встает вопрос об оценке числа элементов смеси, т.е. числа диагностических классов. Были изучены результаты применения обычно рекомендуемого критерия Уилкса для оценки числа элементов смеси. Оказалось (см. статью [9]), что оценка с помощью критерия Уилкса не является состоятельной, асимптотическое распределение этой оценки — геометрическое, как и в случае задачи восстановления зависимости в регрессионном анализе. Итак, продемонстрирована несостоятельность обычно используемых оценок. Для получения состоятельных оценок достаточно связать уровень значимости в критерии Уилкса с объемом выборки, как это было предложено и для задач регрессии [7].

Как уже отмечалось, задачи построения системы диагностических классов целесообразно разбить на два типа: с четко разделенными кластерами (задачи кластер-анализа) и с условными границами, непрерывно переходящими друг в друга классами (задачи группировки). Такое деление полезно, хотя в обоих случаях могут применяться одинаковые алгоритмы. Сколько же существует алгоритмов построения системы диагностических правил? Иногда называют то или иное число. На самом же деле их бесконечно много, в чем нетрудно убедиться.

Действительно, рассмотрим один определенный алгоритм — алгоритм средней связи. Он основан на использовании некоторой меры близости $d(x,y)$ между объектами x и y. Как он работает? На первом шаге каждый объект рассматривается как отдельный кластер. На каждом следующем шаге объединяются две ближайших кластера. Расстояние между объектами рассчитывается как средняя связь (отсюда и название алгоритма), т.е. как среднее арифметическое расстояний между парами объектов, один из которых входит в первый кластер, а другой — во второй. В конце концов все объекты объединяются вместе, и результат работы алгоритма представляет собой дерево последовательных объединений (в терминах теории графов), или «Дендрограмму». Из нее можно выделить кластеры разными способами. Один подход — исходя из заданного числа кластеров. Другой — из соображений предметной области. Третий — исходя из устойчивости (если разбиение долго не менялось при возрастании порога объединения — значит, оно отражает реальность). И т.д.

К алгоритму средней связи естественно сразу добавить алгоритм ближайшего соседа (когда расстоянием между кластерами называется минимальное из расстояний между парами объектов, один из которых входит в первый кластер, а другой — во второй). А также и алгоритм дальнего соседа (когда расстоянием между кластерами называется максимальное из расстояний между парами объектов, один из которых входит в первый кластер, а другой — во второй).

Каждый из трех описанных алгоритмов (средней связи, ближайшего соседа, дальнего соседа), как легко проверить, порождает бесконечное (континуальное) семейство алгоритмов кластер-анализа. Дело в том, что величина $d^a(x,y)$, $a>0$, также является мерой близости между x и y и порождает новый алгоритм. Если параметр a пробегает отрезок, то получается бесконечно много алгоритмов классификации.

Каким из них пользоваться при обработке данных? Дело осложняется тем, что практически в любом пространстве данных мер близости различных видов существует весьма много. Именно в связи с обсуждаемой проблемой следует указать на принципиальное различие между кластер-анализом и задачами группировки.

Если классы реальны, естественны, существуют на самом деле, четко отделены друг от друга, то любой алгоритм кластер-анализа их выделит. Следовательно, ***в качестве критерия естественности классификации следует рассматривать устойчивость относительно выбора алгоритма кластер-анализа.***

Проверить устойчивость можно, применив к данным несколько подходов, например, столь непохожие алгоритмы, как «ближнего соседа» и «дальнего соседа». Если полученные результаты содержательно близки, то они адекватны действительности. В противном случае следует предположить, что естественной классификации не существует, задача кластер-анализа не имеет решения, и можно проводить только группировку.

Как уже отмечалось, часто применяется т.н. агломеративный иерархический алгоритм «Дендрограмма», в котором вначале все элементы рассматриваются как отдельные кластеры, а затем на каждом шагу объединяются два наиболее близких кластера. Для работы «Дендрограммы» необходимо задать правило вычисления расстояния между кластерами. Оно вычисляется через расстояние $d(x,y)$ между элементами x и y. Поскольку $d^a(x,y)$ при $0<a<1$ также расстояние, то, как правило, существует бесконечно много различных вариантов этого алгоритма. Представим себе, что они применяются для обработки одних и тех же реальных данных. Если при всех a получается одинаковое разбиение элементов на кластеры, т.е. результат работы алгоритма устойчив по отношению к изменению a (в смысле общей схемы устойчивости, рассмотренной в главе 1.4), то имеем «естественную» классификацию. В противном случае результат зависит от субъективно выбранного исследователем параметра a, т.е. задача кластер-анализа неразрешима (предполагаем, что выбор *a* нельзя специально обосновать). Задача группировки в этой ситуации имеет много решений. Из них можно выбрать одно по дополнительным критериям.

Следовательно, получаем эвристический критерий: если решение задачи кластер-анализа существует, то оно находится с помощью любого алгоритма. Целесообразно использовать наиболее простой.

Проблема поиска естественной классификации. Существуют различные точки зрения на эту проблему. Естественная классификация обычно противопоставляется искусственной. На Всесоюзной школе-семинаре «Использование математических методов в задачах классификации» (г. Пущино, 1986 г.), в частности, были высказаны мнения, что естественная классификация:

- закон природы;
- основана на глубоких закономерностях, тогда как искусственная классификация — на неглубоких;
- для конкретного индивида та, которая наиболее быстро вытекает из его тезауруса;
- удовлетворяет многим целям; цель искусственной классификации задает человек;
- классификация с точки зрения потребителя продукции;
- классификация, позволяющая делать прогнозы;
- имеет критерием устойчивость.

Приведенные высказывания уже дают представление о больших расхождениях в понимании «естественной классификации». Этот термин следует признать нечетким, как, впрочем, и многие другие термины, и профессиональные — социально-экономические, научно-технические, и используемые в обыденном языке. Нетрудно подробно обоснована нечеткость естественного языка и тот факт, что «мы мыслим нечетко», что, однако, не слишком мешает нам решать производственные и жизненные проблемы. Кажущееся рациональным требование выработать сначала строгие определения, а потом развивать науку — невыполнимо. Следовать ему — значит отвлекать силы от реальных задач. При системном подходе к теории классификации становится ясно, что строгие определения можно надеяться получить на последних этапах построения теории. Мы же сейчас находимся чаще всего на первых этапах. Поэтому, не давая определения понятиям «естественная классификация» и «естественная диагностика», обсудим, как проверить на «естественность» классификацию (набор диагностических классов), полученную расчетным путем.

Можно выделить два критерия «естественности», по поводу которых имеется относительное согласие:

А. Естественная классификация должна быть реальной, соответствующей действительному миру, лишенной внесенного исследователем субъективизма;

Б. Естественная классификация должна быть важной или с научной точки зрения (давать возможность прогноза, предсказания новых свойств, сжатия информации и т.д.), или с практической.

Пусть классификация проводится на основе информации об объектах, представленной в виде матрицы «объект-признак» или матрицы попарных расстояний (мер близости). Пусть алгоритм классификации дал разбиение на кластеры. Как можно получить доводы в пользу естественности этой классификации? Например, уверенность в том, что она — закон природы, может появиться только в результате ее длительного изучения и практического применения. Это соображение относится и к другим из перечисленных выше критериев, в частности к Б (важности). Сосредоточимся на критерии А (реальности).

Понятие «реальности» кластера требует специального обсуждения. (оно начато в работе [9]). Рассмотрим существо различий между понятиями «классификация» и «группировка». Пусть, к примеру, необходимо деревья, растущие в определенной местности, разбить на группы находящихся рядом друг с другом. Ясна интуитивная разница между несколькими отдельными рощами, далеко отстоящими друг от друга и разделенными полями, и сплошным лесом, разбитым просеками на квадраты с целью лесоустройства.

Однако формально определить эту разницу столь же сложно, как определить понятие «куча зерен», чем занимались еще в Древней Греции. Ясно, что одно зерно не составляет кучи, два зерна не составляют кучи… Если к тому, что не составляет кучи, добавить еще одно зерно, то куча не получится. Значит — по принципу математической индукции — никакое количество зерен не составляет кучи. Но ясно, что миллиард зерен — большая куча зерен — подсчитайте объем!

Переформулируем сказанное в терминах «кластер-анализа» и «методов группировки». Выделенные с помощью первого подхода кластеры реальны, а потому могут рассматриваться как кандидаты в «естественные». Группировка дает «искусственные» классы, которые не могут быть «естественными».

Выборку из унимодального распределения можно, видимо, рассматривать как «естественный», «реальный» кластер. Применим к ней какой-либо алгоритм классификации («средней связи», «ближайшего соседа» и т.п.). Он даст какое-то разбиение на классы, которые, разумеется, не являются «реальными», поскольку отражают прежде всего свойства алгоритма, а не исходных данных. Как отличить такую ситуацию от противоположной, когда имеются реальные кластеры и алгоритм классификации более или менее точно их выделяет? Как известно, «критерий истины — практика», но слишком много времени необходимо для применения подобного критерия. Поэтому представляет интерес критерий, оценивающий «реальность» выделяемых с помощью алгоритма классификации кластеров одновременно с его применением.

Такой показатель существует — это критерий устойчивости. Устойчивость — понятие широкое. Общая схема формулирования и изучения проблем устойчивости рассмотрена в главе 1.4. В частности, поскольку значения признаков всегда измеряются с погрешностями, то «реальное» разбиение должно быть устойчиво (т.е. не меняться или меняться слабо) при малых отклонениях исходных данных. Алгоритмов классификации существует бесконечно много, и «реальное» разбиение должно быть устойчиво по отношению к переходу к другому алгоритму. Другими словами, если «реальное» разбиение на классы возможно, то оно находится с помощью любого алгоритма автоматической классификация. Следовательно, критерием естественности классификации может служить совпадение результатов работы двух достаточно различающихся алгоритмов, например «ближайшего соседа» и «дальнего соседа».

Выше рассмотрены два типа «глобальных» критериев «естественности классификации», касающихся разбиения в целом. «Локальные» критерии относятся к отдельным кластерам. Простейшая постановка такова: достаточно ли однородны два кластера (две совокупности) для их объединения? Если объединение возможно, то кластеры не являются «естественными». Преимущество этой постановки в том, что она допускает применение статистических критериев однородности двух выборок. В одномерном случае (классификация по одному признаку) разработано большое число подобных критериев — Крамера — Уэлча, Смирнова, омега-квадрат (Лемана — Розенблатта), Вилкоксона, Ван-дер-Вардена, Лорда, Стьюдента и др. (см. главу 3.1 и справочник [4]). Имеются критерии и для многомерных данных. Для одного из видов объектов нечисловой природы — люсианов — статистические методы выделения «реальных» кластеров развиты в работе [10].

Что касается глобальных критериев, то для изучения устойчивости по отношению к малым отклонениям исходных данных естественно использовать метод статистических испытаний и проводить расчеты по «возмущенным» данным. Некоторые теоретические утверждения, касающиеся влияния «возмущений» на кластеры различных типов, получены в работе [9].

Опишем практический опыт реализации анализа устойчивости. Несколько алгоритмов классификации были применены к данным, полученным при проведении маркетинга образовательных услуг и приведенным в работе [11]. Для анализа данных были использованы широко известные алгоритмы «ближайшего соседа», «дальнего соседа» и алгоритм кластер-анализа из работы [12]. С содержательной точки зрения полученные разбиения отличались мало. Поэтому есть основания считать, что с помощью этих алгоритмов действительно выявлена «реальная» структура данных.

Идея устойчивости как критерия «реальности» иногда реализуется неадекватно. Так, для однопараметрических алгоритмов иногда предлагают выделять разбиения, которым соответствуют наибольшие интервалы устойчивости по параметру, т.е. наибольшие приращения параметра между очередными объединениями кластеров. Для данных работы [11] это предложение не дало полезных результатов — были получены различные разбиения: три алгоритма — три разбиения. И с теоретической точки зрения предложение этого специалиста несостоятельно. Покажем это.

Действительно, рассмотрим алгоритм «ближайшего соседа», использующий меру близости $d(x,y)$, и однопараметрическое семейство алгоритмов с мерой близости $d^a(x,y)$, $a>0$, также являющихся алгоритмами «ближайшего соседа». Тогда дендрограммы, полученные с помощью этих алгоритмов, совпадают при всех a, поскольку при их реализации происходит лишь сравнение мер близости между объектами. Другими словами, дендрограмма, полученная с помощью алгоритма «ближайшего соседа», является адекватной в порядковой шкале (измерения меры близости $d(x,y)$), т.е. сохраняется при любом строго возрастающем преобразовании этой меры. Однако выделенные по обсуждаемому методу «устойчивые разбиения» меняются. В частности, при достаточно большом a «наиболее объективным» в соответствии с рассматриваемым предложением будет, как нетрудно показать, разбиение на два кластера! Таким образом, разбиение, выдвинутое им как «устойчивое», на самом деле оказывается весьма неустойчивым.

Рассмотрим с позиций прикладной статистики несколько конкретных вопросов теории классификации.

Вероятностная теория кластер-анализа. Как и для прочих статистических методов, свойства алгоритмов кластер-анализа необходимо изучать на вероятностных моделях. Это касается, например, условий естественного объединения двух кластеров.

Вероятностные постановки нужно применять, в частности, при перенесении результатов, полученных по выборке, на генеральную совокупность. Вероятностная теория кластер-анализа и методов группировки различна для исходных данных типа таблиц «объект Ч признак» и матриц сходства. Для первых параметрическая вероятностно-статистическая теория называется «расщеплением смесей». Непараметрическая теория основана на непараметрических оценках плотностей вероятностей и их мод. Основные результаты, связанные с непараметрическими оценками плотности, обсуждались в разделе 2.1.

Если исходные данные — матрица сходства $\|d(x,y)\|$, то необходимо признать, что развитой вероятностно-статистической теории пока нет. Подходы к ее построению намечены в работе [9]. Одна из основных проблем — проверка «реальности» кластера, его объективного существования независимо от расчетов исследователя. Проблема «реальности» кластера давно обсуждается специалистами различных областей. Типичное рассуждение таково. Предположим, что результаты наблюдений можно рассматривать как выборку из некоторого распределения с монотонно убывающей плотностью при увеличении расстояния от некоторого центра. Примененный к подобным данным какой-либо алгоритм кластер-анализа порождает некоторое разбиение. Ясно, что оно — чисто формальное, поскольку выделенным таксонам (кластерам) не соответствуют никакие «реальные» классы. Другими словами, задача кластер-анализа не имеет решения, а алгоритм дает лишь группировку. При обработке реальных данных мы не знаем вида плотности. Проблема состоит в том, чтобы определить, каков результат работы алгоритма (реальные кластеры или формальные группы).

Частный случай этой проблемы — проверка обоснованности объединения двух кластеров, которые мы рассматриваем как два множества объектов, а именно множества $\{a_1, a_2, \ldots, a_k\}$ и $\{b_1, b_2, \ldots, b_m\}$. Пусть, например, используется алгоритм типа «Дендрограмма». Естественной представляется следующая идея. Пусть есть две совокупности мер близости. Одна — меры близости между объектами, лежащими внутри одного кластера, т.е. $d(a_i,a_j)$, $1\le i<j\le k$, $d(b_\alpha,b_\beta)$, $1\le\alpha<\beta\le m$. Другая совокупность — меры близости между объектами, лежащими в разных кластерах, т.е. $d(a_i,b_\alpha)$, $1\le i\le k$, $1\le\alpha\le m$. Эти две совокупности мер близости предлагается рассматривать как независимые выборки и проверять гипотезу о совпадении их функций распределения. Если гипотеза не отвергается, объединение кластеров считается обоснованным; в противном случае — объединять нельзя, алгоритм прекращает работу.

В рассматриваемом подходе есть две некорректности (см. также работу [9, разд. 4]). Во-первых, меры близости не являются независимыми случайными величинами. Во-вторых, не учитывается, что объединяются не заранее фиксированные кластеры (с детерминированным составом), а полученные в результате работы некоторого алгоритма, и их состав (в частности, количество элементов) оказывается случайным. От первой из этих некорректностей можно частично избавиться. Справедливо следующее утверждение.

Теорема 1. Пусть a_1, a_2, ..., a_k, b_1, b_2, ..., b_m — независимые одинаково распределенные случайные величины (со значениями в произвольном пространстве). Пусть случайная величина $d(a_1,a_2)$ имеет все моменты. Тогда при $k,m \to \infty$ распределение статистики:

$$\frac{8\sqrt{3}U - 3(k+m)(k+m-1)(k(k+1)+m(m+1))}{2(k+m)\sqrt{km(k^2+m^2)}}$$

(где U — сумма рангов элементов первой выборки в объединенной выборке; первая выборка составлена из внутрикластерных расстояний (мер близости) $d(a_i,a_j)$, $1 \leq i < j \leq k$, и $d(b_\alpha,b_\beta)$, $1 \leq \alpha < \beta \leq m$, а вторая — из межкластерных расстояний $d(a_i,b_\alpha)$, $1 \leq i \leq k$, $1 \leq \alpha \leq m$) сходится к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1.

На основе теоремы 1 очевидным образом формулируется правило проверки обоснованности объединения двух кластеров. Другими словами, мы проверяем статистическую гипотезу, согласно которой объединение двух кластеров образует однородную совокупность. Если величина U слишком мала, статистическая гипотеза однородности отклоняется (на заданном уровне значимости), и возможность объединения отбрасывается. Таким образом, хотя расстояния между объектами в кластерах зависимы, но эта зависимость слаба, и доказана математическая теорема о допустимости применения критерия Вилкоксона для проверки возможности объединения кластеров.

О вычислительной сходимости алгоритмов кластер-анализа. Алгоритмы кластер-анализа и группировки зачастую являются итерационными. Например, формулируется правило улучшения решения задачи кластер-анализа шаг за шагом, но момент остановки вычислений не обсуждается. Примером является известный алгоритм «Форель», в котором постепенно улучшается положение центра кластера. В этом алгоритме на каждом шаге строится шар определенного заранее радиуса, выделяются элементы кластеризуемой совокупности, попадающие в этот шар, и новый центр кластера строится как центр тяжести выделенных элементов. При анализе алгоритма «Форель» возникает проблема: завершится ли процесс улучшения положения центра кластера через конечное число шагов или же он может быть бесконечным. Она получила название «проблема остановки». Для широкого класса так называемых «эталонных алгоритмов» проблема остановки была решена в работе [9]: процесс улучшения остановится через конечное число шагов.

Отметим, что алгоритмы кластер-анализа могут быть модифицированы разнообразными способами. Например, описывая алгоритм «Форель» в стиле статистики объектов нечисловой природы, заметим, что вычисление центра тяжести для совокупности многомерных точек — это нахождение эмпирического среднего для меры близости, равной квадрату евклидова расстояния. Если взять более естественную меру близости — само евклидово расстояние, то получим алгоритм кластер-анализа «Медиана», отличающийся от «Форели» тем, что новый центр строится не с помощью средних арифметических координат элементов, попавших в кластер, а с помощью медиан.

Проблема остановки возникает не только при построении диагностических классов. Она принципиально важна, в частности, и при оценивании параметров вероятностных распределений методом максимального правдоподобия. Обычно не представляет большого труда выписать систему уравнений максимального правдоподобия и предложить решать ее каким-либо численным методом. Однако остаются вопросы: когда остановиться, сколько итераций сделать, какая точность оценивания будет при этом достигнута? Общий ответ, видимо, невозможно найти, но обычно нет ответа и для конкретных семейств распределения вероятностей. Именно поэтому нет оснований рекомендовать решать системы уравнений максимального правдоподобия. Вместо них целесообразно использовать т.н. одношаговые оценки (подробнее см. об этих оценках главу 2.2). Эти оценки задаются конечными формулами, но асимптотически столь же хороши (на профессиональном языке — эффективны), как и оценки максимального правдоподобия.

О сравнении алгоритмов диагностики по результатам обработки реальных данных. Перейдем к этапу применения диагностических правил, когда классы, к одному из которых нужно отнести вновь поступающий объект, уже выделены.

В прикладных исследованиях применяют различные методы дискриминантного анализа, основанные на вероятностно-статистических моделях, а также с ними не связанные, т.е. эвристические, использующие детерминированные методы анализа данных. Независимо от «происхождения», каждый подобный алгоритм должен быть исследован как на параметрических и непараметрических вероятностно-статистических моделях порождения данных, так и на различных массивах реальных данных. Цель исследования — выбор наилучшего алгоритма в определенной области применения, включение его в стандартные программные продукты, методические материалы, учебные программы и пособия. Но для этого надо уметь сравнивать алгоритмы по качеству. Как это делать?

Часто используют такой показатель качества алгоритма диагностики, как «вероятность правильной классификации» (при обработке конкретных данных — «частота правильной классификации»). Чуть ниже мы покажем, что этот показатель качества некорректен, а потому пользоваться им не рекомендуется. Целесообразно применять другой показатель качества алгоритма диагностики — оценку специального вида т.н. «расстояния Махаланобиса» между классами. Изложение проведем на примере разработки программного продукта для специалистов по диагностике материалов. Прообразом является диалоговая система «АРМ материаловеда», разработанная Институтом высоких статистических технологий и эконометрики для ВНИИ эластомерных материалов.

При построении информационно-исследовательской системы диагностики материалов (ИИСДМ) возникает задача сравнения прогностических правил «по силе». Прогностическое правило — это алгоритм, позволяющий по характеристикам материала прогнозировать его свойства. Если прогноз дихотомичен («есть» или «нет»), то правило является алгоритмом диагностики, при котором материал относится к одному из двух классов. Ясно, что случай нескольких классов может быть сведен к конечной последовательности выбора между двумя классами.

Прогностические правила могут быть извлечены из научно-технической литературы и практики. Каждое из них обычно формулируется в терминах небольшого числа признаков, но наборы признаков сильно меняются от правила к правилу. Поскольку в ИИСДМ должно фиксироваться лишь ограниченное число признаков, то возникает проблема их отбора. Следует отбирать лишь те их них, которые входят в наборы, дающие наиболее «надежные» прогнозы. Для придания точного смысла термину «надежный» необходимо иметь способ сравнения алгоритмов диагностики по прогностической «силе».

Результаты обработки реальных данных с помощью некоторого алгоритма диагностики в рассматриваемом случае двух классов описываются долями: правильной диагностики в первом классе κ; правильной диагностики во втором классе λ; долями классов в объединенной совокупности π_i, $i = 1.2$; $\pi_1 + \pi_2 = 1$.

При изучении качества алгоритмов классификации их сравнивают по результатам дискриминации вновь поступающей контрольной выборки. Именно по контрольной выборке определяются величины $\kappa, \lambda, \pi_1, \pi_2$. Однако иногда вместо контрольной используют обучающую выборку, т.е. указанные величины определяются ретроспективно, в результате анализа уже имеющихся данных. Обычно это связано с трудоемкостью получения данных. Тогда κ и λ зависимы. Однако в случае, когда решающее правило основано на использовании дискриминантной поверхности, параметры которой оцениваются по обучающим выборкам, величины κ и λ асимптотически (при безграничном росте объемов выборок) независимы [9], что позволяет использовать приводимые ниже результаты и в этом случае.

Нередко как показатель качества алгоритма диагностики (прогностической «силы») используют долю правильной диагностики:

$$\mu = \pi_1 \kappa + \pi_2 \lambda.$$

Однако показатель μ определяется, в частности, через характеристики π_1 и π_2, частично заданные исследователем (например, на них влияет тактика отбора образцов для изучения). В аналогичной медицинской задаче величина μ оказалась больше для тривиального прогноза, согласно которому у всех больных течение заболевания будет благоприятно. Тривиальный прогноз сравнивался с алгоритмом выделения больных с прогнозируемым тяжелым течением заболевания. Он был разработан группы под руководством академика АН СССР И.М. Гельфанда. Применение этого алгоритма с медицинской точки зрения вполне оправдано [13].

Другими словами, по доле правильной классификации алгоритм академика И.М. Гельфанда оказался хуже тривиального — объявить всех больных легкими, не требующими специального наблюдения. Этот вывод очевидно нелеп. И причина появления нелепости вполне понятна. Хотя доля тяжелых больных невелика, но смертельные исходы сосредоточены именно в этой группе больных. Поэтому целесообразна гипердиагностика — рациональнее часть легких больных объявить тяжелыми, чем сделать ошибку в противоположную сторону. Применение теории статистических решений в рассматриваемой постановке вряд ли возможно, поскольку оценить количественно потери от смерти больного нельзя по этическим соображениям. Поэтому, на наш взгляд, долю правильной диагностики μ нецелесообразно использовать как показатель качества алгоритма диагностики.

Применение теории статистических решений требует знания потерь от ошибочной диагностики, а в большинстве научно-технических и экономических задач определить потери, как уже отмечалось, сложно. В частности, из-за необходимости оценивать человеческую жизнь в денежных единицах. По этическим соображениям это, на наш взгляд, недопустимо. Сказанное не означает отрицания пользы страхования, но, очевидно, страховые выплаты следует рассматривать лишь как способ первоначального смягчения потерь от утраты близких.

Для выявления информативного набора признаков целесообразно использовать *метод пересчета на модель линейного дискриминантного анализа,* согласно которому статистической оценкой прогностической «силы» является:

$$\delta^* = \Phi(d^*/2),\ \ d^* = \Phi^{-1}(\kappa) + \Phi^{-1}(\lambda),$$

где $\Phi(x)$ — функция стандартного нормального распределения вероятностей с математическим ожиданием 0 и дисперсией 1, а $\Phi^{-1}(y)$ — обратная ей функция.

Пример 1. Если доли правильной классификации $\kappa = 0{,}90$ и $\lambda = 0{,}80$, то $\Phi^{-1}(\kappa) = 1{,}28$ и $\Phi^{-1}(\lambda) = 0{,}84$, откуда $d^* = 2{,}12$ и прогностическая сила $\delta^* = \Phi^{-1}(1{,}06) = 0{,}86$. При этом доля правильной классификации μ может принимать любые значения между 0,80 и 0,90, в зависимости от доли элементов того или иного класса среди анализируемых данных.

Если классы описываются выборками из многомерных нормальных совокупностей с одинаковыми матрицами ковариаций, а для классификации применяется классический линейный дискриминантный анализ Р. Фишера, то величина d^* представляет собой состоятельную статистическую оценку так называемого расстояния Махаланобиса между рассматриваемыми двумя совокупностями (конкретный вид этого расстояния сейчас не имеет значения), независимо от порогового значения, определяющего конкретное решающее правило. В общем случае показатель δ^* вводится как эвристический.

Пусть алгоритм классификации применялся к совокупности, состоящей из m объектов первого класса и n объектов второго класса.

Теорема 2. Пусть $m, n \to \infty$. Тогда для всех x:

$$P\left\{\frac{\delta^* - \delta}{A(\kappa,\lambda)} < x\right\} \to \Phi(x),$$

где δ — истинная «прогностическая сила» алгоритма диагностики; δ^* — ее эмпирическая оценка,

$$A^2(\kappa,\lambda) = \frac{1}{4}\left\{\left[\frac{\varphi(d^*/2)}{\varphi(\Phi^{-1}(\kappa))}\right]^2 \frac{\kappa(1-\kappa)}{m} + \left[\frac{\varphi(d^*/2)}{\varphi(\Phi^{-1}(\lambda))}\right]^2 \frac{\lambda(1-\lambda)}{n}\right\};$$

$\varphi(x) = \Phi'(x))$ — плотность стандартного нормального распределения вероятностей с математическим ожиданием 0 и дисперсией 1.

С помощью теоремы 2 по κ и λ обычным образом определяют доверительные границы для «прогностической силы» δ .

Пример 2. В условиях примера 1 при m = n = 100 найдем асимптотическое среднее квадратическое отклонение A(0,90; 0,80).

Поскольку $\varphi(\Phi^{-1}(\kappa)) = \varphi(1{,}28) = 0{,}176$, $\varphi(\Phi^{-1}(\lambda)) = \varphi(0{,}84) = 0{,}280$, $\varphi(d^*/2) = \varphi(1{,}06) = 0{,}227$, то подставляя в выражение для A^2 численные значения, получаем, что:

$$A^2(0{,}90;\ 0{,}80) = \frac{0{,}0372}{m} + \frac{0{,}0265}{n}$$

(численные значения плотности стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1 и функции, обратной к функции этого распределения, можно было взять, например, из справочника [4]).

При $m = n = 100$ имеем А(0,90; 0,80) = 0,0252. При доверительной вероятности $\gamma = 0{,}95$ имеем $u(0{,}95) = \Phi^{-1}(1{,}0{,}975) = 1{,}96$, а потому нижняя доверительная граница для прогностической силы δ есть $\delta_Н$ = 0,86 — 1,96 Ч 0,0252 = 0,81, а верхняя доверительная граница такова: $\delta_В$ = 0,86 + 1,96 Ч 0,0252 = 0,91. Аналогичный расчет при $m = n = 1000$ дает $\delta_Н = 0{,}845$, $\delta_В = 0{,}875$.

Как проверить обоснованность пересчета на модель линейного дискриминантного анализа? Допустим, что классификация состоит в вычислении некоторого прогностического индекса y и сравнении его с заданным порогом c. Объект относят к первому классу, если $y \leq c$, ко второму, если $y > c$. Прогностический индекс — это обычно линейная функция от характеристик рассматриваемых объектов. Другими словами, от координат векторов, описывающих объекты.

Возьмем два значения порога c_1 и c_2. Если пересчет на модель линейного дискриминантного анализа обоснован, то , как можно показать, «прогностические силы» для обоих правил совпадают: $\delta(c_1) = \delta(c_2)$. Выполнение этого равенства можно проверить как статистическую гипотезу.

Пусть κ_1 — доля объектов первого класса, для которых $y \leq c_1$, а κ_2 — доля объектов первого класса, для которых $c_1 < y \leq c_2$. Аналогично пусть λ_2 — доля объектов второго класса, для которых $c_1 < y \leq c_2$, а λ_3 — доля объектов второго класса, для которых $y > c_2$. Тогда можно рассчитать две оценки одного и того же расстояния Махаланобиса. Они имеют вид:

$$d^*(c_1) = \Phi^{-1}(\kappa_1) + \Phi^{-1}(\lambda_2 + \lambda_3), \quad d^*(c_2) = \Phi^{-1}(\kappa_1 + \kappa_2) + \Phi^{-1}(\lambda_3).$$

Теорема 3. Если истинные прогностические силы двух правил диагностики совпадают, $\delta(c_1) = \delta(c_2)$, то при $m \to \infty, n \to \infty$ при всех *x:*

$$P\left\{\frac{d^*(c_1) - d^*(c_2)}{B} < x\right\} \to \Phi(x),$$

где

$$B^2 = \frac{1}{m}T(\kappa_1;\kappa_2) + \frac{1}{n}T(\lambda_3;\lambda_2);$$

$$T(x;y) = \frac{x(1-x)}{\varphi^2(\Phi^{-1}(x))} + \frac{(x+y)(1-x-y)}{\varphi^2(\Phi^{-1}(x+y))} - \frac{2x(1-x-y)}{\varphi(\Phi^{-1}(x))\varphi(\Phi^{-1}(x+y))}.$$

Из теоремы 3 вытекает метод проверки рассматриваемой гипотезы: при выполнении неравенства:

$$\left|\frac{d^*(c_1) - d^*(c_2)}{B}\right| \le \Phi^{-1}(1 - \frac{\alpha}{2})$$

она принимается на уровне значимости, асимптотически равном α, в противном случае — отвергается.

Пример 3. Пусть данные примеров 1 и 2 соответствуют порогу c_1. Пусть порогу c_2 соответствуют $\kappa' = 0{,}95$ и $\lambda' = 0{,}70$. Тогда в обозначениях теоремы 3 $\kappa_1 = 0{,}90$, $\kappa_2 = 0{,}05$, $\lambda_2 = 0{,}10$, $\lambda_3 = 0{,}70$. Далее $d^*(c_1) = 2{,}12$ (пример 1), $d^*(c_2) = 2{,}17$, $T(\kappa_1, \kappa_2) = 2{,}22$, $T(\lambda_3, \lambda_2) = 0{,}89$. Гипотеза о совпадении прогностических сил на двух порогах принимается на уровне значимости $\alpha = 0{,}05$ тогда и только тогда, когда

$$\frac{0{,}05^2}{\frac{2{,}22}{m} + \frac{0{,}89}{n}} \le 1{,}96^2,$$

т.е. когда

$$\frac{2{,}22}{m} + \frac{0{,}89}{n} \ge 0{,}00065.$$

Так, гипотеза принимается при $m = n = 1\,000$ и отвергается при $m = n = 5\,000$.

Подходы к построению прогностических правил. Для решения задач диагностики используют два подхода — параметрический и непараметрический. Первый из них обычно основан на использовании того или иного индекса и сравнения его с порогом. Индекс может быть построен по статистическим данным, например, как в уже упомянутом линейном дискриминантном анализе Фишера. Часто индекс представляет собой линейную функцию от характеристик, выбранных специалистами предметной области, коэффициенты которой подбирают эмпирически. Непараметрический подход связан с леммой Неймана-Пирсона в математической статистике и с теорией статистических решений. Он опирается на использование непараметрических оценок плотностей распределений вероятностей, описывающих диагностические классы.

Обсудим ситуацию подробнее. Математические методы диагностики, как и статистические методы в целом, делятся на параметрические и непараметрические. Первые основаны на предположении, что классы описываются распределениями из некоторых параметрических семейств. Обычно рассматривают многомерные нормальные распределения, при этом зачастую принимают гипотезу о том, что ковариационные матрицы для различных классов совпадают. Именно в таких предположениях сформулирован классический дискриминантный анализ Фишера. Как известно, обычно нет оснований считать, что наблюдения извлечены из нормального распределения.

Поэтому более корректными, чем параметрические, являются непараметрические методы диагностики. Исходная идея таких методов основана на лемме Неймана-Пирсона, входящей в стандартный курс математической статистики. Согласно этой лемме решение об отнесении вновь поступающего объекта (сигнала, наблюдения и др.) к одному из двух классов принимается на основе отношения плотностей $f(x)/g(x)$, где $f(x)$ — плотность распределения, соответствующая первому классу, а $g(x)$ — плотность распределения, соответствующая второму классу. Если плотности распределения неизвестны, то применяют их непараметрические оценки, построенные по обучающим выборкам. Пусть обучающая выборка объектов из первого класса состоит из n элементов, а обучающая выборка для второго класса — из m объектов. Тогда рассчитывают значения непараметрических оценок плотностей $f_n(x)$ и $g_m(x)$ для первого и второго классов соответственно, а диагностическое решение принимают по их отношению. Таким образом, для решения задачи диагностики достаточно научиться строить непараметрические оценки плотности для выборок объектов произвольной природы.

Методы построения непараметрических оценок плотности распределения вероятностей в пространствах произвольной природы рассмотрены в главе 2.4.

6.5. Методы снижения размерности

В многомерном статистическом анализе каждый объект описывается вектором, размерность которого произвольна (но одна и та же для всех объектов). Однако человек может непосредственно воспринимать лишь числовые данные или точки на плоскости. Анализировать скопления точек в трехмерном пространстве уже гораздо труднее. Непосредственное восприятие данных более высокой размерности невозможно. Поэтому вполне естественным является желание перейти от многомерной выборки к данным небольшой размерности, чтобы «на них можно было посмотреть».

Кроме стремления к наглядности, есть и другие мотивы для снижения размерности. Те факторы, от которых интересующая исследователя переменная не зависит, лишь мешают статистическому анализу. Во-первых, на сбор информации о них расходуются ресурсы. Во-вторых, как можно доказать, их включение в анализ ухудшает свойства статистических процедур (в частности, увеличивает дисперсию оценок параметров и характеристик распределений). Поэтому желательно избавиться от таких факторов.

Обсудим с точки зрения снижения размерности пример использования регрессионного анализа для прогнозирования объема продаж, рассмотренный в разделе 6.3. Во-первых, в этом примере удалось сократить число независимых переменных с 17 до 12. Во-вторых, удалось сконструировать новый фактор — линейную функцию от 12 упомянутых факторов, которая лучше всех иных линейных комбинаций факторов прогнозирует объем продаж. Поэтому можно сказать, что в результате размерность задачи уменьшилась с 18 до 2. А именно, остался один независимый фактор (приведенная в подразделе 3.2.3 линейная комбинация) и один зависимый — объем продаж.

При анализе многомерных данных обычно рассматривают не одну, а множество задач, в частности, по-разному выбирая независимые и зависимые переменные. Поэтому рассмотрим задачу снижения размерности в следующей формулировке. Дана многомерная выборка. Требуется перейти от нее к совокупности векторов меньшей размерности, максимально сохранив структуру исходных данных, по возможности не теряя информации, содержащихся в данных. Задача конкретизируется в рамках каждого конкретного метода снижения размерности.

Метод главных компонент является одним из наиболее часто используемых методов снижения размерности. Основная его идея состоит в последовательном выявлении направлений, в которых данные имеют наибольший разброс. Пусть выборка состоит из векторов, одинаково распределенных с вектором $X = (x(1), x(2), \ldots , x(n))$. Рассмотрим линейные комбинации:

$$Y(\lambda(1), \lambda(2), \ldots, \lambda(n)) = \lambda(1)x(1) + \lambda(2)x(2) + \ldots + \lambda(n)x(n),$$

где

$$\lambda^2(1) + \lambda^2(2) + \ldots + \lambda^2(n) = 1.$$

Здесь вектор $\lambda = (\lambda(1), \lambda(2), \ldots, \lambda(n))$ лежит на единичной сфере в n-мерном пространстве.

В методе главных компонент прежде всего находят направление максимального разброса, т.е. такое λ, при котором достигает максимума дисперсия случайной величины $Y(\lambda) = Y(\lambda(1), \lambda(2), \ldots, \lambda(n))$. Тогда вектор λ задает первую главную компоненту, а величина $Y(\lambda)$ является проекцией случайного вектора X на ось первой главной компоненты.

Затем, выражаясь терминами линейной алгебры, рассматривают гиперплоскость в n-мерном пространстве, перпендикулярную первой главной компоненте, и проектируют на эту гиперплоскость все элементы выборки. Размерность гиперплоскость на 1 меньше, чем размерность исходного пространства.

В рассматриваемой гиперплоскости процедура повторяется. В ней находят направление наибольшего разброса, т.е. вторую главную компоненту. Затем выделяют гиперплоскость, перпендикулярную первым двум главным компонентам. Ее размерность на 2 меньше, чем размерность исходного пространства. Далее — следующая итерация.

С точки зрения линейной алгебры речь идет о построении нового базиса в n-мерном пространстве, ортами которого служат главные компоненты.

Дисперсия, соответствующая каждой новой главной компоненте, меньше, чем для предыдущей. Обычно останавливаются, когда она меньше заданного порога. Если отобрано k главных компонент, то это означает, что от n-мерного пространства удалось перейти к k-мерному, т.е. сократить размерность с n-до k, практически не исказив структуру исходных данных.

Для визуального анализа данных часто используют проекции исходных векторов на плоскость первых двух главных компонент. Обычно хорошо видна структура данных, выделяются компактные кластеры объектов и отдельно выделяющиеся вектора.

Метод главных компонент является одним из методов **факторного анализа** [14]. Различные алгоритмы факторного анализа объединены тем, что во всех них происходит переход к новому базису в исходном n-мерном пространстве. Важным является понятие «нагрузка фактора», применяемое для описания роли исходного фактора (переменной) в формировании определенного вектора из нового базиса.

Новая идея по сравнению с методом главных компонент состоит в том, что на основе нагрузок происходит разбиение факторов на группы. В одну группу объединяются факторы, имеющие сходное влияние на элементы нового базиса. Затем из каждой группы рекомендуется оставить одного представителя. Иногда вместо выбора представителя расчетным путем формируется новый фактор, являющийся центральным для рассматриваемой группы. Снижение размерности происходит при переходе к системе факторов, являющихся представителями групп. Остальные факторы отбрасываются.

Описанная процедура может быть осуществлена не только с помощью факторного анализа. Речь идет о кластер-анализе признаков (факторов, переменных). Для разбиения признаков на группы можно применять различные алгоритмы кластер-анализа. Достаточно ввести расстояние (меру близости, показатель различия) между признаками. Пусть X и Y — два признака. Различие $d(X,Y)$ между ними можно измерять с помощью выборочных коэффициентов корреляции:

$$d_1(X,Y) = 1 - |r_n(X,Y)|,\ d_2(X,Y) = 1 - |\rho_n(X,Y)|,$$

где $r_n(X,Y)$ — выборочный линейный коэффициент корреляции Пирсона, $\rho_n(X,Y)$ — выборочный коэффициент ранговой корреляции Спирмена.

Многомерное шкалирование. На использовании расстояний (мер близости, показателей различия) $d(X,Y)$ между признаками X и Y основан обширный класс методов многомерного шкалирования [15, 16]. Основная идея этого класса методов состоит в представлении каждого объекта точкой геометрического пространства (обычно размерности 1, 2 или 3), координатами которой служат значения скрытых (латентных) факторов, в совокупности достаточно адекватно описывающих объект. При этом отношения между объектами заменяются отношениями между точками — их представителями. Так, данные о сходстве объектов — расстояниями между точками, данные о превосходстве — взаимным расположением точек [17].

В практике используется ряд различных моделей многомерного шкалирования. Во всех них встает проблема оценки истинной размерности факторного пространства. Рассмотрим эту проблему на примере обработки данных о сходстве объектов с помощью метрического шкалирования.

Пусть имеется n объектов $O(1)$, $O(2)$, …, $O(n)$, для каждой пары объектов $O(i)$, $O(j)$ задана мера их сходства $s(i,j)$. Считаем, что всегда $s(i,j) = s(j,i)$. Происхождение чисел $s(i,j)$ не имеет значения для описания работы алгоритма. Они могли быть получены либо непосредственным измерением, либо с использованием экспертов, либо путем вычисления по совокупности описательных характеристик, либо как-то иначе.

В евклидовом пространстве рассматриваемые n объектов должны быть представлены конфигурацией n точек, причем в качестве меры близости точек-представителей выступает евклидово расстояние $d(i,j)$ между соответствующими точками. Степень соответствия между совокупностью объектов и совокупностью представляющих их точек определяется путем сопоставления матриц сходства $\|s(i,j)\|$ и расстояний $\|d(i,j)\|$. Метрический функционал сходства имеет вид:

$$S = \sum_{i<j} | s(i,j) - d(i,j) |^2 .$$

Геометрическую конфигурацию надо выбирать так, чтобы функционал S достигал своего наименьшего значения [17].

Замечание. В неметрическом шкалировании вместо близости самих мер близости и расстояний рассматривается близость упорядочений на множестве мер близости и множестве соответствующих расстояний. Вместо функционала S используются аналоги ранговых коэффициентов корреляции Спирмена и Кендалла. Другими словами, неметрическое шкалирование исходит из предположения, что меры близости измерены в порядковой шкале (см. главу 11.3).

Пусть евклидово пространство имеет размерность m. Рассмотрим минимум среднего квадрата ошибки:

$$\alpha_m = \frac{2}{n(n-1)} \min S ,$$

где минимум берется по всем возможным конфигурациям n точек в m-мерном евклидовом пространстве. Можно показать, что рассматриваемый минимум достигается на некоторой конфигурации. Ясно, что при росте m величина α_m монотонно убывает (точнее, не возрастает). Можно показать, что при $m \geq n - 1$ она равна 0 (если $s(i,j)$ — метрика). Для увеличения возможностей содержательной интерпретации желательно действовать в пространстве возможно меньшей размерности. При этом, однако, размерность необходимо выбрать так, чтобы точки представляли объекты без больших искажений. Возникает вопрос: как рационально выбирать размерность, т.е. натуральное число m?

В рамках детерминированного анализа данных обоснованного ответа на этот вопрос, видимо, нет. Следовательно, необходимо изучить поведение α_m в тех или иных вероятностных моделях. Если меры близости $s(i,j)$ являются случайными величинами, распределение которых зависит от «истинной размерности» m_0 (и, возможно, от каких-либо еще параметров), то можно в классическом математико-статистическом стиле ставить задачу оценки m_0, искать состоятельные оценки и т.д.

Начнем строить вероятностные модели. Примем, что объекты представляют собой точки в евклидовом пространстве размерности k, где k достаточно велико. То, что «истинная размерность» равна m_0, означает, что все эти точки лежат на гиперплоскости размерности m_0. Примем для определенности, что совокупность рассматриваемых точек представляет собой выборку из кругового нормального распределения с дисперсией $\sigma^2(0)$. Это означает, что объекты $O(1), O(2), \ldots, O(n)$ являются независимыми в совокупности случайными векторами, каждый из которых строится как $\zeta(1)e(1) + \zeta(2)e(2) + \ldots + \zeta(m_0)e(m_0)$, где $e(1), e(2), \ldots, e(m_0)$ — ортонормальный базис в подпространстве размерности m_0, в котором лежат рассматриваемые точки, а $\zeta(1), \zeta(2), \ldots, \zeta(m_0)$ — независимые в совокупности одномерные нормальные случайные величины с математическим ожиданием) и дисперсией $\sigma^2(0)$.

Рассмотрим две модели получения мер близости $s(i,j)$. В первой из них $s(i,j)$ отличаются от евклидова расстояния между соответствующими точками из-за того, что точки известны с искажениями. Пусть $c(1), c(2), \ldots, c(n)$ — рассматриваемые точки. Тогда:

$$s(i,j) = d(c(i) + \varepsilon(i), c(j) + \varepsilon(j)),\ i,j = 1, 2, \ldots, n,$$

где d — евклидово расстояние между точками в k-мерном пространстве, вектора $\varepsilon(1), \varepsilon(2), \ldots, \varepsilon(n)$ представляют собой выборку из кругового нормального распределения в k-мерном пространстве с нулевым математическим ожиданием и ковариационной матрицей $\sigma^2(1)I$, где I — единичная матрица. Другими словами, $\varepsilon(i) = \eta(1)e(1) + \eta(2)e(2) + \ldots + \eta(k)e(k)$, где $e(1), e(2), \ldots, e(k)$ — ортонормальный базис в k-мерном пространстве, а $\{\eta(i,t),\ i = 1, 2, \ldots, \mathrm{n},\ t = 1, 2, \ldots, \mathrm{k}\}$ — совокупность независимых в совокупности одномерных случайных величин с нулевым математическим ожиданием и дисперсией $\sigma^2(1)$.

Во второй модели искажения наложены непосредственно на сами расстояния:

$$s(i,j) = d(c(i), c(j)) + \varepsilon(i,j),\ i,j = 1, 2, \dots , n,\ i \neq j,$$

где $\{\varepsilon(i,j),\ i,j = 1, 2, \dots , n\}$ — независимые в совокупности нормальные случайные величины с математическим ожиданием) и дисперсией $\sigma^2(1)$.

В работе [18] показано, что для обеих сформулированных моделей минимум среднего квадрата ошибки α_m при $n \to \infty$ сходится по вероятности к

$$f(m) = f_1(m) + \sigma^2(1)(k - m),\ m = 1, 2, \dots, k,$$

где

$$f_1(m) = \begin{cases} \sigma^2(0)(m_0 - m), & m < m_0, \\ 0, & m \geq m_0. \end{cases}$$

Таким образом, функция $f(m)$ линейна на интервалах $[1, m_0]$ и $[m_0, k]$, причем на первом интервале она убывает быстрее, чем на втором. Отсюда следует, что статистика:

$$m^* = \underset{m}{Arg} \min \{\alpha_{m+1} - 2\alpha_m + \alpha_{m-1})$$

является состоятельной оценкой истинной размерности m_0.

Итак, из вероятностной теории вытекает рекомендация — в качестве оценки размерности факторного пространства использовать m^*. Отметим, что подобная рекомендация была сформулировано как эвристическая одним из основателей многомерного шкалирования Дж. Краскалом [15]. Он исходил из опыта практического использования многомерного шкалирования и вычислительных экспериментов. Вероятностная теория позволила обосновать эту эвристическую рекомендацию.

6.6. Индексы и их применение

Индекс (лат. *index* — показатель, список) — статистический относительный показатель, характеризующий соотношение во времени (динамический индекс) или в пространстве (территориальный индекс) социально-экономических явлений. Речь идет о ценах на товары и услуги, объемах производства, себестоимости, объемах продаж и др. Индексы делятся на индивидуальные и сводные. Так, индивидуальный динамический индекс описывает изменение тех или иных явлений во времени. Например, изменения цены на отдельный товар, объема выплавки стали, урожайности картофеля. Для вычисления индивидуального индекса значение измеряемой величины в текущем периоде делят на ее значение в базисном периоде. Сводный индекс служит для сопоставления непосредственно несоизмеримых, разнородных явлений. Например, объемов продаж различных продовольственных товаров (в килограммах). Для требуемого сопоставления необходимо составные элементы несоизмеримых явлений сделать соизмеримыми, выразив их общей мерой: стоимостью, трудовыми затратами и т.д. Сводные индексы обычно имеют один из трех видов:

$$I_1 = \frac{\sum x_1 f_0}{\sum x_0 f_0}, \quad I_2 = \frac{\sum x_1 f_1}{\sum x_0 f_1}, \quad I_3 = \frac{\sum x_1 f_1}{\sum x_0 f_0},$$

где x — индексируемая величина, f — веса индексов, 0 и 1 — знаки соответственно базисного и текущего периодов, суммирование ведется по одному и тому же множеству «индексов суммирования» [19, с. 154]. (Обратите внимание, что в статистических методах термин «индекс» может использоваться в разных смыслах.) Таким образом, индексы, как и коэффициенты корреляции, зависят от двух переменных — индексируемой величины x и весов индексов f.

В качестве примера построения и использования индексов рассмотрим индекс потребительских цен, он же — индекс инфляции.

Под инфляцией понимаем рост (изменение) цен [6]. При анализе экономических процессов, протяженных во времени, необходимо переходить к сопоставимым ценам. Это невозможно сделать без расчета индекса роста цен, т.е. индекса инфляции. Проблема состоит в том, что цены на разные товары растут с различной скоростью, и необходимо эти скорости усреднять.

Рассмотрим конкретного покупателя товаров и услуг, т.е. конкретного экономического субъекта: физическое лицо, домохозяйство или фирму. Он покупает не один товар, а много. Обозначим через n количество типов товаров или услуг (далее кратко — товаров), которые он хочет и может купить. Пусть:

$$Q_i = Q_i(t),\ i=1, 2, ..., n,$$

— объемы покупок этих товаров в момент времени t по ценам:

$$r_i = r_i(t),\ i=1, 2, ..., n.$$

(Имеется в виду цена за единицу измерения соответствующего товара, например, за штуку или килограмм.)

Подход к измерению роста цен основан на выборе и фиксации потребительской корзины ($Q_1(t)$, $Q_2(t)$, ..., $Q_n(t)$), не меняющейся со временем, т.е. ($Q_1(t)$, $Q_2(t)$, ..., $Q_n(t)$) ≡ (Q_1, Q_2, ..., Q_n). Затем необходимо сравнить стоимости потребительской корзины (Q_1, Q_2, ..., Q_n) в старых $r_i(t_1)$, $i=1, 2, \dots, n$, и новых $r_i(t_2)$, $i=1, 2, \dots, n$, ценах.

Определение. *Индексом инфляции называется*:

$$I(t_1, t_2) = \frac{\sum_{1 \le i \le n} r_i(t_2) Q_i}{\sum_{1 \le i \le n} r_i(t_1) Q_i}.$$

Таким образом, каждой потребительской корзине соответствует свой индекс инфляции. Однако согласно теореме сложения для индекса инфляции [6] он является средним взвешенным арифметическим роста цен на отдельные товары. Поэтому индексы инфляции, рассчитанные по разным достаточно обширным и представительным потребительским корзинам, достаточно близки между собой (см. конкретные данные в [6]).

Институт высоких статистических технологий и эконометрики (ИВСТЭ) использовал для измерения инфляции минимальную потребительскую корзину физиологически необходимых продовольственных товаров [6]. Она была разработана на основе исходных данных Института питания Российской академии медицинских наук (РАМН). Данные о динамике индекса инфляции приведены в табл. 3.

Таблица 3

Индекс инфляции и стоимость потребительской корзины

№ п/п	**Дата снятия цен *t***	**Стоимость потребительской корзины *S*(*t*) (руб.) в месяц**	**Индекс инфляции *I*(31.03.91; *t*)**
1	31.03.1991	26,60	1,00
2	14.08.93	17 691,00	665,08
3	15.11.93	28 050,00	1054,51
4	14.03.94	40 883,00	1536,95
5	14.04.94	44 441,00	1670,71
6	28.04.94	47 778,00	1796,17
7	26.05.94	52 600,00	1977,44

№ п/п	Дата снятия цен t	Стоимость потребительской корзины $S(t)$ (руб.) в месяц	Индекс инфляции $I(31.03.91; t)$
8	08.09.94	58 614,00	2203,53
9	06.10.94	55 358,00	2081,13
10	10.11.94	72 867,00	2739,36
11	01.12.94	78 955,00	2968,23
12	29.12.94	97 897,00	3680,34
13	02.02.95	129 165,00	4855,83
14	02.03.95	151 375,00	5690,79
15	30.03.95	160 817,00	6045,75
16	27.04.95	159 780,00	6006,77
17	01.06.95	167 590,00	6300,38
18	29.06.95	170 721,00	6418,08
19	27.07.95	175 499,00	6597,71
20	31.08.95	173 676,00	6529,17
21	28.09.95	217 542,00	8178,27
22	26.10.95	243 479,00	9153,35
23	30.11.95	222 417,00	8361,54
24	28.12.95	265 716,00	9989,32
25	01.02.96	287 472,55	10 807,24
26	05.03.96	297 958,00	11 201,43
27	05.04.96	304 033,44	11 429,83
28	08.05.96	305 809,55	11 496,60
29	05.06.96	302 381,69	11 367,73
30	03.07.96	306 065,21	11 506,21
31	03.08.96	308 963,42	11 615,17
32	07.09.96	288 835,07	10 858,46
33	01.10.96	278 235,35	10 459,98
34	05.11.96	287 094,77	10 793,04
35	04.12.96	298 024,76	11 203,94
36	03.01.97	314 287,16	11 815,31
37	04.02.97	334 738,24	12 584,14
38	04.01.98	345,72	12,997
39	03.01.99	630,07	20,395
40	05.01.2000	737,80	27,737
41	03.01.01	886,84	33,340
42	03.01.02	1051,79	39,541
43	03.01.03	1210,62	45,512
44	03.01.04	1355,91	50,974
45	14.05.04	1369,10	51,470
46	11.01.05	1463,98	55,037
47	10.01.06	1525,62	57,354
48	26.11.06	1571,26	59,070

№ п/п	Дата снятия цен t	Стоимость потребительской корзины $S(t)$ (руб.) в месяц	Индекс инфляции $I(31.03.91; t)$
49	10.01.07	1580,89	59,432
50	02.07.07	1644,38	61,819
51	03.01.08	1891,04	71,092
52	03.07.08	2286,54	85,960
53	10.01.09	2458,03	92,407
54	02.07.09	2682,08	100,83
55	11.01.10	2833,04	106,48
56	01.07.10	2954,86	111,07
57	11.01.11	3082,35	115,85
58	01.07.11	3267,29	122,80
59	11.01.12	3496,00	131,40
60	01.07.12	3740,72	141,91
61	11.01.13	4227,01	160,36
62	01.07.13	4507,77	170,79
63	11.01.14	4776,53	181,20
64	01.07.14	5206,42	197,51
65	11.01.15	5349,77	202,94
66	01.07.15	5681,46	215,52
67	11.01.16	6039,82	229,12
68	01.07.16	6291,41	238,28
69	01.01.17	6731,81	254,96
70	01.07.17	6799,13	257,51
71	01.01.18	6867,12	260,08
72	01.07.18	7004,46	265,29
73	01.01.19	7144,55	270,59
74	01.01.20	7358,89	278,71
75	01.01.21	7719,47	292,37

Примечание. Учитывается проведенная 01.01.98 деноминация рубля. Стоимость потребительской корзины приводится без включения группы «прочие» (6 % от стоимости основной части корзины).

Использования индекса инфляции в экономических расчетах при принятии решений. Хорошо известно, что стоимость денежных единиц со временем меняется. Например, на один доллар США полвека назад можно было купить примерно в восемь раз больше материальных ценностей (например, продовольствия), чем сейчас (см. таблицу пересчета в учебнике [20]), а если сравнивать с временами Тома Сойера — примерно в 100 раз больше. Причем стоимость денежных единиц с течением времени, как правило, падает. Этому есть две основные причины — банковский процент и инфляция. В экономике есть инструменты для учета изменения стоимости денежных единиц с течением времени. Один из наиболее известных — расчет *NPV* (*Net Present Value*) — чистой текущей стоимости. Однако бухгалтерский учет и построенный на данных баланса предприятия экономический анализ финансово-хозяйственной деятельности предприятия пока что, как правило, игнорируют сам факт наличия инфляции. Обсудим некоторые возможности использования индекса инфляции в экономических расчетах в процессе подготовки и принятия решений. Чтобы избежать обсуждения современного экономического положения, рассмотрим условные годы.

Переход к сопоставимым ценам. Индекс инфляции даст возможность перехода к сопоставимым ценам, расходам, доходам и другим экономическим величинам. Например, по данным табл. 3 индекс инфляции за 4 года — с 14.03.Х+1 г. по 16.03.Х+5 г. — составил 5 936. Это означает, что покупательной способности 1 рубля марта Х+1 г. соответствует примерно 6 000 (а точнее 5 936) руб. марта Х+5 г.

Рассмотрим приведение доходов к неизменным ценам. Пусть Иван Иванович Иванов получал в Х г. 300 руб. в месяц, а в мае Х+5 г. — 1 миллион руб. в месяц. Увеличились его доходы или уменьшились?

Номинальная заработная плата выросла в 1 000 000/300 = 3 333 раза. Однако индекс инфляции на 18 мая Х+5 г. составлял 7 080. Это значит, что 1 руб. Х г. соответствовал по покупательной способности 7 080 руб. в ценах на 18.05.Х+5 г. Следовательно, в ценах Х г. доход И.И. Иванова составлял 1 000 000/7 080 = 142 руб. 24 коп., т.е. 47,4 % от дохода в Х г.

Можно поступить наоборот, привести доход Х г. к ценам на 18 мая Х+5 г. Для этого достаточно умножить его на индекс инфляции: доход Х г. соответствует $300 \cdot 7\,080$ = 2 миллиона 124 тыс. руб. в ценах мая Х+5 г.

Средняя зарплата. По данным Госкомстата РФ средняя заработная плата составляла в Х г. 297 руб., в октябре Х+3 г. — 93 тыс. руб., в январе Х+5 г. — 303 тыс. руб. Поскольку зарплата тратится в основном в следующем месяце после получки, то рассмотрим индексы инфляции на 15.11.Х+3 г. и 2.02.Х+5 г., равные 1 045 и 4 811 соответственно. В ценах Х г. средняя зарплата составила 89 руб. и 62 руб.98 коп. соответственно, т.е. 30 и 21,2 % от зарплаты Х г.

Средняя зарплата рассчитывается путем деления фонда оплаты труда на число работников. При этом объединяются доходы и низкооплачиваемых лиц и сравнительно высокооплачиваемых. Известно, что распределение доходов резко асимметрично, большому числу низкооплачиваемых работников соответствует малое число лиц с высокими доходами. За Х+1 — Х+5-е годы дифференциация доходов резко увеличилась. Это означает, что доходы основной массы трудящихся сдвинулись влево относительно средней зарплаты. По нашей оценке, 50 % получают не более 70 % от средней зарплаты, т.е. не более 212 100 руб. по состоянию на январь X+5 г., а наиболее массовой является оплата в 50 % от средней, т.е. около 150 тыс. руб. в месяц.

Доходы отдельных слоев трудящихся снизились еще существеннее. Зарплата профессора Московского государственного института электроники и математики (технического университета) составляла в марте X+4 г. — 42 руб.92 коп. (в ценах X г.), в июле X+5 г. — 43 руб. 01 коп., т.е. с X г. (400 руб.) снизилась в 9,3 раза, дошла до уровня прежней студенческой стипендии. А студенческие стипендии снизились примерно в той же пропорции и составляли 4–5 руб. в ценах X г.

Кроме того, необходимо учесть, что Госкомстат (и Росстат) учитывает начисленную зарплату, а не выплаченную. В отдельные периоды отечественной истории выплата заработной платы откладывалась надолго.

Минимальная зарплата и прожиточный минимум. Минимальная зарплата в сентябре X+4 г. (22 500 руб.) и в мае X+5 г. (43 700 руб.) составляла 38 % и 23,4 % соответственно от стоимости минимальной физиологически необходимой продовольственной корзины. После подъема до 55 тыс. руб. она в сентябре X+5 г. составляла около 26,34 % от стоимости корзины, т.е. реально уменьшилась в 1,44 раза по сравнению с сентябрем X+4 г. В дальнейшем уменьшение стало еще более заметным.

Минимальная зарплата вместе с единой тарифной сеткой во многом определяла зарплату работников бюджетной сферы. Учитывая снижение коэффициентов тарифной сетки, проведенное весной X+5 г., снижение в 1,5 раза покупательной способности минимальной зарплаты, необходимо заключить, что в сентябре X+5 г. доход бюджетников в 2 раза меньше, чем год назад.

Оценим прожиточный минимум. Бюджетные обследования X года показали, что для лиц с низкими доходами расходы на продовольствие составляют около 50 % всех расходов, т.е. на промтовары и услуги идет около 50 % доходов. Это соотношение подтвердило и проведенное ИВСТЭ бюджетное обследование конца X+5 г. Исходя из него, среднедушевой прожиточный минимум можно оценить, умножая на 2,0 стоимость минимальной продовольственной корзины ИВСТЭ. Например, на 1 сентября X+5 г. — 418 220 руб. Прожиточный уровень для семьи из трех человек — муж, жена и ребенок — должен был на 1 сентября X+5 г. составлять 1,25 миллиона руб. (в месяц). Например, муж должен получать 800 тыс. руб., жена — 450 тыс. руб. в месяц. Очевидно, доходы большинства трудящихся меньше прожиточного уровня.

Численные значения стоимостей потребительских корзин и индексов инфляции рассчитаны ИВСТЭ в основном по ценам на продукты в Москве и Подмосковье. Однако для других регионов численные значения отличаются мало. Для Москвы индекс инфляции на 1.09.Х+5 г. — 7 759, а для Иванова на 1.08.Х+5 г. — 7 542. Поскольку потребительская корзина на 14.03.Х+1 г. в Иванове была на 95 коп. дешевле, то и на 1.08.Х+5 г. она несколько дешевле — 195 337 руб., а прожиточный минимум равен 390 673 руб. Приведенные выше численные значения для Москвы в качестве первого приближения можно использовать для различных регионов России.

Индексы инфляции с помощью описанной выше методики можно рассчитать для любого региона, профессиональной или социальной группы, отдельного предприятия или даже конкретной семьи. Эти значения могут быть эффективно использованы на трехсторонних переговорах между профсоюзами, работодателями и представителями государства.

Проценты по вкладам в банк, плата за кредит и инфляция. Рассмотрим банк, честно выполняющий свои обязательства. Пусть он дает 10 % в месяц по депозитным вкладам. Тогда 1 руб., положенный в банк, через месяц превращается в 1,1 руб., а через 2 — по формуле сложных процентов — в $1{,}1^2 = 1{,}21$ руб., ..., через год — в $1{,}1^{12} = 3{,}14$ руб. Однако за год росли не только вклады, но и цены. Например, с 19.05.Х+4 г. по 18.05.Х+5 г. индекс инфляции составил 3,73. Значит, в ценах на момент оформления вкладов итог годового хранения равен 3,14 / 3,73 = 0,84 руб. Хранение оказалось невыгодным — реальная стоимость вклада уменьшилась на 16 %, несмотря на, казалось бы, очень выгодные условия банка.

Пусть фирма получила кредит под 200 % годовых. Значит, вместо 1 рубля, полученного в настоящий момент в кредит, через год ей надо отдать 3 рубля. Пусть она взяла кредит 19.05.Х+4 г., а отдает 18.05.Х+5 г. Тогда в ценах на момент взятия кредита она отдает 3/3,73 = 0,80 руб. за 1 руб. кредита. Таким образом, кредит частично превратился в подарок — возвращать надо на 20 % меньше, чем получил, реальная ставка кредита отрицательна, она равна (– 20) %! Такова была типичная ситуация в России в течение ряда лет начиная с 1992 г., особенно в Х+2 — Х+4 гг. Но бесплатных подарков в бизнесе не бывает — за них надо платить по другим каналам, как правило, криминальным.

Сколько стоит доллар? Например, в июле 1995 г. индекс инфляции около 7 000, а курс доллара США — около 4 500 руб. за доллар. Следовательно, доллар США стоит 4 500 / 7 000 = 0,64 руб. в ценах 1990 г., т.е. примерно соответствует официальному обменному курсу в 1980-х гг. В сентябре 1994 г. курс доллара был около 2 000, а индекс инфляции — около 2 200, т.е. доллар стоил около 0,90 руб. в ценах 1990 г. Реальная покупательная способность доллара упала за 10 месяцев в 1,42 раза.

В середине 2003 г. курс доллара был несколько больше 30 руб. (30 руб. 38 коп.), индекс инфляции составлял 52.56, следовательно, 1 доллар США по своей покупательной способности в России на июль 2003 г. соответствовал 58 копейкам начала 1991 г. В начале 2004 г. курс доллара составлял около 28 руб. 50 коп. при индексе инфляции 56.068, следовательно, 1 доллар США по своей покупательной способности в России на начало 2004 г. соответствовал 51 копейке начала 1991 г. Всего за полгода доллар подешевел на 12 %.

Инфляция, показатели работы предприятия и ВВП. Индексы инфляции используются для пересчета номинальных цен в неизменные (сопоставимые). Другими словами, для приведения доходов и расходов к ценам определенного момента времени. Потребительские корзины для промышленных предприятий, конечно, должны включать промышленные товары, а потому отличаться от потребительских корзин, ориентированных для изучения жизненного уровня.

Сколько стоит предприятие? Важно оценить основные фонды. Для этого нужно взять их стоимость в определенный момент времени и умножить на индекс инфляции (и учесть амортизационные отчисления).

Валовой внутренний продукт, валовой национальный продукт и другие характеристики экономического положения страны рассчитываются в текущих ценах. Для перехода к неизменным ценам, грубо говоря, надо поделить на индекс инфляции (т.е. умножить на дефлятор).

Проблема учета инфляции при экономическом анализе финансово-хозяйственной деятельности предприятия. Как известно, разработана и широко применяется развернутая система коэффициентов, используемых при экономическом анализе финансово-хозяйственной деятельности предприятия [21]. Она основана на данных бухгалтерского баланса. Естественно, опирается на два столбца баланса — данные на «начало периода» и данные на «конец периода». Записывают в эти столбцы номинальные значения. В настоящее время инфляцию полностью игнорируют. Это приводит к искажению реального положения предприятия. Денежные средства преувеличиваются, а реальная стоимость основных фондов занижается. По официальной отчетности предприятие может считаться получившим хорошую прибыль, а по существу — не иметь средств для продолжения производственной деятельности.

Ясно, что учитывать инфляцию надо. Вопрос в другом — как именно. Потребительская корзина должна, видимо, состоять из тех товаров и услуг, которые предприятие покупает. Стоимость основных фондов может не убывать в соответствии с амортизацией, а возрастать согласно отраслевому темпу инфляции, и т.д.

Литература

1. *Крамер, Г.* Математические методы статистики / Г. Крамер. — Москва : Мир, 1975. — 648 с.

2. *Красильников, В.В.* Статистика объектов нечисловой природы / В.В. Красильников. — Набережные Челны : Изд-во Камского политехнического института, 2001. — 144 с.

3. *Кендэл, М.* Ранговые корреляции / М. Кендэл. — Москва : Статистика, 1975. — 216 с.

4. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1983. — 416 с.

5. *Себер, Дж.* Линейный регрессионный анализ / Дж. Себер. — Москва : Мир, 1980. — 456 с.

6. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 2-е изд., испр. и доп. — Москва : Экзамен, 2003. — 576 с.

7. *Орлов, А.И.* Оценка размерности модели в регрессии / А.И. Орлов // Алгоритмическое и программное обеспечение прикладного статистического анализа. Ученые записки по статистике. — Т. 36. — Москва : Наука, 1980. — С. 92–99.

8. *Кендалл, М.Дж.* Многомерный статистический анализ и временные ряды / М.Дж. Кендалл, А. Стьюарт. — Москва : Наука, 1976. — 736 с.

9. *Орлов, А.И.* Некоторые вероятностные вопросы теории классификации / А.И. Орлов // Прикладная статистика. Ученые записки по статистике. — Т. 45. — Москва : Наука, 1983. — С. 166–179.

10. *Орлов, А.И.* Парные сравнения в асимптотике Колмогорова / А.И. Орлов // Экспертные оценки в задачах управления. — Москва : Изд-во ИПУ, 1982. — С. 58–66.

11. *Орлов, А.И.* Математические методы в изучении способных к математике школьников / А.И. Орлов, Г.А. Гусейнов // Исследования по вероятностно-статистическому моделированию реальных систем. — Москва : ЦЭМИ АН СССР, 1977. — С. 80–93.

12. *Куперштох, В.Л.* Сумма внутренних связей как показатель качества классификации / В.Л. Куперштох, Б.Г. Миркин, В.А. Трофимов // Автоматика и телемеханика. — 1976. — № 3. — С. 91–98.

13. Прогнозирование исхода инфаркта миокарда с помощью программы «Кора-3» / И.М. Гельфанд, М.А. Алексеевская, Ш.А. Губерман и др. // Кардиология. — 1977. — Т. 17. — № 6. — С. 19–23.

14. *Харман, Г.* Современный факторный анализ / Г. Харман. — Москва : Статистика, 1972. — 488 с.

15. *Терехина, А.Ю.* Анализ данных методами многомерного шкалирования / А.Ю. Терехина. — Москва : Наука, 1986. — 168 с.

16. *Перекрест, В.Т.* Нелинейный типологический анализ социально-экономической информации: Математические и вычислительные методы / В.Т. Перекрест. — Ленинград : Наука, 1983. — 176 с.

17. Анализ нечисловой информации / Ю.Н. Тюрин, Б.Г. Литвак, А.И. Орлов и др. — Москва : Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1981. — 80 с.

18. Орлов А.И. Общий взгляд на статистику объектов нечисловой природы / А.И. Орлов // Анализ нечисловой информации в социологических исследованиях. — Москва : Наука, 1985. — С. 58–92.

19. Статистический словарь / главный редактор М.А. Королев. — Москва : Финансы и статистика, 1989. — 623 с.

20. *Макконнелл, К.Р.* Экономикс: Принципы, проблемы и политика. В 2 томах / К.Р. Макконнелл, С.Л. Брю ; перевод с английского 11-го издания — Москва : Республика, 1992.

21. *Баканов, М.И.* Теория экономического анализа / М.И. Баканов, А.Д. Шеремет. — Москва : Финансы и статистика, 2000. — 416 с.

22. *Robinson, D.E.* Estimates for the points of intersection of two polynomial regressions / D.E. Robinson // Journal of American Statistical Association. — 1964. — V. 19. — № 2. — P. 214–238.

23. *Орлов, А.И.* Программно-алгоритмическое обеспечение статистических методов в САПР стандартов / А.И. Орлов, В.Н. Медведев // Тезисы докладов III Всесоюзной школы-семинара «Программно-алгоритмическое обеспечение прикладного многомерного статистического анализа». — Москва : ЦЭМИ АН СССР, 1987. — С. 313–314.

24. *Орлов, А.И.* Вероятностно-статистические модели корреляции и регрессии / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 160. — С. 130–162.

25. *Луценко, Е.В.* Когнитивные функции как обобщение классического понятия функциональной зависимости на основе теории информации в системной нечеткой интервальной математике / Е.В. Луценко, А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 95. — С. 122–183.

26. *Орлов, А.И.* Математические методы теории классификации / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 95. — С. 23–45.

27. *Орлов, А.И.* Базовые результаты математической теории классификации / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 110. — С. 219–239.

28. *Орлов, А.И.* Основные требования к методам анализа данных (на примере задач классификации) / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 159. — С. 239–267.

29. *Орлов, А.И.* Прогностическая сила — наилучший показатель качества алгоритма диагностики / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 99. — С. 33–49.

30. *Луценко, Е.В.* Методы снижения размерности пространства статистических данных / Е.В. Луценко, А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 119. — С. 92–107.

31. *Орлов, А.И.* Оценивание размерности вероятностно-статистической модели / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 162. — С. 1–36.

32. *Орлов, А.И.* Компьютерно-статистические методы: состояние и перспективы / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 103. — С. 163–195.

33. *Орлов, А.И.* Оценка инфляции по независимой информации / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 108. — С. 259–287.

34. *Куликова, С.Ю.* Контроллинг динамики потребительских цен и прожиточного минимума / С.Ю. Куликова, В.С. Муравьева, А.И. Орлов // Научный журнал КубГАУ. — 2017. — № 126. — С. 403–421.

35. *Найдис, О.А.* Потребительские корзины, контроллинг уровня цен и МРОТ / О.А. Найдис, И.О. Найдис // Контроллинг. — 2019. — № 74. — С. 40–53.

Контрольные вопросы и задачи

1. Имеются данные за несколько лет о торговом обороте Y западногерманского предприятия и его расходах на рекламу X. Данные представлены в табл. 4.

Таблица 4

Расходы на рекламу и торговый оборот предприятия

Годы, t	68	69	70	71	72	73	74	75
Расходы на рекламу $x(t)$, тыс. марок	4	4	5	6	8	8	10	11
Торговый оборот $y(t)$, млн марок	4	5	6	6	8	10	12	13

Вычислите линейный коэффициент корреляции между случайными величинами X и Y. С помощью метода наименьших квадратов определите коэффициенты линейной регрессии $Y = aX + b$. Постройте график (заданные точки (x_i, y_i) и прямую $y = a^*x + b^*$). Найдите доверительные границы для регрессионной зависимости (при доверительной вероятности $\gamma = 0{,}95$). Нанесите доверительные границы на график. Сделайте точечный и интервальный прогноз для торгового оборота при расходах на рекламу, равных 15 (тыс. марок ФРГ).

Аналогичным образом изучите зависимости расходах на рекламу X и торгового оборота Y от времени t.

2. Семь школьников выполняют несколько заданий по математике и физике, которые оцениваются баллами 1–5, затем вычисляется средний балл для каждого школьника по каждому предмету: по математике — x_i, по физике — y_j. Данные представлены в табл. 5. Определите, существует ли корреляция (т.е. связь) между этими оценками, вычислив коэффициент ранговой корреляции Спирмена.

Таблица 5

Средние баллы по математике и физике

Школьник	Средний балл по математике x_i	Средний балл по физике y_i
A	1,8	3,2
B	3,0	2,8
C	3,5	4,0
D	4,0	5,0
E	5,0	3,6
F	3,8	2,4
G	2,0	1,2

3. Исходные данные (табл. 6) — набор n пар чисел (t_k, x_k), $k = 1, 2, ..., n$, где t_k — независимая переменная (например, время), а x_k — зависимая (например, индекс инфляции). Предполагается, что переменные связаны зависимостью

$$x_k = a\, t_k + b + e_k,\ k = 1, 2, ..., n,$$

где a и b — параметры, неизвестные статистику и подлежащие оцениванию, а e_k — погрешности, искажающие зависимость.

Таблица 6

Исходные данные для расчетов по методу наименьших квадратов

t_k	1	3	4	7	9	10
x_k	12	20	20	32	35	42

Методом наименьших квадратов оцените параметры a и b линейной зависимости. Выпишите восстановленную зависимость.

Вычислите восстановленные значения зависимой переменной, сравните их с исходными значениями (найдите разности) и проверьте условие точности вычислений (при отсутствии ошибок в вычислениях сумма исходных значений должна равняться сумме восстановленных).

Найдите остаточную сумму квадратов и оцените дисперсию погрешностей.

Выпишите точечный прогноз, а также верхнюю и нижнюю доверительные границы для него (для доверительной вероятности 0,95).

Рассчитайте прогнозное значение и доверительные границы для него для момента t = 12.

Как изменятся результаты, если доверительная вероятность будет увеличена? А если она будет уменьшена?

4. Как в методе наименьших квадратов используются преобразования переменных?

5. Как соотносятся задачи группировки и задачи кластер-анализа?

6. В табл. 7 приведены попарные расстояния между десятью социально-психологическими признаками способных к математике школьников [11].

Примените к этим данным алгоритмы ближнего соседа, средней связи и дальнего соседа. Для каждого из трех алгоритмов выделите наиболее устойчивые разбиения на кластеры.

Таблица 7

Попарные расстояния между социально-психологическими признаками

Номер признака	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**9**	**10**
2	1028	–	–	–	–	–	–	–	–
3	1028	608	–	–	–	–	–	–	–
4	1050	688	610	–	–	–	–	–	–
5	1012	686	636	634	–	–	–	–	–
6	1006	566	538	616	562	–	–	–	–
7	1012	1026	748	692	774	732	–	–	–
8	960	1088	1144	1122	1120	1130	1110	–	–
9	1026	878	874	830	836	802	904	1040	–
10	990	744	674	744	718	580	814	1090	830

7. Расскажите о динамике индекса инфляции в России.

Темы докладов, рефератов, исследовательских работ

1. Примеры практического использования методов многомерного статистического анализа.

2. Для непараметрической модели метода наименьших квадратов в случае линейной функции одной переменной разработайте алгоритмы

а) расчета доверительных границ для коэффициентов модели;

б) проверки гипотез относительно этих коэффициентов.

3. Докажите, что сумма исходных значений зависимой переменной должны быть равна сумме восстановленных значений.

4. Критерии качества регрессионной модели.

5. Использование непараметрических оценок плотности для восстановления зависимости.

6. В задаче о точке встречи оцените асимптотические дисперсии оценок уровня качества в момент встречи x^* и временного лага L^* и используйте полученные результаты для доверительного оценивания и проверки гипотез.

7. Многообразие методов кластер-анализа.

8. Теоремы умножения и сложения для индекса инфляции.

9. Экспериментальная работа: соберите данные о ценах и рассчитайте индекс инфляции для своего региона (данные о потребительской корзине ИВСТЭ и ценах на базовый момент времени приведены в [6]).

10. Учет инфляции при проведении анализа финансово-хозяйственной деятельности предприятия.

ГЛАВА 7. СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ДИНАМИКИ

Анализ динамики — это анализ временных рядов. Под временными рядами понимают детерминированные или случайные функции от времени. При этом время предполагается дискретным, в противном случае говорят о случайных процессах, а не о временных рядах.

7.1. Методы анализа и прогнозирования временных рядов

Модели стационарных и нестационарных временных рядов. Пусть $t = 0, \pm 1, \pm 2, \pm 3, \ldots$ Рассмотрим временной ряд $X(t)$. Пусть сначала временной ряд принимает числовые значения. Это могут быть, например, цены на батон хлеба в соседнем магазине или курс обмена доллара на рубли в ближайшем обменном пункте. Обычно в поведении временного ряда выявляют две основные тенденции — тренд и периодические колебания.

При этом под трендом понимают зависимость от времени линейного, квадратичного или иного типа, которую выявляют тем или иным способом сглаживания (например, экспоненциального сглаживания) либо расчетным путем, в частности, с помощью метода наименьших квадратов. Другими словами, тренд — это очищенная от случайностей основная тенденция временного ряда.

Временной ряд обычно колеблется вокруг тренда, причем отклонения от тренда часто обнаруживают правильность. Часто это связано с естественной или назначенной периодичностью, например, сезонной или недельной, месячной или квартальной (например, в соответствии с графиками выплаты заплаты и уплаты налогов). Иногда наличие периодичности и тем более ее причины неясны, и задача статистика — выяснить, действительно ли имеется периодичность.

Элементарные методы оценки характеристик временных рядов обычно достаточно подробно рассматриваются в курсах «Общей теории статистики» (см., например, учебники [1, 2]), поэтому нет необходимости подробно разбирать их здесь. О некоторых современных методах оценивания длины периода и самой периодической составляющей речь пойдет ниже в разделе 7.2.

Характеристики временных рядов. Для более подробного изучения временных рядов используются вероятностно-статистические модели. При этом временной ряд $X(t)$ рассматривается как случайный процесс (с дискретным временем). Основными характеристиками $X(t)$ являются *математическое ожидание* $X(t)$:

$$a(t) = MX(t),$$

дисперсия $X(t)$:

$$\sigma^2(t) = DX(t)$$

и *автокорреляционная функция* временного ряда $X(t)$:

$$\rho(t,s) = \frac{M(X(t)-a(t))(X(s)-a(s))}{\sigma(t)\sigma(s)},$$

т.е. функция двух переменных, равная коэффициенту корреляции между двумя значениями временного ряда $X(t)$ и $X(s)$.

В теоретических и прикладных исследованиях рассматривают широкий спектр моделей временных рядов. Выделим сначала *стационарные* модели. В них совместные функции распределения $F(t_1, t_2, ..., t_k)$ для любого числа моментов времени k, а потому и все перечисленные выше характеристики временного ряда *не меняются со временем*. В частности, математическое ожидание и дисперсия являются постоянными величинами, автокорреляционная функция зависит только от разности $t - s$. Временные ряды, не являющиеся стационарными, называются *нестационарными*.

Линейные регрессионные модели с гомоскедастичными и гетероскедастичными, независимыми и автокоррелированными остатками. Как видно из сказанного выше, основное — это «очистка» временного ряда от случайных отклонений, т.е. оценивание математического ожидания. В отличие от простейших моделей регрессионного анализа, рассмотренных в главе 6, здесь естественным образом появляются более сложные модели. Например, дисперсия может зависеть от времени. Такие модели называют гетероскедастичными, а те, в которых нет зависимости от времени — гомоскедастичными. (Точнее говоря, эти термины могут относиться не только к переменной «время», но и к другим переменным.)

Далее, в главе 6 предполагалось, что погрешности независимы между собой. В терминах настоящей главы это означало бы, что автокорреляционная функция должна быть вырожденной — равняться 1 при равенстве аргументов и 0 при их неравенстве. Ясно, что для реальных временных рядов так бывает отнюдь не всегда. Если естественный ход изменений наблюдаемого процесса является достаточно быстрым по сравнению с интервалом между последовательными наблюдениями, то можно ожидать «затухания автокорреляции» и получения практически независимых остатков, в противном случае остатки будут автокоррелированы.

Идентификация моделей. Под идентификацией моделей обычно понимают выявление их структуры и оценивание параметров. Поскольку структура — это тоже параметр, хотя и нечисловой, то речь идет об одной из типовых задач прикладной статистики — оценивании параметров.

Проще всего задача оценивания решается для линейных (по параметрам) моделей с гомоскедастичными независимыми остатками. Восстановление зависимостей во временных рядах может быть проведено на основе методов наименьших квадратов и наименьших модулей оценивания параметров в моделях линейной (по параметрам) регрессии. На случай временных рядов переносятся результаты, связанные с оцениванием необходимого набора регрессоров, в частности, легко получить предельное геометрическое распределение оценки степени тригонометрического полинома.

Однако на более общую ситуацию такого простого переноса сделать нельзя. Так, например, в случае временного ряда с гетероскедастичными и автокоррелированными остатками снова можно воспользоваться общим подходом метода наименьших квадратов, однако система уравнений метода наименьших квадратов и, естественно, ее решение будут иными. Формулы в терминах матричной алгебры, о которых упоминалось в главе 6, будут отличаться. Поэтому рассматриваемый метод называется «*обобщенный метод наименьших квадратов* (ОМНК)».

Замечание. Как уже отмечалось в главе 6, простейшая модель метода наименьших квадратов допускает весьма далекие обобщения, особенно в области системам одновременных эконометрических уравнений для временных рядов. Для понимания соответствующей теории и алгоритмов необходимо владение методами матричной алгебры. Поэтому мы отсылаем тех, кому это интересно, к литературе по системам эконометрических уравнений [3, 4] и непосредственно по временным рядам [5, 6], в которой особенно много интересуются спектральной теорией, т.е. выделением сигнала из шума и разложением его на гармоники. Подчеркнем еще раз, что за каждой главой настоящей книги стоит большая область научных и прикладных исследований, вполне достойная того, чтобы посвятить ей много усилий. Однако из-за ограниченности объема книги мы вынуждены изложение сделать конспективным.

7.2. Системы эконометрических уравнений

В качестве первоначального примера рассмотрим эконометрическую модель временного ряда, описывающего рост индекса потребительских цен (индекса инфляции). Пусть $I(t)$ — рост цен в месяц t (подробнее об этой проблематике см. главу 7 в [7]). По мнению некоторых экономистов естественно предположить:

$$I(t) = cI(t-1) + a + bS(t-4) + e, \qquad (1)$$

где $I(t-1)$ — рост цен в предыдущий месяц (а c — некоторый коэффициент затухания, предполагающий, что при отсутствии внешний воздействий рост цен прекратится), a — константа (она соответствует линейному изменению величины $I(t)$ со временем), $bS(t-4)$ — слагаемое, соответствующее влиянию эмиссии денег (т.е. увеличения объема денег в экономике страны, осуществленному Центральным Банком) в размере $S(t-4)$ и пропорциональное эмиссии с коэффициентом b, причем это влияние проявляется не сразу, а через 4 месяца; наконец, е — это неизбежная погрешность.

Модель (1), несмотря на свою простоту, демонстрирует многие характерные черты гораздо более сложных эконометрических моделей. Во-первых, обратим внимание на то, что некоторые переменные определяются (рассчитываются) внутри модели, такие, как $I(t)$. Их называют *эндогенными (внутренними)*. Другие задаются извне (это *экзогенные* переменные). Иногда, как в теории управления, среди экзогенных переменных, выделяют *управляемые* переменные — те, с помощью выбора значений которых можно привести систему в нужное состояние.

Во-вторых, в соотношении (1) появляются переменные новых типов — с лагами, т.е. аргументы в переменных относятся не к текущему моменту времени, а к некоторым прошлым моментам.

В-третьих, составление эконометрической модели типа (1) — это отнюдь не рутинная операция. Например, запаздывание именно на 4 месяца в связанном с эмиссией денег слагаемом $bS(t-4)$ — это результат достаточно изощренной предварительной статистической обработки. Далее, требует изучения вопрос зависимости или независимости величин $S(t-4)$ и $I(t)$ в различные моменты времени t. От решения этого вопроса зависит, как выше уже отмечалось, конкретная реализация процедуры метода наименьших квадратов.

С другой стороны, в модели (1) всего 3 неизвестных параметра, и постановку метода наименьших квадратов выписать нетрудно. Положим:

$$f(a,b,c) = \sum_{1 \le t \le k} (I(t) - cI(t-1) - a - bS(t-4))^2.$$

Путем минимизации этой функции находим оценки параметров.

Проблема идентифицируемости. Представим теперь модель тапа (1) с большим числом эндогенных и экзогенных переменных, с лагами и сложной внутренней структурой. Вообще говоря, ниоткуда не следует, что существует хотя бы одно решение у такой системы. Поэтому возникает не одна, а две проблемы. Есть ли хоть одно решение (проблема идентифицируемости)? Если да, то как найти наилучшее решение из возможных? (Это — проблема статистической оценки параметров.)

И первая, и вторая задача достаточно сложны. Для решения обеих задач разработано множество методов, обычно достаточно сложных, лишь часть из которых имеет научное обоснование. В частности, достаточно часто пользуются статистическими оценками, не являющимися состоятельными (строго говоря, их даже нельзя назвать оценками).

Коротко опишем некоторые распространенные приемы при работе с системами линейных эконометрических уравнений.

Система линейных одновременных эконометрических уравнений. Чисто формально можно все переменные выразить через переменные, зависящие только от текущего момента времени. Например, в случае уравнения (1) достаточно положить:

$$H(t) = I(t-1),\ G(t) = S(t-4).$$

Тогда уравнение примет вид:

$$I(t) = cH(t) + a + bG(t) + e. \qquad (2)$$

Отметим здесь же возможность использования регрессионных моделей с переменной структурой путем введения фиктивных переменных. Эти переменные при одних значениях времени (скажем, начальных) принимают заметные значения, а при других — сходят на нет (становятся фактически равными 0). В результате формально (математически) одна и та же модель описывает совсем разные зависимости.

Косвенный, двухшаговый и трехшаговый методы наименьших квадратов. Как уже отмечалось, разработана масса методов эвристического анализа систем эконометрических уравнений. Они предназначены для решения тех или иных проблем, возникающих при попытках найти численные решения систем уравнений.

Одна из проблем связана с наличием априорных ограничений на оцениваемые параметры. Например, доход домохозяйства может быть потрачен либо на потребление, либо на сбережение. Значит, сумма долей этих двух видов трат априори равна 1. А в системе эконометрических уравнений эти доли могут участвовать независимо. Возникает мысль оценить их методом наименьших квадратов, не обращая внимания на априорное ограничение, а потом подкорректировать. Такой подход называют косвенным методом наименьших квадратов.

Двухшаговый метод наименьших квадратов состоит в том, что оценивают параметры отдельного уравнения системы, а не рассматривают систему в целом. В то же время трехшаговый метод наименьших квадратов применяется для оценки параметров системы одновременных уравнений в целом. Сначала к каждому уравнению применяется двухшаговый метод с целью оценить коэффициенты и погрешности каждого уравнения, а затем построить оценку для ковариационной матрицы погрешностей. После этого для оценивания коэффициентов всей системы применяется обобщенный метод наименьших квадратов.

Менеджеру и экономисту не следует становиться специалистом по составлению и решению систем эконометрических уравнений, даже с помощью тех или иных программных систем, но он должен быть осведомлен о возможностях этого направления эконометрики, чтобы в случае производственной необходимости квалифицированно сформулировать задание для специалистов по прикладной статистике.

От оценивания тренда (основной тенденции) перейдем ко второй основной задаче эконометрики временных рядов — оцениванию периода (цикла).

7.3. Оценивание длины периоды и периодической составляющей

Рассмотрим достаточно широкий класс практически полезных непараметрических оценок длины периода и периодической составляющей во временных рядах. Из общих результатов об асимптотическом поведении решений экстремальных статистических задач (см. главу 3) вытекает состоятельность этих оценок.

Во многих прикладных задачах рассматривают временной ряд (или случайный процесс):

$$y(t)=x(t)+e(t),$$

где $x(t)$ — детерминированная периодическая функция от времени t, т.е. $x(t)=x(t+T)$ при некотором T, где T — длина периода (минимальная из возможных, поскольку $2T$, $3T$, $4T$ — тоже, как легко видеть, длины периодов), а $e(t)$ — «шумы», случайные погрешности, искажающие периодический сигнал. Требуется оценить (минимальную) длину периода $T=T_0$ и периодическую составляющую $x(t)$. При этом не предполагается, что функция $x(t)$ входит в какое-либо параметрическое семейство, например, конечных сумм синусов и косинусов, т.е. рассматривается задача непараметрического оценивания (минимальной) длины периода и периодической составляющей сигнала.

Приведем примеры прикладных постановок.

1. По акустическим сигналам необходимо установить тип двигателя (и его национальную принадлежность). Предполагается, что двигатели различаются по длине периода и виду основного периодического сигнала. Процедура идентификации основана на оценивании длины периода и периодической составляющей регистрируемого сигнала. Очевидна важность такой задачи при быстрой технической диагностике. В частности, высокая производительность, а потому и высокая экономическая эффективность при ремонте впрямую зависят от умения решать поставленную задачу. Не менее важно по шуму двигателя подводной лодки определить ее тип и национальную принадлежность.

2. В предположении цикличности экономических процессов требуется по статистическим данным установить длину цикла и на основе вида периодической составляющей построить прогноз, например, прогноз урожайности, емкости рынка тех или иных товаров или экономической активности в целом. Часто говорят об экономических циклах, но почти никогда не дают строгого определения понятия цикла. (Под строгим определением понимаем такое, согласно которому можно отличить «цикл» от «не цикла», можно выделить начало и конец цикла, отделить один цикл от другого, короче, однозначно выделить цикл как самостоятельный объект экономического изучения.)

3. По мнению авторов работы [8], для среднесрочного прогнозирования развития социокультурной сферы (социально-политического «климата», живописи, музыки, архитектуры, поэзии и т.д.) необходимо выявить ее цикличность с помощью объективных измерений на базе субъективных первичных данных (т.е. на базе оценок экспертов).

4. В исторических событиях, описываемых согласно распространенной в настоящее время т.н. скалигеровской хронологии, автор работы [9] обнаруживает цикличность. Эта цикличность полностью объясняется новой статистической хронологией (см., например, [10]), построенной с помощью специальных методов статистики объектов нечисловой природы (см. главу 8.5), предна-

значенных для анализа текстов исторических хроник, и одновременно служит еще одним подтверждением этой хронологии.

Описание метода оценивания. Пусть рассматриваемые функции $y(t)$, $x(t)$, $e(t)$ определены на отрезке $[0; A]$. При фиксированном T рассмотрим «куски» сигнала $y(t)$ на последовательных отрезках длины T, т.е. на отрезках $[0;T]$, $[T;2T]$, $[2T;3T]$, ... Удобно ввести последовательность функций на отрезке $[0;T]$, полученную сдвигами этих кусков к началу координат:

$$y_1(t)=y(t),\ y_2(t)=y(t+T),\ y_3(t)=y(t+2T),\ \ldots$$

Все они определены на отрезке $[0;T]$. Число этих функций равно числу полных периодов длины T, укладывающихся на отрезке $[0;A]$, т.е. равно целой части числа A/T. Отметим еще раз, что если T — период, то $2T$, $3T$, $4T$, ... — тоже периоды. В дальнейшем из всех периодов будем рассматривать и оценивать, как правило, только наименьший.

Если $T=T_0$ — истинный период (или кратный ему) и погрешности $e(t)$ отсутствуют, то все введенные в предыдущем абзаце функции совпадают между собой и с периодической составляющей:

$$x(t)=y_1(t)=y_2(t)=y_3(t)=\ldots$$

при всех t из $[0;T]$. При наличии погрешностей полного совпадения не будет. Однако отклонения определяются лишь шумами в различные моменты времени. При этом в качестве оценки периодической составляющей $x(t)$ естественно взять среднее арифметическое $y_{ср}(t)$ функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ... (могут быть использованы и другие виды средних величин).

Если же T отличается от истинного периода T_0 (и кратных ему величин), то различия функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ... между собой определяются также и различием значений $x(t)$ в точках, отстоящих друг от друга на интервалы, длина которых кратна T.

В предположении отсутствия погрешностей (т.е. когда $e(t)$ тождественно равно 0) рассмотрим поведение функции $y_{ср}(t)$ на отрезке $[0;T]$ при росте длины интервала A наблюдения сигнала, а потому и при росте числа периодов — целой части числа A/T. Если $T = T_0$ или T кратно T_0, то, как уже сказано, $y_{ср}(t)$ совпадает с периодической составляющей $x(t)$. Если число T/T_0 иррационально, то можно показать, что значения $t+mT(\mathrm{mod}T_0)$, где m — натуральные числа такие, что $t+mT<A$, асимптотически (при росте A) равномерно заполняют отре-

зок $[0;T_0]$, а потому при выполнении соответствующих условий регулярности, например, непрерывности периодической составляющей сигнала, функция $y_{ср}(t)$ приближается к константе — среднему значению периодического сигнала $x(t)$, т.е. интегралу от $x(t)$ по отрезку $[0;T_0]$, деленному на T_0. При этом при конечных A функция $y_{ср}(t)$ отлична от константы. (Здесь запись $t+mT(\mathrm{mod}T_0)$ означает теоретико-числовое сравнение по модулю T_0 , т.е. взятие дробной части от числа $(t+mT)/T_0$, что соответствует вычитанию соответствующего количества целых периодов T_0.

Если же число T/T_0 рационально, то наблюдаем промежуточный случай по сравнению с двумя описанными выше, в котором $y_{ср}(t)$, как можно показать, приближается к периодической функции с периодом $T=T_0/n$ при некотором натуральном n. Эта функция получена усреднением n последовательных участков длины T_0/n периодического сигнала $x(t)$. Она не является константой, хотя разброс ее значений меньше, чем для исходного периодического сигнала, поскольку T_0 — минимальная длина периода.

Из сказанного вытекает, что для оценивания T целесообразно ввести два показателя: показатель разброса $F(T;Y)=F(T;\ y_1(t),\ y_2(t),\ y_3(t),\ \ldots)$ множества функций $\{y_1(t), y_2(t), y_3(t), \ldots\}$ на отрезке $[0;T]$ и показатель размаха $G(T;Y)=G(T,\ y_{ср}(t))$ функции $y_{ср}(t)$ на отрезке $[0;T]$. Символ Y означает здесь, что показатели разброса и размаха строятся по функции $y(t)$.) При этом показатель разброса нацелен на оценку различий в значениях семейства функций при одном и том же значении аргумента. А показатель размаха — на различие значений одной и той же функции при различных значениях аргумента. Ниже выписан ряд формул для этих показателей в случае непрерывного времени. Для дискретного времени их можно адаптировать двумя способами: либо заменив sup на max, а интеграл на сумму; либо расширив область определения используемых функций на весь отрезок, например, соединив соседние точки отрезками или использовав для заполнения пропусков сплайны более высокого порядка.

В качестве оценки длины периода по фиксированным показателям разброса $F(T;Y)$ и размаха $G(T;Y)$ представляется рациональным использовать то T, при котором отношение $F(T;Y)/G(T;Y)$ **впервые** (при росте T, начиная с 0) достигает минимума. Впервые — поскольку величины, кратные периоду, сами являются периодами. Поскольку показатели разброса $F(T;Y)$ и размаха $G(T;Y)$ могут быть выбраны многими разными способами, можно указанным выше способом построить целое семейство алгоритмов оценивания длины периода. С каждым из которых может быть связано семейство методов оценивания периодической составляющей путем того или иного способа усреднения функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ...

Показатели разброса и размаха. Ввести показатели разброса $F(T;Y)=F(T;\ y_1(t),\ y_2(t),\ y_3(t),\ ...)$ можно разными способами. Пусть $k=[A/T]$. Можно использовать различные функционалы супремумного типа (здесь и далее число слагаемых k не будем указывать в обозначении функционалов). Первым рассмотрим максимальный разброс непосредственно между значениями функций:

$$F_1(T,Y)=\sup\{|\,y_i(t)-y_j(t)\,|, i, j=1,2,\ldots,k, 0\le t\le T\}.$$

Второй функционал супремумного типа будет учитывать не произвольные отклонения, а только отклонения от «средней функции», т.е. иметь вид:

$$F_2(T,Y)=\sup\{|\,y_i(t)-y_{cp}(t)\,|, i=1,2,\ldots,k, 0\le t\le T\}.$$

Третий функционал показывает, какую зону «заметают» значения функций:

$$F_3(T,Y)=\sup\{y_i(t), i=1,2,\ldots,k, 0\le t\le T\}-\inf\{y_i(t), i=1,2,\ldots,k, 0\le t\le T\}.$$

Для применения функционалов интегрального типа целесообразно сделать замену переменной $q=t/T$ и перейти к функциям $Y_i(q)=y_i(t)=y_i(qT)$, $i=1, 2, ..., k$, $Y_{cp}(q)=y_{cp}(t)=y_{cp}(qT)$, определенным на отрезке [0;1]. В качестве показателя разброса представляется полезным рассмотреть то или иное отклонение совокупности функций $Y_i(q)$, $i=1, 2, ..., k$, друг относительно друга. Можно сказать, что эти функции заполняют некую «трубку», которая тоньше всего при истинном значении периода T, а внутри нее проходит периодическая составляющая $X(q)=x(t)=x(qT)$. Рассмотрим различные функционалы интегрального типа. Например, можно проинтегрировать максимум модулей попарных разностей:

$$F_4(T,Y)=\int_0^1 \max\{|\,Y_i(q)-Y_j(q)\,|, i, j=1,2,\ldots,k\}dq.$$

Вместо максимума можно проинтегрировать сумму:

$$F_5(T,Y)=\int_0^1 \sum_{i,j=1}^{k}|\,Y_i(q)-Y_j(q)\,|\,dq.$$

Как и для функционалов супремумного типа рассмотрим показатели разброса относительно «средней функции»:

$$F_6(T,Y) = \int_0^1 \max\{|Y_i(q) - Y_{cp}(q)|, i = 1,2,\ldots,k\} dq,$$

$$F_7(T,Y) = \int_0^1 \sum_{i=1}^{k} |Y_i(q) - Y_{cp}(q)| \, dq.$$

Следующие четыре функционала, используемые как показатели разброса, аналогичны четырем предыдущим, но включают в себя расчет квадратов:

$$F_8(T,Y) = \int_0^1 [\max\{|Y_i(q) - Y_j(q)|, i, j = 1,2,\ldots,k\}]^2 dq,$$

$$F_9(T,Y) = \int_0^1 \sum_{i,j=1}^{k} \{Y_i(q) - Y_j(q)\}^2 dq,$$

$$F_6(T,Y) = \int_0^1 [\max\{|Y_i(q) - Y_{cp}(q)|, i = 1,2,\ldots,k\}]^2 dq,$$

$$F_7(T,Y) = \int_0^1 \sum_{i=1}^{k} \{Y_i(q) - Y_{cp}(q)\}^2 dq.$$

Список показателей разброса можно существенно расширить. В частности, естественно использовать также расстояния в функциональных пространствах L^p при произвольных $p \geq 1$. А для оценивания периодической составляющей применять не только среднее арифметическое, но и другие виды средних величин.

Показатели размаха также можно ввести самыми различными способами. Например, можно рассмотреть такой показатель:

$$G_1(T,Y) = \sup\{|y_{cp}(t) - y_{cp}(s)|, \ 0 \leq t \leq T, \ 0 \leq s \leq T\} =$$
$$= \sup\{y_{cp}(t), \ 0 \leq t \leq T\} - \inf\{y_{cp}(t), \ 0 \leq t \leq T\}.$$

Пусть сделана замена переменной $q=t/T$ и осуществлен переход к функции $Y_{cp}(q)=y_{cp}(t)=y_{cp}(qT)$. Возможными показателями размаха являются:

$$G_2(T,Y) = \int_0^1 \int_0^1 |Y_{cp}(q) - Y_{cp}(r)| \, dq dr,$$

$$G_2(T,Y) = \int_0^1 \int_0^1 (Y_{cp}(q) - Y_{cp}(r))^2 \, dq dr.$$

Введем среднее значение оценки периодической составляющей:

$$Y_{cp} = \int_0^1 Y_{cp}(q)dq.$$

К естественным показателям размаха относятся, например, такие:

$$G_4(T,Y) = \sup\{|Y_{cp}(q) - Y_{cp}|,\ 0 \le q \le 1\},$$

$$G_5(T,Y) = \int_0^1 |Y_{cp}(q) - Y_{cp}|\,dq,$$

$$G_6(T,Y) = \int_0^1 (Y_{cp}(q) - Y_{cp})^2\,dq.$$

Список показателей размаха, как и список показателей разброса, можно значительно расширить. В частности, естественно использовать расстояния в функциональных пространствах L^p при произвольном $p \ge 1$. А для оценивания периодической составляющей применять не только среднее арифметическое, но и другие виды средних — медиану, среднее геометрическое и др. Вопрос о выборе наилучших (в каком-либо смысле) показателей размаха и разброса здесь не обсуждается. Некоторые из причин этого отказа от оптимизации системы показателей рассмотрены ниже.

Алгоритмы оценивания. С прикладной точки зрения остается численно минимизировать один или несколько из 66 описанных выше функционалов $F_i(T;Y)/G_j(T;Y)$, $i = 1, 2, ..., 11$, $j = 1, 2, ..., 6$.

Численная минимизация по одному параметру (возможной длине периода) для современных ЭВМ не вызывает проблем, даже если попросту перебирать возможные значения периода с шагом 0,001. По нескольким реальным или смоделированным сигналам можно установить, какой из функционалов позволяет оценить период и периодическую составляющую реально встречающихся сигналов наиболее точно. Возможно и одновременное использование всех или части функционалов, что в соответствии с методологией устойчивости (см. главу 14) позволяет установить чувствительность оценок к выбору метода оценивания, найти интервал их разброса. Проведенные в Институте высоких статистических технологий и эконометрики расчеты по реальным и смоделированным данным о временных рядах показали, что описанные выше алгоритмы позволяют оценивать длину периода и восстанавливать периодическую составляющую временного ряда достаточно точно с практической точки зрения.

В обширной литературе по временным рядам (см., например, монографии [5–7], дающие представление обо всем массиве литературы по этой тематике) проблеме оценивания периода не уделяется большого внимания. Фактически рекомендуют пользоваться либо периодограммой, либо автокорреляционной функцией. С помощью периодограммы (несостоятельной оценки спектральной плотности) можно выделить лишь синусоидальные составляющие, в то время как в кратко рассмотренных выше прикладных задачах периодическая составляющая представляет интерес сама по себе, без разложения на гармоники. Вторая рекомендация более полезна. В качестве оценки периода можно взять наименьшее положительное число, в котором достигается локальный максимум автокорреляционной функции. Эмпирический коэффициент автокорреляции — еще один функционал типа тех, что перечислены выше.

При поверхностном взгляде на проблемы статистического оценивания, как и на иные проблемы прикладной математики, часто возникает желание обсудить «оптимальность» тех или иных процедур. При более глубоком анализе становятся очевидными два обстоятельства. Во-первых, оптимальность имеет быть лишь в рамках той или иной теоретической модели, при отклонениях от которой оптимальность оценки, как правило, пропадает. Например, выборочное среднее арифметическое как оценка математического ожидания случайной величины оптимальна тогда и только тогда, когда распределение результатов наблюдений — гауссово (доказательство этого утверждения приведено в монографии [11]). С другой стороны, для практически любой статистической процедуры можно подобрать свойство оптимальности так, чтобы эта процедура оказалась оптимальной (как подобрать — это уже дело профессионала). Так, например, метод наименьших модулей оптимален, если погрешности имеют распределение Лапласа, а метод наименьших квадратов — когда их распределение гауссово. Поскольку реальные распределения — не Лапласа и не Гаусса, то указанные математические результаты не могут иметь большого практического значения.

Однако представляется полезным получить доказательства состоятельности оценок изучаемых параметров в возможно более широких, например, непараметрических, постановках. Хотя на основе самого факта сходимости нельзя оценить близость оценок к интересующим исследователя параметрам, но получение доказательства состоятельности — первый шаг при изучении скорости сходимости (см. парараф 14.6).

Состоятельность оценок. Наиболее общий подход к установлению асимптотического поведения решений экстремальных статистических задач развит в статистике нечисловых данных для случая пространств произвольной природы (см. главу 8 монографии [7] и раздел 3.3). Согласно этому подходу сначала при фиксированном T доказывается сходимость (по вероятности) при $A \to \infty$ значений функционала (показателя разброса) к некоторой предельной функции, а затем проверяются условия, обеспечивающие сходимость *Argmin* допредельного случайного процесса к *Argmin* этой детерминированной функции.

Свойства алгоритмов приходится изучать в рамках тех или иных вероятностно-статистических моделей. Моделей может быть много. Достаточно вспомнить историю Центральной Предельной Теоремы (ЦПТ) теории вероятностей. Она на протяжении более 200 лет доказывалась во все более и более широких условиях, вплоть до необходимых и достаточных условий Линдеберга — Феллера (после чего начались обобщения на зависимые слагаемые, на суммы случайных элементов гильбертовых пространств и др.). Отметим, что иногда математические модели далеко выходят за пределы, достаточные для обоснования алгоритмов анализа реальных данных. Так, почти всегда распределения реальных величин дискретны и финитны, а потому, в частности, существуют все моменты. Однако условия финитности и дискретности в вероятностно-статистических моделях часто необоснованно ослабляются. В результате возникают проблемы, не имеющие отношения к реальным данным, например, связанные с измеримостью относительно тех или иных сигма-алгебр. Поэтому ограничимся здесь наиболее простыми моделями из адекватных реальным постановкам. Считаем, что читатель знаком с основными определениями, относящимися к теории случайных процессов.

Теорема 1. Пусть случайный процесс $e(t)$ имеет нулевое математическое ожидание, является стационарным и эргодическим (т.е. выполнена теорема Биркгофа — Хинчина) с непрерывными траекториями. Тогда при фиксированном T и $A \to \infty$ имеем:

$$\sup\{|E_{cp}(q)|,\ 0 \le q \le 1\} \to 0$$

(сходимость по вероятности), где $E_{cp}(q) = Y_{cp}(q) - X_{cp}(q)$, т.е. $E_{cp}(q)$ — среднее арифметическое погрешностей $e(qT)$, $e(qT+T)$, $e(qT+2T)$,...

Доказательство теоремы 1 проводится стандартными методами теории стационарных временных рядов (с шагом T) с использованием известного условия достаточно быстрого убывания элементов матрицы Лорана по мере удаления от ее главной диагонали (т.е. условия, необходимого и достаточного для справедливости теоремы Биркгофа — Хинчина). С помощью теоремы 1 можно найти асимптотику введенных выше показателей разброса и размаха.

Теорема 2. В предположениях теоремы 1 при фиксированном T и $\mathrm{A} \to \infty$ пронормированные показатели разброса $F_i(T;Y)$ для наблюдаемого сигнала $У$ сближаются по распределению с соответствующими положительными случайными величинами $\mathrm{W}_i(\mathrm{T},\mathrm{X},\omega)$, зависящими от T, характеристик случайного процесса $e(t)$ и периодической составляющей X, т.е. существуют числовые последовательности $s_i(k)$ такие, что:

$$s_i(k)F_i(T,Y) \Rightarrow W_i(T,X,\omega), \quad i = 1,2,\ldots,11.$$

Доказательство теоремы 2 проводится с помощью достаточно трудоемких (в частности, из-за числа функционалов), но стандартных рассуждений. Они относятся к теории случайных процессов как части теории вероятностей. Эти рассуждения посвящены максимумам (не супремумам, т.к. траектории функции $x(t)$ и случайного процесса $e(t)$ непрерывны) случайных процессов и интегралам от них, с использованием принципа инвариантности (см., например, учебное пособие [12]) и ряда результатов теории стационарных случайных процессов (см., например, монографию [13]). Таким образом, пронормированные функционалы разброса асимптотически не зависят от числа слагаемых — в этом и состоит основной смысл теоремы 2.

Теорема 3. В предположениях теоремы 1 при фиксированном T и $\mathrm{A} \to \infty$ показатели размаха для наблюдаемого сигнала $У$ сближаются с соответствующими показателями для периодической составляющей X:

$$G_j(T,Y) - G_j(T,X) \to 0, \quad j = 1,2,\ldots,6.$$

Для доказательства используются стандартные оценки, основанные на виде конкретных функционалов, задающих показатели размаха. В отличие от теоремы 2 предельные показатели детерминированы.

Аналоги теорем 2 и 3 верны также и при использовании (в качестве показателей разброса и размаха) расстояний в функциональных пространствах L^p при произвольном $p \geq 1$. А для оценивания периодической составляющей — не только среднего арифметического, но и других видов средних — медианы, среднего квадратического, среднего геометрического, обобщенных средних по Колмогорову (см. раздел 11.4) и др.

Теорема 4. Пусть выполнены условия теоремы 1, периодическая составляющая непрерывна и имеет период T_0. Тогда при фиксированном T и $A \to \infty$ показатели разброса (пронормированные) и размаха стремятся к некоторым детерминированным пределам, зависящим только от T и T_0 , т.е.

$$s_i(k)F_i(T,Y) \to F_i(T,T_0), \quad i = 1,2,...,11,$$

$$G_j(T,Y) \to G_j(T,X), \quad j = 1,2,...,6.$$

(сходимость по вероятности), минимум каждой из функций $F_i(T;T_0)$, i=1, 2, ..., 11, и максимум каждой из функций $G_j(T;T_0)$, j=1, 2, ..., 6, достигается при $T=T_0$ и при T, кратных T_0, причем у показателей разброса $F_i(T;T_0)$ возможны и иные минимумы, а у показателей размаха $G_j(T;T_0)$ других максимумов нет.

Доказательство вытекает из теорем 2 и 3 и свойств усреднения периодической составляющей при росте длины интервала наблюдения сигнала, описанных в начале настоящего подраздела. Отметим, что предельные значения функционала разброса $F_i(T;T_0)$, вообще говоря, показывают разброс случайной погрешности, другими словами, не всегда зависят от периодической составляющей, а потому из-за нормировки на единичный отрезок в ряде случаев оказываются константами. Вместе с тем численные эксперименты показывают, что отмеченная сходимость к пределу является сравнительно медленной. И минимизация непосредственно функционалов разброса (без учета показателей размаха) при конкретной длине сигнала позволяет достаточно точно выделить периодическую составляющую из массива реальных данных. Однако описанные выше теоретические результаты заставили отказаться от первоначальной гипотезы о том, что достаточно использовать только показатели разброса, и привели к необходимости скорректировать алгоритмы, введя деление на показатели размаха.

Теорема 5. В предположениях теоремы 4 оценки, являющиеся первыми локальными минимумами при минимизации по T отношений одного из 11 перечисленных выше показателей разброса к одному из 6 показателей размаха, являются состоятельными оценками истинного периода T_0, а функция $y_{cp}(t)$ является состоятельной оценкой периодической составляющей $x(t)$ на отрезке $[0;T_0]$.

Согласно теоремам 1–4 установлена сходимость (по вероятности) значений допредельных функционалов к предельным при каждом конкретном T. Поэтому для доказательства сходимости минимумов допредельных функционалов к минимумам предельных можно воспользоваться общей теорией асимптотического поведения решения экстремальных статистических задач. Условие асимптотической равномерной разбиваемости, сформулированное в работе [14], выполнено, как можно показать, в силу непрерывности траекторий случайного процесса (непрерывного сглаживания для временного ряда) и его периодической составляющей. Откуда и вытекает заключение теоремы 5, дающей теоретико-статистическое обоснование использованию системы описанных выше эвристических алгоритмов оценивания длины периода и периодической составляющей. При известной или достаточно точно оцененной длине периода сама периодическая составляющая естественным образом оценивается с помощью усреднения перенесенных к началу координат кусков временного ряда, и в силу теоремы 1 эта оценка является состоятельной. Затем для получения оценки математического ожидания сигнала на всей области его определения указанную оценку можно периодически продолжить.

Замечание. При практическом использовании рассматриваемых алгоритмов целесообразно учитывать дополнительные особенности реальных временных рядов. В частности, обратим внимание на неустойчивость супремумов по отношению к выбросам (резко выделяющимся наблюдениям) сравнительно с функционалами интегрального типа. Бывают ситуации, когда методики или аппаратура, регистрирующие значения реальных временных рядов, могут допускать сбои в отдельные моменты времени. Например, если происходит валютный кризис типа «черного вторника», когда курс доллара по отношению к рублю, строго говоря, не определен, другими словами, с точки зрения экономических агентов одновременно существует масса сильно отличающихся курсов. Аналогичная ситуация бывает и в целом ряде других случаев. Набор подходящих ассоциаций вызывают решения руководства страны об обмене денежных знаков, особенно с дискриминационными составляющими. Во всех подобных ситуациях временные ряды дают резкие выбросы (всплески), которые затем, как правило, сглаживаются. Поэтому целесообразно в качестве показателей разброса и размаха использовать функционалы интегрального типа. Вопросам оценивания длины периода и периодической составляющей посвящены многие публикации, в том числе работа [15].

Литература

1. *Елисеева, И.И.* Общая теория статистики / И.И. Елисеева, М.М. Юзбашев. — Москва : Финансы и статистика, 1998. — 368 с.

2. Общая теория статистики. Статистическая методология в изучении коммерческой деятельности / под редакцией А.А. Спирина, О.Э. Башиной. — Москва : Финансы и статистика, 1994. — 296 с.

3. *Доугерти, К.* Введение в эконометрику / К. Доугерти. — Москва : МГУ, 1999. — 402 с.

4. *Нейлор, Т.* Машинные имитационные эксперименты с моделями экономических систем / Т. Нейлор. — Москва : Мир, 1975. — 500 с.

5. *Кендалл, М.Дж.* Многомерный статистический анализ и временные ряды / М.Дж. Кендалл, А. Стьюарт. — Москва : Наука, 1976. — 736 с.

6. *Кендэл, М.* Временные ряды / М. Кендэл. — Москва : Финансы и статистика, 1981. — 199 с.

7. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 2-е изд., испр. и доп. — Москва : Экзамен, 2003. — 576 с.

8. *Петров, В.М.* Цикличность социокультурной сферы и проблемы среднесрочного прогнозирования ее развития / В.М. Петров, Л.А. Мажуль // Математическое и компьютерное моделирование в науках о человеке и обществе : тезисы докладов Всероссийской конференции. — Москва : Изд-во Государственного университетак управления, 1999. — С. 63–66.

9. *Николаев, А.В.* Структура исторического цикла / А.В. Николаев // Математическое и компьютерное моделирование в науках о человеке и обществе. Тезисы докладов Всероссийской конференции. — Москва : Изд-во Государственного университетак управления, 1999. — С. 54–54.

10. *Носовский, Г.В.* Введение в новую хронологию. (Какой сейчас век?) / Г.В.Носовский, А.Т. Фоменко. — Москва : КРАФТ+ЛЕАН, 1999. — 768 с.

11. *Каган, А.М.* Характеризационные задачи математической статистики / А.М. Каган, Ю.В. Линник, С.Р. Рао. — Москва : Наука, 1972. — 656 с.

12. *Биллингсли, П.* Сходимость вероятностных мер / П. Биллингсли. — Москва : Наука, 1977. — 352 с.

13. *Крамер, Г.* Стационарные случайные процессы / Г. Крамер, М. Лидбеттер. — Москва : Мир, 1969.

14. *Орлов, А.И.* Асимптотика решений экстремальных статистических задач / А.И. Орлов // Анализ нечисловых данных в системных исследованиях : сборник трудов. — Вып. 10. — Москва : ВНИИСИ, 1982. — С. 4–12.

15. *Орлов, А.И.* Метод оценивания длины периода и периодической составляющей сигнала / А.И. Орлов // Статистические методы оценивания и проверки гипотез : межвузовский сборник научных трудов. — Пермь : Изд-во Пермского государственного университета, 1999. — С. 38–49.

16. *Орлов, А.И.* Отечественная научная школа в области эконометрики / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 121. — С. 235–261.

17. *Орлов, А.И.* Отечественная научная школа в области организационно-экономического моделирования, эконометрики и статистики / А.И. Орлов // Контроллинг. — 2019. — № 73. — С. 28–35.

18. Исследование метода непараметрических оценок длины периода и периодической составляющей сигнала / В.В. Марченко, О.Г. Берестнева, Д.В. Девятых, Е.Ф. Суханова // Известия Томского политехнического университета. — 2013. — Т. 322. — № 5. — С. 55–59.

19. *Орлов, А.И.* Непараметрические оценки циклов / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 115. — С. 183 — 201.

20. *Луценко, Е.В.* Асимптотический информационный критерий качества шума / Е.В. Луценко, А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 116. — С. 1569–1618.

21. *Орлов, А.И.* Восстановление зависимости методом наименьших квадратов на основе непараметрической модели с периодической составляющей / А.И. Орлов // Научный журнал КубГАУ. — 2013. — № 91. — С. 133–162.

Контрольные вопросы

1. Как соотносятся регрессионный анализ и статистика временных рядов?

2. Чем гетероскедастичные модели рядов динамики отличаются от гомоскедастичных моделей?

3. Опишите различные статистические модели временных рядов.

4. Чем эндогенные переменные отличаются от экзогенных переменных?

5. Какую роль играют показатели разброса и размаха при оценивании длины периода и периодической составляющей?

6. Опишите общую схему оценивания длины периода и периодической составляющей.

Темы докладов, рефератов, исследовательских работ

1. Методы сглаживания временных рядов.

2. Авторегрессионные модели.

3. Различные варианты метода наименьших квадратов при решении систем эконометрических уравнений.

4. Спектральная теория временных рядов.

5. Составьте подробные доказательства теорем, сформулированных в разделе 7.3.

6. Разработайте компьютерную программу, реализующую алгоритмы, описанные в разделе 7.3.

7. Проведите численное изучение скорости сходимости оценок раздела 7.3.

ЧАСТЬ 3. ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Сначала в главе 8 введем основные понятия и рассмотрим примеры вероятностно-статистического моделирования в демографии, логистике, истории и электротехнике. Затем разберем модели динамики (временных рядов), предназначенные для оценки результатов взаимовлияний факторов (глава 9), вероятностно-статистические модели управления качеством (глава 10), проведения экспертных исследований (глава 11), используемые в медицине (глава 12) и социологии (глава 13). Весьма поучительно сопоставить вероятностно-статистические модели, применяемые в различных областях, обнаружить их близость и вместе с тем констатировать некоторые различия.

ГЛАВА 8. ОСНОВЫ ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

8.1. Основные понятия теории статистического моделирования

Начнем с определения используемых понятий.

Модель в общем смысле (обобщенная модель) есть создаваемый с целью получения и (или) хранения информации специфический объект (в форме мысленного образа, описания знаковыми средствами либо материальной системы), отражающий свойства, характеристики и связи объекта-оригинала произвольной природы, существенные для задачи, решаемой субъектом [1, с. 44]. Модели часто выражаются словами или формулами, алгоритмами и иными математическими средствами.

Математические модели. При более тщательном анализе словесных моделей, как правило, не достаточно. Необходимо применение достаточно сложных математических моделей. Так, при принятии решений в менеджменте производственных систем используются:

- модели технологических процессов (прежде всего модели контроля и управления);
- модели обеспечения качества продукции (в частности, модели оценки и контроля надежности);
- модели массового обслуживания;
- модели управления запасами (модели логистики);
- имитационные и эконометрические модели деятельности предприятия в целом, и др.

Часто используют имитационные модели и системы. Имитационная модель позволяет отвечать на вопрос: «Что будет, если…» Имитационная система — это совокупность моделей, имитирующих протекание изучаемого процесса, объединенная со специальной системой вспомогательных программ и информационной базой, позволяющих достаточно просто и оперативно реализовать вариантные расчеты [2, с. 213].

Прежде чем начать рассматривать конкретные математические модели, необходимо определить основные термины, такие, как:

- <u>компоненты системы</u> — части системы, которые могут быть вычленены из нее и рассмотрены отдельно;
- <u>независимые переменные</u> — внешние величины, не зависящие от проходящих в системе процессов;
- <u>зависимые переменные</u> — значения этих переменных есть результат (функция) воздействия на систему независимых внешних переменных;
- <u>управляемые (управляющие) переменные</u> — те, значения которых могут изменяться (в частности, задаваться) исследователем;
- <u>эндогенные переменные</u> — их значения определяются в ходе деятельности компонент системы (т.е. «внутри» системы);
- <u>экзогенные переменные</u> — определяются либо исследователем, либо извне, т.е. в любом случае действуют на систему извне.

При построении любой математической модели желательно придерживаться следующего плана действий:

1) сформулировать цели изучения системы;

2) выбрать те факторы, компоненты и переменные, которые являются наиболее существенными для данной задачи;

3) учесть тем или иным способом посторонние, не включенные в модель факторы;

4) осуществить оценку результатов, проверку модели, оценку полноты модели.

Модели можно делить на различные виды:

1) функциональные модели — выражают прямые зависимости между эндогенными и экзогенными переменными;

2) модели, заданные с помощью систем уравнений относительно эндогенных величин. Выражают балансовые соотношения между различными экономическими показателями (например, модель межотраслевого баланса);

3) модели оптимизационного типа. Основная часть модели — система уравнений относительно эндогенных переменных. Но цель — найти оптималь-

ное решение для некоторого показателя (например, найти такие величины ставок налогов, чтобы обеспечить максимальный приток средств в бюджет за заданное время);

4) имитационные модели — весьма точное отображение явления. Математические уравнения при этом могут содержать сложные, нелинейные, стохастические зависимости.

С другой стороны, модели можно делить на управляемые и прогнозные. Управляемые модели отвечают на вопрос: «Что будет, если ...?»; «Как достичь желаемого?», и содержат три группы переменных: 1) переменные, характеризующие текущее состояние объекта; 2) управляющие воздействия — переменные, влияющие на изменение этого состояния и поддающиеся целенаправленному выбору; 3) исходные данные и внешние воздействия, т.е. параметры, задаваемые извне, и начальные параметры.

В прогнозных моделях управление не выделено явно. Они отвечают на вопросы: «Что будет, если все останется по-старому?»

Далее, модели можно делить по способу измерения времени на непрерывные и дискретные. В любом случае, если в модели присутствует время, то модель называется динамической. Чаще всего в моделях используется дискретное время, т.к. информация поступает дискретно: отчеты, балансы и иные документы приходят периодически. Но с формальной точки зрения непрерывная модель может оказаться более простой для изучения. Отметим, что в физической науке продолжается дискуссия о том, является ли реальное физическое время непрерывным или дискретным.

Обычно в достаточно крупные социально-экономические модели входят материальный, финансовый и социальный разделы. Материальный раздел — балансы продуктов, производственных мощностей, трудовых, природных ресурсов. Это раздел, описывающий основополагающие процессы, это уровень, обычно слабо подвластный управлению, особенно быстрому, поскольку весьма инерционен.

Финансовый раздел содержит балансы денежных потоков, правила формирования и использования фондов, правила ценообразования и т.п. На этом уровне можно выделить много управляемых переменных. Они могут быть регуляторами. Социальный раздел содержит сведения о поведении людей. Этот раздел вносит в модели принятия решений много неопределенностей, поскольку трудно точно правильно учесть такие факторы как трудоотдача, структура потребления, мотивация и т.п.

При построении моделей, использующих дискретное время, часто применяют вероятностно-статистические методы эконометрики. Среди них популярны регрессионные уравнения и их системы. Обычно используют уравнения не выше второго порядка, линейные по параметрам:

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{j=1}^{k} \sum_{f=1}^{k} \beta_{jf} x_{ij} x_{if} + e_i \ ,$$

где Y_i — переменная отклика; x_{ij} — факторы, от которых зависит Y_i; β_j — коэффициенты, которые характеризуют взаимодействие между Y_i и x_{ij}; β_{jf}, — отражают взаимодействие между x_{ij} и x_{if}; e_i — ошибка модели; i — номер наблюдения (измерения, опыта, анализа, испытания), $i = 1, 2, \ldots, n$; j — номер фактора (независимой переменной), $j = 1, 2, \ldots, k$.

Коэффициенты β_j, β_{jf}, находят эконометрическими методами (см. [3]), например, методом наименьших квадратов. Различные системы регрессионных уравнений, построенные для решения практически важных задач, рассмотрены в [4]. Часто используют лаги (запаздывания в реакции). Для систем, нелинейных по параметрам, применение метода наименьших квадратов встречает трудности [3]. Большое количество конкретных эконометрических и математических моделей принятия решений рассмотрено в [4, 5]. Обратим внимание, что популярные в настоящее время подходы к процессам бизнес-реинжиниринга основаны на активном использовании математических и информационных моделей.

Развитие методов математического моделирования. Эта область научно-практической деятельности получила мощный стимул к развитию во время и сразу после второй мировой войны в рамках интеллектуального движения, связанного с терминами «кибернетика», «исследование операций», а позже — «системный анализ», «информатика».

Впрочем, имелась и вполне практическая задача — контроль качества боеприпасов, вышедшая на первый план именно в годы второй мировой войны. Методы статистического контроля качества приносят (по западной оценке, обсуждаемой в [6], и по нашему мнению, основанному на опыте СССР и России, в частности, на анализе организационно-экономических результатов работы служб технического контроля на промышленных предприятиях) наибольший экономический эффект среди всех экономико-математических методов принятия решений. Только дополнительный доход от их применения в промышленности США оценивается как 0,8 % валового национального продукта США.

Для ориентации в практически необозримом море математических моделей экономических явлений и процессов (короче: экономико-математических моделей), необходима их классификация. Первым основанием для классификации служит отношение к практической деятельности. Экономико-математические модели делятся на:

1) ориентированные на практическое использование (примерами служат модели статистического контроля, с помощью которых принимается решение о приемке или забраковании партии конкретной продукции),

2) модели, которые пригодны для теоретических рассуждений, однако в практике их использовать невозможно (примерами служат модели «основного уравнения количественной теории денег» или «спирали ЦЕНЫ — ЗАРПЛАТА» в учебнике макроэкономики [7]).

Экономико-математическое моделирование. Отметим большое практическое значение моделей логистики или, в другой терминологии, управления запасами (см. раздел 8.3). В последние годы интерес вызывает моделирование финансового рынка.

Важная проблема — учет неопределенности. Основное место она занимает в вероятностно-статистических моделях экономических и социально-экономических явлений и процессов. Проблемы устойчивости (к допустимым отклонениям исходных данных и предпосылок модели) для социально-экономических моделей рассматриваются в [5].

Особое место занимают имитационные системы, позволяющие отвечать на вопросы типа: «Что будет, если...?» (Как подчеркнуто в [2, с. 212], «любая модель, в принципе, имитационная, ибо она имитирует реальность».) Основа имитации (смысл которой мы будем понимать как анализ экономического явления с помощью вариантных расчетов) — это математическая модель. Согласно [2, с. 213] имитационная система — это совокупность моделей, имитирующих протекание изучаемого процесса, объединенная со специальной системой вспомогательных программ и информационной базой, позволяющих достаточно просто и оперативно реализовать вариантные расчеты. Таким образом, под имитацией понимается численный метод проведения машинных экспериментов с математическими моделями, описывающими поведение сложных систем в течение продолжительных периодов времени [4, с. 9], при этом имитационный эксперимент состоит из следующих 6 этапов:

1) формулировка задачи;

2) построение математической модели;

3) составление программы для ЭВМ;

4) оценка пригодности модели;

5) планирование эксперимента;

6) обработка результатов эксперимента.

Имитационное моделирование (simulation modelling) широко применяется в различных областях, в том числе в экономике [4]. Наиболее перспективным представляется синтез экспертных систем и математических моделей, впервые осуществленный в нашей стране еще в 1970-е гг.

При построении, изучении и применении экономико-математических моделей принятия решений используются различные математические методы, именуемые в данном контексте экономико-математическими (хотя они, как правило, могут с успехом использоваться вне экономики, как, в частности, эконометрические методы анализа эмпирических экономических данных). Экономико-математические методы можно разделить на несколько групп:

- методы оптимизации;
- методы, учитывающие неопределенность, прежде всего, вероятностно-статистические;
- методы построения и анализа имитационных моделей;
- методы анализа конфликтных ситуаций (теории игр).

Во всех этих группах можно выделить статическую и динамическую постановки. При наличии фактора времени используют дифференциальные уравнения и разностные методы.

Рассмотрим перечисленные группы методов по отдельности.

Методы оптимизации. Со времен классических работ нобелевского лауреата по экономике академика АН СССР Л.В. Канторовича один из основных классов экономико-математических методов — это методы оптимизации. Оптимальному управлению на основе экономико-математических моделей посвящена обширная литература, в ней используются такие термины, как оптимальное программирование и оптимальное планирование. В случае одного критерия принципиальных сложностей нет — применяют диалоговые компьютерные системы. Сложные проблемы — это выбор целевых функций, оценка устойчивости свойств оптимальности [5], многокритериальность. Для построения моделей с целью принятия решений используют теорию полезности.

Вероятностно-статистические модели. Исходная научная база таких моделей — теория вероятностей и математическая статистика. Выделяют как самостоятельное направление прикладную статистику. Она включает в себя прикладную математическую статистику, ее программное обеспечение и методы сбора статистических данных и интерпретации результатов расчетов. Только первая из этих трех областей одновременно входит и в математическую статистику. Последняя включает в себя также чисто математическую область, в которой статистические структуры рассматриваются как математические объекты. Они изучаются внутриматематическими методами. Эту область научных исследований в ряде публикаций называют «аналитической статистикой». Таким образом, математическая статистика состоит из прикладной математической статистики, ориентированной на практическое применение, и ветви чистой математики под названием «аналитическая статистика», полезность которой для применений не подтверждена. Можно всю жизнь доказывать теоремы в аналитической статистике, ни разу не обработав реальные данные и даже не думая об этом. В настоящее время аналитическая статистика постепенно вытесняет прикладную математическую статистику из научных журналов и учебных курсов. Так, в основном в России журнале по теории вероятностей и математической статистике «Теория вероятностей и ее применения» уже почти не встретишь статей, имеющих отношение к работе с реальными данными.

Статистические методы активно применяются в различных областях экономики, причем в России — уже более 150 лет. Как известно, эконометрика (или эконометрия) — это статистические методы анализа эмпирических экономических данных [4]. Однако в XX в. этот термин употреблялся в нашей стране почти исключительно в переводной литературе.

Имеются многочисленные работы по различным конкретным разделам прикладной статистики и эконометрики:

- по регрессионному анализу (методам восстановления зависимости и построения моделей, прежде всего линейных);
- по планированию эксперимента;
- по методам классификации (дискриминантного анализа, кластер-анализа, автоматической классификации, распознавания образов, систематики и типологии, теории группировок);
- по многомерному статистическому анализу экономической информации;
- по методам анализа и прогнозирования временных рядов;
- по теории робастности (robustness), т.е. устойчивости статистических процедур к допустимым отклонениям исходных данных и предпосылок модели;
- по использованию различных индексов, в частности, индекса инфляции.

Основной журнал в России, в котором публикуются исследования по прикладной статистике и, особенно, по планированию эксперимента — это «Заводская лаборатория» (секция «Математические методы исследования»).

Теория конфликтных ситуаций (теория игр). Теория игр (более подходящее название — теория конфликта, или теория конфликтных ситуаций) зародилась как теория рационального поведения двух игроков с противоположными интересами. Она наиболее проста, когда каждый из них стремится минимизировать свой средний проигрыш, т.е. максимизировать свой средний выигрыш. Отсюда ясно, что теория игр склонна излишне упрощать реальное поведение в ситуации конфликта. Участники конфликта могут оценивать свой риск по иным критериям. В случае нескольких игроков возможны коалиции. Большое значение имеет устойчивость точек равновесия и коалиций.

В экономике еще 150 лет назад теория дуополии (конкуренции двух фирм) О. Курно была развита на основе соображений, которые мы сейчас относим к теории игр. Новый толчок дан классической монографией Дж. фон Неймана и О. Моргенштейна [8], вышедшей вскоре после второй мировой войны. В учебниках по экономике обычно разбирается «дилемма заключенного» и точка равновесия по Нэшу (ему присуждена Нобелевская премия по экономике за 1994 г.).

По теории игр имеется обширная литература, часть из которой непосредственно адресована экономистам. Однако в практической работе теория игр почти не используется. Если же это происходит, то она обычно выступает как часть более широкого подхода, ассоциированного с терминами «принятие решений», «конфликтная ситуация».

Критика математической экономики. Второе из указанных выше направлений экономико-математического моделирования, т.е. посвященное моделям, которые непосредственно использовать в практической работе невозможно, обычно связывается с термином «математическая экономика». О нем акад. РАН Н.Н. Моисеев писал:

«...Имеется развитое направление исследований, получившее название математической экономики. В работах, относящихся к этому направлению, изучаются свойства математических моделей, построенных на основе формализации некоторых понятий экономической науки, таких как, например, конкурентное равновесие. Используя некоторые предположения о функциональных зависимостях (например, о выпуклости функций и множеств), исследователи анализируют общие свойства моделей — доказывают теоремы о существовании экстремальных значений тех или иных параметров, изучают свойства точек равновесия, траекторий равновесного роста и т.д. Эти исследования содействовали становлению экономико-математических методов, помогали и помогают отточить математические методы, используемые в прикладных исследованиях. Однако с развитием математической экономики рассматриваемые в ней проблемы все более уходили от экономической реальности и становились чисто математическими. В результате этого в настоящее время математическая экономика представляет собой своеобразный раздел математики, изучающий математические конструкции, которые лишь с большой степенью произвола можно назвать экономическими моделями...» (из предисловия к учебному пособию А.В. Лотова «Введение в экономико-математическое моделирование» [9, с. 6]).

Методы математического моделирования реальных экономических явлений и процессов, разумеется, полезны и необходимы, в частности, для успешной работы менеджеров, экономистов и инженеров, как на предприятиях, так и на государственной службе. Но нужны только те математические результаты, которые помогают экономисту в работе, в частности, методы теории принятия решений, эконометрики, прикладной математической статистики, экспертных оценок (в частности, сценарный метод).

Нельзя не согласиться с тем очевидным утверждением, что некоторые теоретические работы, которые в настоящее время не удается связать с практикой, в будущем могут оказаться полезными для решения реальных задач. Лучший пример — история ядерной физики. Однако нельзя не указать на многочисленные монографии и сборники статей, в которых чисто математические рассуждения даны «под экономическим соусом».

Экономико-математические модели иногда используются в качестве «дымовой завесы» для пропаганды сомнительных с научной точки зрения воззрений. В качестве примера рассмотрим некоторые разделы западных курсов «экономикс».

Они построены, естественно, на обобщении западной экономической жизни. Так, потребитель предполагается совершенно рациональным, точно знающим, что он хочет максимизировать (т.е. знающим свою функцию полезности), а также полностью игнорирующим всех остальных потребителей, действующим совершенно самостоятельно. Общество состоит из эгоистичных индивидуумов-атомов, отстаивающих только свои интересы, т.е. живущих по принципу «человек человеку — волк». Законы правового государства удерживают такое общество от самоуничтожения.

Возможно, такая экономико-математическая модель годится для части жителей западных стран, прежде всего США. Бесспорно, что она не годится для нас, для русских. Мы плохо знаем, что нам нужно, действуем под влиянием друзей, общественного мнения, моды, привыкли жить в коллективе, общине, семье, говорим о соборности, игнорируем экономические стимулы. Несмотря на снижение реальных доходов в несколько раз, пока нет бунтов. Хотя значительная часть предприятий стоит, работники не уходят, а менеджеры (директора) их не увольняют. Сейчас зарплата профессора сравнима с зарплатой уборщицы в метро и в несколько раз меньше дохода продавца коммерческого киоска. Но, вопреки монетаристским экономическим теориям, профессора не рвутся в продавцы. И рабочие зачастую выпускают продукцию, не получая зарплату. И потому Россия жива. Западные экономические теории не годятся не только для России. Они не подходят для исламских стран, для Индии и Китая, и т.д.

Нобелевская премия по экономике за 2002 г. была присуждена Вернону Смиту и Даниэлю Кингману за экспериментальное изучение поведения потребителей. Установлено, что лица, чье потребительское поведение может быть описано функцией полезности, реально составляют очень незначительную часть человеческих популяций. Остальные (в части экономики) действуют преимущественно иррационально, точнее, их действия являются рациональными на другом (более высоком) уровне. Подавляющая часть людей вполне осознанно может пойти на ограничение личных потребностей ради общественного блага.

В некоторых публикациях с помощью экономико-математических моделей сознательно вводят читателей в заблуждение. В качестве примера возьмем учебник Р. Лэйарда [7]. В нем «доказывается», что «инфляционный налог» равен дефициту бюджета. Отсюда рекомендация — для снижения инфляции необходимо ограничивать поступление новых денежных масс в оборот (например, не выдавать зарплату). Однако это утверждение выводится в предположении, что суммарный выпуск постоянен, чего не было у нас — до 1997 г. объем производства монотонно падал. При этом Р. Лэйарда отнюдь не смущает, что в другой главе, говоря о «мультипликаторе Кейнса», он рекомендует увеличивать государственные расходы в период спада производства. Принципиально ошибочно рассмотрение Р. Лэйардом спирали «заработная плата — цены», основанное на математической ошибке (функция принимается за константу). Но вывод каков: чтобы снизить инфляцию, надо, якобы, увеличить безработицу!

Необходимо отметить, что название «Математическая экономика» носят и некоторые публикации, лишенные указанных выше недостатков, например, отличный учебник К. Ланкастера [10].

8.2. Демографические модели

Для подготовки и принятия решений в государственном и муниципальном управлении, в маркетинге, в стратегическом менеджменте и в других областях организационно-экономической деятельности необходимы прогнозы численности населения, его полового и возрастного состава.

Возможны два подхода к прогнозированию демографических характеристик. Можно взять временной ряд численности населения и прогнозировать его с помощью того или иного статистического или экспертного метода. При этом формирование модели формирования населения не предполагается. Несмотря на это, из-за большой инерционности демографических процессов такое «внешнее» прогнозирование оказывается полезным, особенно на небольших интервалах прогноза.

«Внешнее» прогнозирование не всегда позволяет разобраться в существе процесса. Рассмотрим, например, такой показатель, как смертность (число умерших на 1 000 человек). Смертность зависит не только от здоровья населения, но и от его полового возрастного состава. Например, в «молодом» обществе, в котором много детей, а потому его численность растет, смертность будет маленькой, хотя старшее поколение может умирать сравнительно рано, в 40–50 лет. Наоборот, в стабилизировавшемся по численности или вырождающемся обществе, в котором много пожилых людей и мало молодежи, смертность будет велика.

Рост (падение) смертности может быть вызван двумя основными причинами — ухудшением (улучшением) социально-экономического положения населения или же изменением половой и возрастной структуры населения. Однако если наблюдается резкий рост смертности, то это — результат действия первой причины, поскольку структура населения меняется медленно. Например, в СССР в результате антиалкогольной компании 1980-х гг. смертность заметно снизилась (относительно прогнозируемой по предыдущим годам зависимости), а в 1990-х гг. резко повысилась, особенно в начале десятилетия, из-за снижения в разы жизненного уровня населения и тяжелой социально-экономической обстановки.

С рождаемостью ситуация сложнее, поскольку на нее влияют не только перечисленные причины, но и изменение стиля жизни. Если сто и более лет назад интересы большинства женщин были ориентированы на семью с большим числом детей (многие из которых умирали во младенчестве), то в течение XX в. значительная часть женской половины человечества стала заниматься прежде чисто мужскими профессиями. Перед ними возник выбор — рождение детей или профессиональное продвижение. В результате появилось понятие «планирование семьи», а рождаемость сократилась.

Правительства различных стран успешно занимались управлением рождаемостью. Для одних стран, таких, как Китай, Индия, целью было сокращение рождаемости. В странах с сильной государственной властью эффективными оказались административные методы. Например, в Китае в городах долгое время семье разрешалось иметь только одного ребенка. За рождением второго следовали санкции — штрафы, повышенная оплата за детские товары, медицинское обслуживание, обучение. Поскольку в европейских странах, таких, как Франция и Германия, падение рождаемости во второй половине XX в. угрожало депопуляцией, то для стимулирования рождаемости применялись такие экономические меры, как выплата значительных пособий на воспитание детей. Во Франции пособия на трех детей превышали среднюю заработную плату по стране. В результате снижение численности коренного населения удалось предотвратить.

Демографические модели. Для более точного демографического прогнозирования, количественного, а не только качественного, необходимо использовать статистические модели динамики половозрастного состава населения. Такие модели основываются на прослеживании судьбы поколения, рожденного в определенный временной промежуток, например, в год t_0. Пусть всего в этот год родилось $N(t_0)$ человек. Далее каждый год их количество будет уменьшаться, пока не достигнет 0.

Через год численность поколения составит $N(t_0+1)$ человек. За этот год умерли $N(t_0) - N(t_0 + 1)$ человек. Частота смерти за первый год составляет $q(1) = (N(t_0) - N(t_0 + 1))/N(t_0)$, а $1 - q(1)$ — частота того, что человек благополучно перейдет в следующую возрастную категорию. Ясно, что $N(t_0 + 1) = N(t_0)(1 - q(1))$.

Через два года численность поколения составит $N(t_0 + 2)$ человек. За второй год умерли $N(t_0 + 1) - N(t_0 + 2)$ человек. Частота смерти за второй год составляет $q(2) = (N(t_0 + 1) - N(t_0 + 2))/N(t_0 + 1)$. При этом $(1 - q(2))$ — частота того, что человек благополучно перейдет в следующую возрастную категорию. Ясно, что $N(t_0 + 2) = N(t_0)(1 - q(1))(1 - q(2))$. И так далее.

Через k лет, k=1, 2, 3, ..., численность поколения составит $N(t_0 + k)$ человек. За k-й год умерли $N(t_0 + (k - 1)) - N(t_0 + k)$ человек. Частота смерти за второй год составляет $q(k) = (N(t_0 + (k - 1)) - N(t_0 + k))/N(t_0 + (k - 1))$. Очевидно, $1 - q(k)$ – частота того, что человек благополучно перейдет в следующую возрастную категорию. Ясно, что $N(t_0 + k) = N(t_0)\,(1 - q(1))(1 - q(2))...(1 - q(k))$.

Возрастные коэффициенты смертности. Здесь $q(1)$, $q(2)$, ... $q(k)$, ... — погодовые возрастные коэффициенты смертности. Это — частоты некоторых событий, т.е. оценки соответствующих вероятностей. Но поскольку наблюдений много, можно не делать различия между частотами и вероятностями. Возрастной коэффициент смертности $q(k)$ есть условная вероятность того, что человек, проживший $(k - 1)$ лет, умрет в течение k-го года. Возрастные коэффициенты смертности рассчитывают не только для всего населения в целом, но отдельно для мужчин и женщин, для различных регионов, социальных групп (например, жителей крупных городов и сельских поселений). Часто используют более длинные промежутки времени, чем год, особенно при выделении поколений. Например, поколением считаются все, родившиеся с 1945 г. по 1954 г. Очевидно, возрастные коэффициенты смертности зависят от начала отсчета t_0 — от «даты рождения» поколения. Прослеживают судьбу поколения в течение 90–100 лет, поскольку в настоящее время лишь мизерные доли процента численности поколения живут дольше.

На возрастные коэффициенты смертности конкретных поколений никак не влияет половозрастная структура всего населения. А вот изменение социально-экономического положения и уровня медицинского обслуживания населения, как показывают статистические наблюдения, отражаются в изменении возрастных коэффициентов смертности. Например, в 1999 г. возрастные коэффициенты смертности российских мужчин и женщин выросли (по сравнению с предыдущим годом) примерно на 10 % для всех поколений от 20 до 60 лет [11, с. 95]. Это наглядно говорит об ухудшении социально-экономической обстановки и условий жизни в 1999 г., что можно связать с двукратным падением уровня жизни среднего жителя России в результате экономического кризиса, последовавшего за дефолтом в августе 1998 г. [3, гл. 8].

Погодовые возрастные коэффициенты смертности $q(1)$, $q(2)$, … $q(k)$, … позволяют рассчитать различные демографические характеристики поколения. Например, вероятность $p(k)$ того, что продолжительность жизни представителя поколения составляет ровно k лет, равна вероятности того, что он переживет $(k–1)$ контрольных сроков и умрет до наступления k-го, т.е. равна:

$$p(k) = (1-q(1))(1-q(2))\ldots(1-q(k-1))q(k). \tag{1}$$

Таким образом, продолжительность жизни моделируется дискретной случайной величиной X с распределением:

$$P(X = k) = p(k), \quad k = 1,2,3,\ldots$$

Средняя продолжительность жизни в рассматриваемом поколении, т.е. математическое ожидание продолжительности жизни, есть

$$MX = \sum_{л=1}^{100} kp(k).$$

Строго говоря, все эти характеристики можно рассчитать только после того, как поколение закончит свой жизненный путь (по крайней мере через сто лет после начала отсчета t_0). Однако судьба ушедшего поколения волнует немногих. И возрастные коэффициенты смертности используются в демографии по-другому.

Базовая модель в демографии может быть описана так: возрастные коэффициенты смертности постоянны и равны значениям, рассчитанным за последний год. Это значит, что демографическая ситуация в год t описывается вероятностями $p_1(k)$, $k = 1, 2, \ldots$, причем $p_1(k) = p(k)$ для поколения, чей жизненный путь начался в году $t - k$. Таким образом, данным 2021 г. вероятность $p_1(1)$ оценивается для рожденных в 2020 г., $p_1(2)$ — для детей 2019 г. рождения, $p_1(3)$ — для поколения 2018 г., …, $p_1(13)$ — для тех, кто родился в 2008 г., …, $p_1(33)$ — для поколения 1988 г., …, $p_1(63)$ — для родившихся в 1958 г., …, $p_1(93)$ — для лиц 1928 г. рождения... Можно сконструировать случайную величину X_1 — продолжительность жизни в предположении постоянства имеющихся в фиксированный момент временных коэффициентов смертности. Ее распределение имеет вид:

$$P(X_1 = k) = p_1(k),\ \ k = 1,2,3,\ldots$$

Средней ожидаемой продолжительностью предстоящей жизни (СОППЖ), или, короче и менее точно, ожидаемой продолжительностью жизни называется:

$$MX_1 = \sum_{л=1}^{100} kp_1(k).$$

Средняя ожидаемая продолжительность предстоящей жизни — наиболее общая характеристика качества жизни. Чем она выше, тем лучше живут люди, тем лучше социально-экономические условия, медицина, обычаи.

Изменения СОППЖ в различные периоды времени имели различные причины. В первой половине XX в. СОППЖ росла в основном за счет снижения детской смертности. Затем наступила очередь снижения потерь в рабочих возрастах (20–60 лет). «Естественная» продолжительность жизни, при отсутствии ранних заболеваний, травм, убийств и самоубийств, сравнительно мало меняется во времени. Есть сведения [12], что она равна 89±5 лет.

Проблемы демографии — одна из наиболее острых тем обсуждений в современном обществе. Чтобы не накалять страсти, будем демонстрировать методические приемы применения демографических данных на примерах прошлых лет. Анализ текущих данных предоставляем читателю.

СОППЖ в России. Приведем взятую из [11, с. 84] табл. 1 значений СОППЖ в России за 1960–1999 гг. Ее анализ показывает, что самым лучшим с точки зрения СОППЖ был 1965 г. Действительно, это был год, когда положение СССР в мире было самым лучшим за весь XX век. Последствия войн ушли в прошлое. Успехи СССР в мирном строительстве, в космосе, в области атомной энергии были общепризнанными. Социально-психологическое состояние населения было наилучшим за век.

По значениям СОППЖ выделяются два периода. В первом, с 1960 г. по 1990 г., СОППЖ для мужчин колеблется между 61,4 (1980 г.) и 64,6 (1964 г.), для женщин между 72,0 (1960 г.) и 74,3 (1990 г.). В дальнейшие годы (1991–1999) все значения СОППЖ для мужчин меньше, чем минимум за годы первого периода. Сначала идет резкое монотонное падение — с 63,8 лет в 1990 г. до 57,6 лет в 1994 г., всего на 6,2 года. Затем медленный подъем до 61,3 лет в 1998 г. И опять спад до 59,9 лет. Для женщин картина аналогичная, но менее выраженная. Сначала идет монотонное падение — с 74,3 лет в 1990 г. до 71,2 лет в 1994 г., всего на 3,1 года. Затем медленный подъем до 72,9 лет в 1998 г. И опять спад до 72,4 лет. Это меньше, чем любое значение за период 1965–1990 гг. Табл. 1 демонстрирует корреляцию изменений СОППЖ и социально-экономического положения населения. Наибольшее падение СОППЖ как для мужчин, так и для женщин приходится на годы наиболее резкого снижения жизненного уровня, а именно, 1992–1994 гг. и 1999 г.

Таблица 1

Средняя ожидаемая продолжительность предстоящей жизни в России

Год	Мужчины	Женщины	Среднегодовой прирост	
			Мужчины	Женщины
1960	63,6	72,0	–	–
1965	64,6	73,4	0,19	0,28
1970	63,1	73,4	– 0,29	0,01
1975	62,5	73,2	– 013	– 005
1980	61,4	72,9	– 0,22	– 0,07
1985	62,8	73,3	0,27	0,10
1990	63,8	74,3	0,21	0,21

Год	Мужчины	Женщины	Среднегодовой прирост	
			Мужчины	Женщины
1991	63,5	74,3	– 0,27	– 0,08
1992	62,0	73,8	– 1,50	– 0,50
1993	58,9	71,9	– 3,09	– 1,92
1994	57,6	71,2	– 1,32	– 0,70
1995	58,3	71,7	0,68	0,52
1996	59,8	72,5	1,48	0,79
1997	60,8	72,9	1,00	0,40
1998	61,3	72,9	0,55	0,04
1999	59,9	72,5	– 1,37	– 0,55

Артефакт или факт? Как и иные данные о социально-экономических явлениях и процессах, демографические данные зачастую пытаются интерпретировать в угоду политическим концепциям. Например, в [11, с. 90] утверждается, что «небывалый в мирное время подъем смертности в России в первой половине 1990-х гг. — артефакт». Значительное снижение СОППЖ (табл. 1), значительное возрастание смертности в рабочих возрастах (см. табл. 2 и [11, с. 87]) и другие факты неопровержимо свидетельствуют, что небывалый в мирное время подъем смертности в России в первой половине 1990-х гг. — достоверный факт.

Таблица 2

Возрастные коэффициенты смертности (число смертей на 1 000 человек) в России в 1991–1998 гг. [13, с. 309]

Возрастные группы	1991	1994	1998	1994/1991(%)
Мужчины				
15–19	1,7	2,1	1,9	123
20–24	2,7	4,0	4,1	148
25–29	3,5	5,5	4.6	157
30–34	4,5	7,7	5,8	171
35–39	5,9	10,6	7,5	180
40–44	8,0	15,2	10,2	190
45–49	11,6	20,8	14,4	179
50–54	16,5	29,1	19,5	176
55–59	23,3	36,2	28,6	155
60–64	34,6	51,0	38,1	136
65–69	47,3	64,2	55,3	136
70 и старше	104,0	121,4	97,0	116
Женщины				
15–19	0,7	0,8	0,8	114
20–24	0,7	1,0	1,0	143

Возрастные группы	1991	1994	1998	1994/1991(%)
25–29	0,9	1,3	1,2	143
30–34	1,1	1,8	1,5	164
35–39	1,6	2,7	2,0	169
40–44	2,5	4,2	2,8	168
45–49	3,8	6,2	4,4	163
50–54	5,5	9,0	6,4	163
55–59	8,6	12,3	10,0	143
60–64	13,6	18,4	14,5	135
65–69	22,0	27,1	24,1	123
70 и старше	78,1	89,6	77,5	115

Как же авторы [11] пытаются его опровергнуть? Во-первых, они пользуются не возрастными коэффициентами смертности, а данными об общем чисел умерших [11, с. 86], пытаясь скрыть потери от «реформ» среди всех умерших, число которых монотонно возрастает в соответствии с общим постарением населения. Во-вторых, они анализируют совместно два последовательных процесса — снижение смертности в 1985–1987 гг. в результате антиалкогольной компании и повышение смертности в 1991–1994 гг. вследствие «реформ». Они показывают, что мужские поколения, родившиеся между 1945 г. и 1984 г., и женские поколения, родившиеся между 1945 и 1974 г. (т.е. основные рабочие возраста), от совместного действия этих дух процессов проиграли, остальные (дети и пожилые люди) несколько выиграли [11, с. 87]. Это говорит только о том, что положительный эффект антиалкогольной компании оказался перекрыт потерями населения в первые годы «реформ», но не может привести к выводу об отрицании этих потерь.

Из двух указанных ошибок первая является основной. Анализ возрастных коэффициентов — основа изучения демографических процессов. Как отмечено вначале пункта, попытки анализировать сводные характеристики непосредственно порождают ошибки.

Рождаемость. При изучении рождаемости необходимо, во-первых, знать численность женских контингентов (поколений) различных возрастов, во-вторых, иметь возрастные коэффициенты рождаемости, например, число рождений на 1000 женщин соответствующего возраста.

Как и при изучении смертности, возможны два подхода к моделированию процесса рождаемости. В первом из них прослеживается судьба женской части поколения. Пусть $N_1(t_0)$ — численность поколения (число девочек, родившихся в год t_0). Число женщин, доживших до k лет, обозначим $N(t_0)p_2(k)$. (Обратите внимание, что вероятности $p_2(k)$ имеют иной смысл, чем при рассмотрении смертности. Они являются аналогом функции распределения, а не плотности.) Число детей, рожденных женщинами этого поколения в k-й год, обозначим $\alpha(k)N(t_0)p_2(k)$. Здесь $\alpha(k)$ — возрастной коэффициент рождаемости.

Во втором подходе возрастные коэффициенты рождаемости для всех поколений принимаются равными, причем равными тем, что имеют быть в последний рассматриваемый год (подробнее этот подход расписан выше при обсуждении сводных характеристик смертности). Пусть $\alpha_1(k)$, $k = 1, 2, \ldots$, — соответствующие возрастные коэффициенты рождаемости, $p_3(k)$ — доли женщин, доживших до k лет (на анализируемый год). Пусть случайная величина Y — число детей, рожденных одной конкретной женщиной. Основная сводная характеристика рождаемости — коэффициент суммарной рождаемости, т.е. среднее число рожденных детей на одну женщину, т.е.

$$M(Y) = \sum_{n=1}^{100} \alpha_1(k) p_3(k).$$

В табл. 3 приведены коэффициенты суммарной рождаемости по годам с разбивкой на городских и сельских жителей [11, с. 40, 44]. Очевидно, в перспективе (в следующем поколении через 20–30 лет) население будет расти, если в среднем на одну женщину приходится более двух детей, и уменьшаться, если менее двух. Если коэффициент суммарной рождаемости около 2,0, то ситуация является пограничной и неустойчивой.

Таблица 3

Коэффициенты суммарной рождаемости в России

Годы	Коэффициенты суммарной рождаемости		
	Город	Село	Всего
1970	1,70	3,38	2,00
1980	1,71	2,50	1,89
1990	1,83	2,63	1,89
1995	1,24	1,84	1,34
1999	1,07	1,48	1,17
2005	1,207	1,576	1,294
2010	1,439	1,983	1,582
2015	1,678	2,111	1,777
2019	1,427	1,754	1,504

Из табл. 3 очевидно, что из пограничного состояния 1970 г. страна перешла в область депопуляции (сокращения численности населения), причем ситуация резко ухудшилась в годы «реформ». Обратим внимание на различие между городом и селом. В 1970–1990 гг. сельское (по месту рождения) население увеличивалось, в то время как город в течение всего рассматриваемого периода не обеспечивал воспроизводство населения. Это различие указывает на принципиальную возможность управления демографическими процессами.

Миграция. Большое значение в изменении численности населения страны играет миграция. Разность между числом прибывших в страну и числом уехавших называется чистой миграцией. Разность между числом родившихся и числом умерших называется естественным приростом. Общий прирост населения есть сумма естественного прироста и чистой миграции. В табл. 4 [11, с. 7] приведены сведения о численности населения России во второй половине XX в. и компонентах ее изменения. Из нее видно, что чистая миграция составляет заметную часть от изменения численности населения.

Таблица 4

Изменение численности населения России

Год	Численность населения в конце года (тыс. чел.)	Изменение численности населения		
		Общий прирост (тыс. чел.)	В том числе	
			Естественный прирост	Чистая миграция
1955	112 266	9 321	9 160	161
1960	120 766	8 500	9 515	– 1015
1965	127 189	6 423	7 067	– 644
1970	130 704	3 515	4 180	– 195
1975	134 690	3 986	4 180	– 195
1980	139 028	4 338	3 730	607
1985	143 835	4 807	3 939	869
1990	148 543	4 707	3 649	1 058
1991	148 704	162	110	52
1992	148 673	– 31	– 207	176
1993	148 366	– 308	– 738	430
1994	148 306	– 60	– 870	810
1995	147 976	– 330	– 832	502
1996	147 502	– 474	– 818	344

Год	Численность населения в конце года (тыс. чел.)	Изменение численности населения		
		Общий прирост (тыс. чел.)	В том числе	
			Естественный прирост	Чистая миграция
1997	147 105	– 398	– 751	353
1998	146 693	– 411	– 696	285
1999	145 925	– 768	– 923	155
2000	145 185	– 740	– 954	214

Замечание. Демографические данные получают в результате учета населения, проводимого государственными органами. Система учета такова, что прибывшие в страну мигранты в дальнейшем «растворяются» среди коренного населения, смертность и рождаемость среди недавних мигрантов не отделяются от таковых среди коренного населения. Это приводит к тому, что возрастные коэффициенты типа $p_1(k)$ рассчитываются с учетом мигрантов. Прослеживание судьбы поколения и нахождение коэффициентов типа $p(k)$ становится невозможным из–за эмиграции части поколения. Если возрастные коэффициенты коренного населения и мигрантов сильно различаются, то искажение итоговых демографических характеристик населения может быть заметным. Однако в российской ситуации начала XXI века указанным различием можно пренебречь.

Прогнозирование демографической ситуации. Считая возрастные коэффициенты смертности и рождаемости постоянными, можно рассчитать прогнозируемое изменение половозрастной структуры населения (при отсутствии миграции).

Обычно прогнозирование проводят, исходя из некоторого числа сценариев развития ситуации. Например, в 2000 г. опубликован официальный прогноз Госкомстата РФ изменения численности структуры населения на 2000–2015 гг. [14]. В нем рассмотрено три сценария — оптимистический, вероятный и пессимистический.

Оптимистический сценарий предполагает экономический рост, повышение уровня жизни, а потому увеличение рождаемости, средней ожидаемой продолжительности предстоящей жизни (СОППЖ), а также достаточно высокую миграционную подвижность.

В вероятном сценарии предполагается постепенное улучшение социально-экономической ситуации в России, но гораздо более медленными темпами, чем в оптимистическом сценарии.

В пессимистическом сценарии стагнация сложившейся экономической ситуации влечет сохранение негативных тенденций в области смертности и миграции и делает маловероятным повышение рождаемости.

Динамика коэффициентов суммарной рождаемости приведена в табл. 5. Отметим, что даже для оптимистического сценария невозможен выход на уровень простого воспроизводства. Это связано с накопившимися негативными тенденциями в структуре населения и невозможностью быстрого изменения возрастных коэффициентов рождаемости.

Таблица 5

Среднее число детей на одну женщину

Сценарий		**1999**	**2005**	**2010**	**2015**
Пессимистический	Город	1,072	0,946	0,873	0,800
	Село	1,479	1,223	1,052	0,932
Вероятный	Город	1,072	1,062	1,131	1,200
	Село	1,479	1,429	1,382	1,360
Оптимистический	Город	1,072	1,177	1,388	1,600
	Село	1,479	1,634	1,711	1,788

В табл. 6 приведены прогнозные значения продолжительности жизни. В вероятном сценарии предполагается выйти на уровень 1960–1990-х гг., в оптимистическом превзойти его.

Таблица 6

Средняя ожидаемая продолжительность предстоящей жизни, лет

Сценарий		**1999**	**2005**	**2010**	**2015**
Пессимистический	Мужчины	59,9	59,7	59,7	59,7
	Женщины	72,4	72,3	72,3	72,3
Вероятный	Мужчины	59,9	60,8	61,5	62,3
	Женщины	72,4	73,0	73,4	73,8
Оптимистический	Мужчины	59,9	62,0	63,4	64,9
	Женщины	72,4	73,8	74,6	75,4

Оказалось, что средняя ожидаемая продолжительность предстоящей жизни в России для всего населения в целом составила в 2019 г. 73,3 года, в 2020 г. — 71,5 лет.

Труднее всего спрогнозировать будущую миграцию, поскольку она определяется развитием внутриполитической и социально-экономической ситуации во многих странах, связанных с Россией миграционными потоками. Мнение Госкомстата РФ отражено в табл. 7.

Таблица 7

Прогноз чистой миграции (тыс. человек в год) на 2000–2015 гг.

Сценарий	1999	2005	2010	2015
Пессимистический	154,6	– 0,5	– 15,0	– 22,1
Вероятный	154,6	167,1	148,4	133,3
Оптимистический	154,6	323,2	300,1	278,3

Прогноз фактической численности населения России приведен в табл. 8. К 2015 г. население уменьшится. По пессимистическому сценарию — на 20,2 миллиона человек по сравнению с 1999 г.; по вероятному — на 11,5 миллиона; по оптимистическому — на 3,1 миллиона. И только для оптимистического варианта намечается небольшой подъем с 2010 г. по 2015 г. (на 0,1 миллиона человек).

Таблица 8

Прогноз численности населения России (млн человек)

Сценарий	1999	2005	2010	2015
Пессимистический	145,9	139,3	133,0	125,7
Вероятный	145,9	141,2	137,9	134,4
Оптимистический	145,9	143,1	142,7	142,8

Из приведенных результатов прогнозирования вытекают вполне определенные выводы, в частности, в области стратегического менеджмента. Сокращение численности населения означает, грубо говоря, сокращение внутреннего рынка. При более подробных расчетах можно получить прогнозируемую динамику для различных возрастных групп и соответствующие экономические выводы. Например, падение общего числа рождений в 1990-х годах (грубо говоря, в 2 раза) влечет в дальнейшем сокращение рынка образовательных услуг и рынка труда, а также числа лиц, годных к военной службе. После 2005 г. численность населения в рабочих возрастах начнет быстро снижаться, а доля пенсионеров в численности населения — возрастать, что, очевидно, вызовет трудности в пенсионном обеспечении.

Различные виды прогнозов. При рассмотрении моделей взаимосвязи факторов типа ЖОК мы увидим (глава 9), что при моделировании развития какого-либо социально-экономического процесса целесообразно рассматривать три типа сценариев: «Пассивный», «Активный» и «Цель». Первый из них соответствует случаю, когда не меняются эндогенные параметры (т.е. внешние по отношению к моделируемой системе параметры, обычно параметры управления). Сценарии второго типа соответствуют заданным законам изменения параметров управления. Для сценариев типа «Цель» для заданных целевых ограничений или значений параметров находятся способы управления (правила изменения параметров управления), обеспечивающие оптимальное (в том или ном смысле) достижение цели.

В Центре демографии и экологии человека Института народнохозяйственного прогнозирования РАН разработана система сценариев демографических прогнозов на период до 2050 г. [11]. Четыре сценария рассчитаны в предположении нулевой чистой миграции (эндогенными переменными, т.е. переменными управления, являются показатели рождаемости и смертности, а именно, число рождений на одну женщину, средняя продолжительность жизни мужчин и женщин):

Сценарий 1: низкая рождаемость (1,3 рождения на одну женщину; в 1999 г. в России — 1,17 рождений на одну женщину), высокая смертность (СОППЖ для мужчин — 59,9 лет, для женщин — 72,5 лет, как в 1999 г.).

Сценарий 2: низкая рождаемость (как в сценарии 1), снижающаяся смертность (СОППЖ растет и к 2050 г. достигает 77,0 лет для мужчин и 83,0 лет для женщин).

Сценарий 3: растущая рождаемость (к 2050 г. поднимается до 2,0 рождений на одну женщину), высокая смертность (как в 1999 г.).

Сценарий 4: растущая рождаемость (как в сценарии 3) и снижающая смертность (как в сценарии 2).

Сценарий 1 относится к типу «Пассивный», сценарии 2–4 — к типу «Активный». Прогноз численности населения приведен в табл. 9. Только сценарий 4 соответствует выходу России из демографического тупика — достижению к 2050 г. воспроизводимости населения и повышению средней продолжительности жизни до уровня передовых в этом отношении стран. Однако и в этом случае из-за накопившихся к настоящему моменту проблем населению России предстоит сократиться в ближайшие 50 лет на 33,5 млн человек, т.е. на 23 %.

Таблица 9

Прогноз численности населения России (млн человек)

Год	Сценарий 1	Сценарий 2	Сценарий 3	Сценарий 4
2000	145,2	145,2	145,2	145,2
2025	121,4	128,0	122,2	128,8
2050	86,5	103,3	94,5	111,7

Введем в рассмотрение миграцию как параметр управления, позволяющий достигать те или иные цели. Первая цель — сохранить численность населения России на уровне 2000 г. Значения чистой миграции, необходимые для достижения этой цели, приведены в табл. 10 (в предположениях сценариев 1–4 о характеристиках рождаемости и смертности).

Таблица 10

Среднегодовой миграционный прирост (тыс. человек), обеспечивающий сохранение численности населения России

Год/период	Сценарий 1	Сценарий 2	Сценарий 3	Сценарий 4
2000	214	214	214	214
2000–2025	1 040	738	1 002	702
2025–2050	1 712	1 066	1 263	677

Приведенные в табл. 10 значения среднегодового миграционного прироста в несколько раз превышают его значение на 2000 г. Динамика среднегодового миграционного прироста в 1990-х гг. (табл. 4) показывает, что после всплеска в первые годы после развала СССР наблюдается тенденция к стабилизации этого показателя. Следовательно, нет веских оснований надеяться на то, что за счет сложившегося механизма миграции удастся стабилизировать численность населения России.

Тем более не удастся обеспечить его увеличение, например, на 0,5 % в год (в 1970–1980-х годах численность населения России росла на 0,6–0,7 % в год). Соответствующие ежегодному росту на 0,5 % в 2000–2050 гг. значения миграционного прироста приведены в табл. 11. Эти значения по сравнению с данными табл. 10 выросли примерно в 2 раза.

Анализ рассмотренных сценариев типа «Цель» показывают, что для обеспечения сохранения или, тем более, роста численности населения России за счет миграции следует изменить сложившийся механизм миграции. Из табл. 11 следует, что за 50 лет чистая миграция должна составить 75–140 миллионов человек. Русская диаспора составляет порядка 30 миллионов человек и может дать не более половины нужного числа мигрантов.

Таблица 11

Среднегодовой миграционный прирост (тыс. человек), дающий прирост численности населения России на 0,5 % в год

Год/период	Сценарий 1	Сценарий 2	Сценарий 3	Сценарий 4
2000	214	214	214	214
2000–2025	1 878	1 550	1 836	1 510
2025–2050	2 825	2 028	2 268	1 542

Методы управления демографическими процессами. Государственная власть и общественные движения могут влиять на демографические процессы, т.е. осуществлять управления ими. Среди методов управления можно выделить административно-правовые, экономические, социально-психологические.

К административно-правовым относятся, например, правила, регулирующие возможность искусственного прерывания беременности. Исторический опыт показывает, что запрещение абортов, хотя и влечет некоторые отрицательные последствия, в целом приводит к повышению рождаемости. В некоторых странах прибегают к административно-правовым методам сокращения рождаемости. Речь идет, например, о популярной в США идее о принудительной стерилизации безработных, уже имеющих детей.

К снижению смертности приводит борьба с алкоголизмом и наркоманией. Ранее обсуждалось увеличение средней продолжительности жизни (СОППЖ) в результате антиалкогольной компании 1980-х годов. Законы, регулирующие миграцию и права переселенцев, очевидным образом влияют на численность населения. А иногда, при поощрении миграции лиц определенных возрастных групп, эти законы влияют также на показатели смертности и рождаемости.

К экономическим методам управления демографическими процессами относится, прежде всего, система пособий на детей, позволяющая в ряде случаев нейтрализовать материальные потери, которые несут семьи при появлении детей (потеря в заработке матери и расходы непосредственно на детей). Весьма эффективны вложения в систему здравоохранения, позволяющие сократить младенческую смертность и продлить жизнь пожилым людям. Необходимо упомянуть расходы на охрану труда и повышение промышленной безопасности, снижающие смертность в рабочих возрастах. Весьма важным является экономическое стимулирование различных форм оздоровления. Ясно, что экономическая поддержка мигрантов (первоначальные «подъемные», приобретение пенсионных прав и др.) способствует росту населения страны.

Многообразны социально-психологические методы управления демографическими процессами. Господствующие в обществе взгляды весьма сильно влияют, например, на рождаемость. В традиционном обществе она существенно выше (это наглядно видно при сравнении сельского и городского населения России). Большое влияние (как на рождаемость, так и на смертность) имеет сложившийся в определенный момент времени оптимистический или пессимистический настрой в обществе. Различные движения «за здоровый образ жизни» могут снизить смертность, а терпимость к вредным привычкам — повысить ее. На миграционные процессы сильно влияет отношение масс к иммиграции и эмиграции. Совершенно очевидно, что государство и общественные структуры могут влиять на психологию масс, в частности, через средства массовой информации.

Ясно, что управление демографическими процессами должно быть комплексным. Методы управления многообразны, невозможно выделить только один или небольшое число методов в качестве единственно перспективных для применения. Необходимо разрабатывать и реализовывать развернутые программы действий, включающие весьма много конкретных мероприятий. Это могут быть федеральные или региональные программы, программы партий или общественных объединений.

Управление персоналом на конкретном предприятии также требует разработки моделей движения персонала. Первый этап — прогнозирование этого движения. Кроме рождаемости в женской части коллектива, ухода на пенсию и смертности необходимо учитывать поступление на работу и увольнения, а также движение кадров внутри предприятия.

На основе прогнозирования, в том числе с помощью сценариев типа «Активный» и «Цель», проводится процесс принятия решений, в частности, планирование, а в процессе осуществления плана — контроль. На современном этапе весьма актуален контроллинг в области демографического менеджмента.

8.3. Статистические модели движения товарных потоков

Термин «логистика» происходит от французского слова «*loger*» (размещение, расквартирование), которое употребляется в военной терминологии для определения движения военных грузов, их складирования и размещения, а также для описания процесса размещения и расквартирования военных подразделений. В настоящее время термин «логистика» широко используется в деловом мире и определяет теорию и практику движения сырья, материалов, комплектующих изделий, производственных, трудовых и финансовых ресурсов, готовой продукции от их источников к потребителям.

Логистика — наука о планировании, управлении и контроле за движением материальных, информационных и финансовых ресурсов в различных производственно-экономических системах. Предметом логистики является комплексное управление всеми материальными и нематериальными потоками в таких системах. Новизна концепции логистики в области управления промышленными системами состоит во всестороннем подходе к вопросам движения материальных благ в процессе производства и управления. Логистическая система должна охватывать и согласовывать процессы производства, закупок и распределения продукции, а также быть основой при стратегическом планировании и прогнозировании. Итак, логистика — это экономическая дисциплина, занимающаяся оптимальной организацией материальных, финансовых и информационных потоков.

Одна из основных частей логистики — теория управления запасами. Сколько товара держать на складе? Много — будут омертвляться оборотные средства, вложенные в запас. Мало — слишком часто надо будет заниматься получением новых партий товара и нести соответствующие расходы. Значит, надо рассчитать и использовать оптимальный размер запаса. А для этого необходимо построить соответствующую математическую модель.

Управление запасами (другими словами, материально-техническое снабжение) — неотъемлемая часть работы фирм и организаций. Речь идет о запасах сырья, топлива, материалов, инструментов, комплектующих изделий, полуфабрикатов, готовой продукции на промышленном (или сельскохозяйственном) предприятии, о запасах товаров на оптовых базах, складах магазинов, на рабочих местах продавцов, наконец, у потребителей. Запасы постоянно расходуются и пополняются по тем или иным правилам, принятым на предприятии. Оптимизация этих правил, т.е. оптимальное управление запасами, дает большой экономический эффект.

Математическая теория управления запасами является крупной областью экономико-математических исследований. Предложенная еще в 1915 г. Ф. Харрисом классическая модель теории управления запасами является одним из наиболее простых и наглядных примеров применения математического аппарата для принятия решений в экономической области. Эту модель обычно называют моделью Вильсона, так как она получила известность после публикации работы Р.Г. Вильсона в 1934 г.

Формула оптимального размера заказа, полученная в модели Вильсона, широко применяется на различных этапах производства и распределения продукции, поскольку оказывается практически полезной для принятия решений при управлении запасами, в частности, приносящей заметный экономический эффект [5]. Рассмотрим эту модель подробнее.

Классическая модель управления запасами. Пусть $y(t)$ — величина запаса некоторого товара на складе в момент времени t, $t \geq 0$. Дефицит не допускается, т.е. $y(t) \geq 0$ при всех t. Товар пользуется равномерным спросом с интенсивностью μ, т.е. за интервал времени Δt со склада извлекается и поступает потребителям часть запаса величиной $\mu\Delta t$. В моменты времени $t_0 = 0$, t_1, t_2, … пополняется запас на складе — приходят поставки величиной Q_0, Q_1, Q_2, … соответственно. Таким образом, изменение во времени величины запаса $y(t)$ товара на складе изображается ломаной зубчатой линией (рис. 1), состоящей из наклонных и вертикальных звеньев, причем наклонные отрезки параллельны.

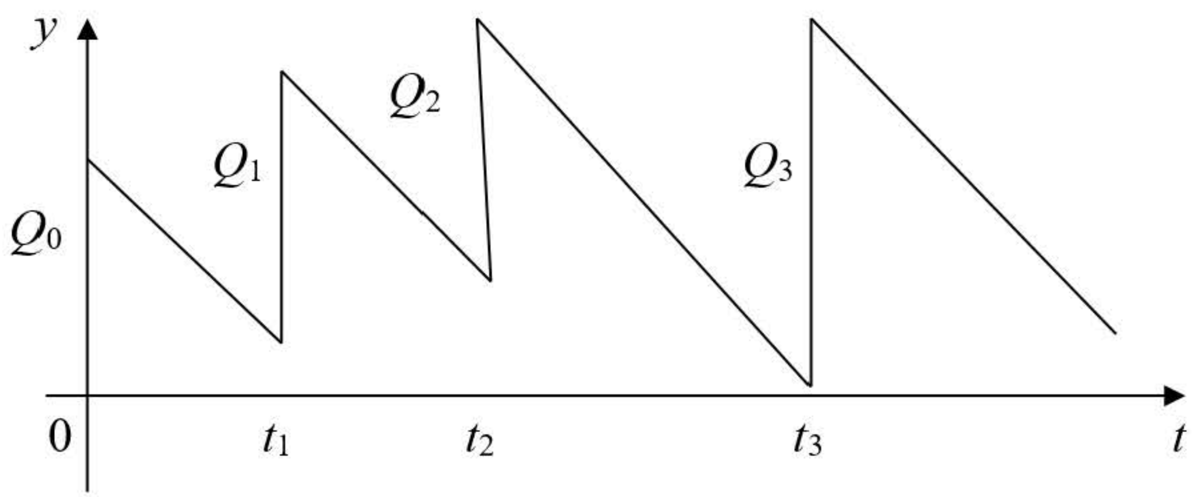


Рис. 1. График изменения величины запаса на складе

Таким образом, в момент t_i величина запаса на складе $y(t)$ скачком увеличивается на Q_i. Следовательно, функция $y(t)$ имеет разрывы в точках t_1, t_2, … Для определенности будем считать, что эта функция непрерывна справа.

Пусть s — плата за хранение единицы товара в течение единицы времени. Поскольку можно считать, что величина запаса $y(t)$ не меняется в течение интервала времени $(t;\ t + dt)$, где dt — дифференциал, т.е. бесконечно малая, то плата за хранение всего запаса в течение этого интервала времени равна $sy(t)dt$. Следовательно, затраты за хранение в течение интервала времени $[0;T)$, где T — интервал планирования, пропорциональны (с коэффициентом пропорциональности s) площади под графиком уровня запаса на складе $y(t)$ и равны:

$$s\int_0^T y(t)dt.$$

Пусть g — плата за доставку одной партии товара. Примем для простоты, что она не зависит от размера поставки. Позже покажем, что если эта плата равна $g + g_1Q$, где Q — размер поставки, то оптимальный план поставки — тот же, что и при отсутствии линейного члена. Будет проанализирована и более сложная модель, в которой предусмотрена скидка с ростом поставки, приводящая к выражению $g + g_1Q + g_2Q^2$ для платы за доставку одной партии товара размером Q.

Пусть $n(T)$ — количество поставок, пришедших в интервале $[0;T)$. При этом включаем поставку в момент $t = 0$ и не включаем поставку в момент $t = T$ (если такая поставка происходит). Тогда суммарные издержки на доставку товара равны $gn(T)$. Следовательно, общие издержки (затраты, расходы) за время T равны:

$$F(T;y) = F(y(t),\ 0 \le t < T) = gn(T) + s\int_0^T y(t)dt.$$

Запись $F(T;y) = F(y(t),\ 0 \le t < T)$ означает, что общие издержки зависят от значений функции $y = y(t)$ при всех $0 \le t < T$. Символ y обозначает функцию как целое. Другими словами, область определения $F(T;y)$ при фиксированном горизонте планирования T — не множество чисел, а множество функций.

Общие издержки, очевидно, возрастают при росте горизонта планирования T. Поэтому часто используют средние издержки, приходящиеся на единицу времени. Средние издержки за время T равны:

$$f(T;y) = f(y(t),\ 0 \le t < T) = \frac{1}{T} F(T;y) = \frac{1}{T}\left\{ gn(T) + s\int_0^T y(t)dt \right\}.$$

Поскольку товар отпускается со склада с постоянной интенсивностью (скоростью), дефицит не допускается, то доходы от работы склада пропорциональны горизонту планирования, средние доходы постоянны. Следовательно, максимизация прибыли эквивалентна минимизации издержек или средних издержек.

Если задать моменты прихода поставок и величины партий, то будет полностью определена функция $y=y(t)$ при всех $0\le t<T$. Верно и обратное утверждение — фиксация функции $y=y(t)$, $0\le t<T$, рассматриваемого вида (рис.1) полностью определяет моменты прихода поставок и величины партий. И то, и другое будем называть *планом* поставок или планом работы системы управления запасами. Для ее оптимизации необходимо выбрать моменты времени $t_0 = 0$, t_1, t_2, … пополнения запаса на складе и размеры поставляемых партий товара Q_0, Q_1, Q_2, … так, чтобы минимизировать средние издержки $f_T(y)$ при фиксированном T.

Модель производственной ситуации (т.е. работы склада) описывается четырьмя параметрами — μ (интенсивность спроса), s (стоимость хранения единицы продукции в течение единицы времени), g (стоимость доставки партии товара), T (горизонт планирования).

Поставленная задача оптимизации работы склада интересна тем, что неизвестно число $2n(T) - 1$ параметров, определяющих план поставок. Поэтому ее решение не может быть проведено с помощью стандартных методов теории оптимизации.

Решим эту задачу в три этапа. На первом установим, что оптимальный план следует искать среди тех планов, у которых все зубцы доходят до оси абсцисс, т.е. запас равен 0 в момент доставки очередной партии. Цель второго этапа — доказать, что все зубцы должны быть одной и той же высоты. Наконец, на третьем находим оптимальный размер поставки.

Оптимальный план. Найдем наилучший план поставок. План, для которого в моменты доставок очередных партий запас равен 0 (т.е. $y(t) = 0$), назовем *напряженным.*

Утверждение 1. Для любого плана поставок, не являющегося напряженным, можно указать напряженный план, для которого средние издержки меньше.

Покажем, как можно от произвольного плана перейти к напряженному плану, уменьшив при этом издержки. Пусть с течением времени при приближении к моменту t_1 прихода поставки Q_1 уровень запаса не стремится к 0, а лишь уменьшается до положительного значения $y(t_1\text{-})$ (где знак «минус» означает предел слева функции $y(t)$ в точке t_1). Тогда рассмотрим новый план поставок с теми же моментами поставок и их величинами, за исключением величин поставок в моменты $t = 0$ и $t = t_1$. А именно, заменим Q_0 на $Q_{01} = Q_0 - y(t_1-)$, а Q_1 на $Q_{11} = Q_0 + y(t_1-)$. Тогда график уровня запаса на складе параллельно сдвинется вниз на интервале $(0; t_1)$, достигнув 0 в t_1, и не изменится правее точки t_1. Следовательно, издержки по доставке партий не изменятся, а издержки по хранению уменьшатся на величину, пропорциональную (с коэффициентом пропорциональности s) площади параллелограмма, образованного прежним и новым положениями графика уровня запаса на интервале $(0; t_1)$ (см. рис.2).

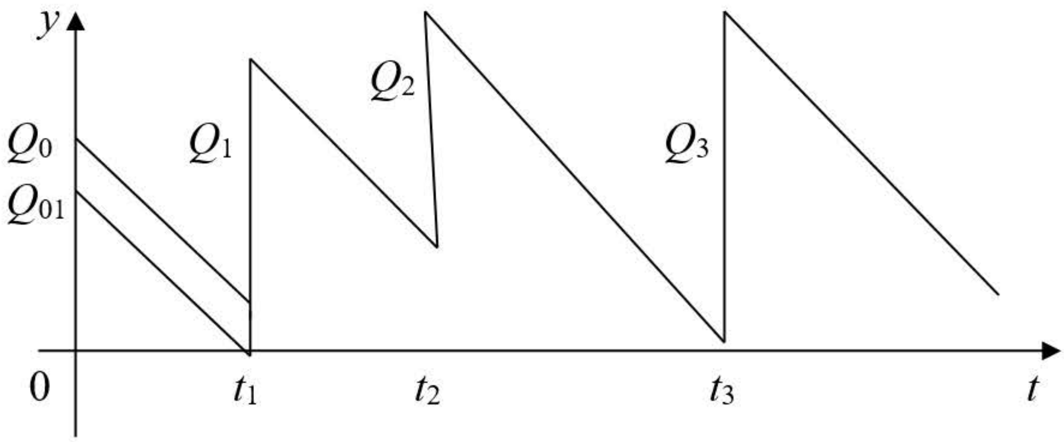


Рис. 2. Первый шаг перехода к напряженному плану

Итак, в результате первого шага перехода получен план, в котором крайний слева зубец достигает оси абсцисс. Следующий шаг проводится аналогично, только момент времени $t = 0$ заменяется на $t = t_1$. Если есть такая возможность, второе наклонное звено графика уровня запаса на складе параллельно сдвигается вниз, достигая в крайней правой точке t_2 оси абсцисс.

Аналогично поступаем со всеми остальными зубцами, двигаясь слева направо. В результате получаем напряженный план. На каждом шагу издержки по хранению либо сокращались, либо оставались прежними (если соответствующее звено графика не опускалось вниз). Следовательно, для полученного в результате описанного преобразования напряженного плана издержки по хранению меньше, чем для исходного плана, либо равны (если исходный план уже являлся напряженным).

Из утверждения 1 следует, что оптимальный план следует искать только среди напряженных планов. Другими словами, план, не являющийся напряженным, не может быть оптимальным.

Утверждение 2. Среди напряженных планов с фиксированным числом поставок минимальные издержки имеет тот, в котором все интервалы между поставками равны.

При фиксированном числе поставок затраты на доставку партий не меняются. Следовательно, достаточно минимизировать затраты на хранение.

Для напряженных планов размеры поставок однозначно определяются с помощью интервалов между поставками:

$$Q_{i-1} = \mu(t_i - t_{i-1}),\ i = 1,2,\ldots,n(T)-1,\ Q_{n(T)-1} = \mu(T - t_{n(T)-1}).$$

Действительно, очередная поставка величиной Q_{i-1} совпадает с размером запаса на складе в момент t_{i-1}, расходуется с интенсивностью μ единиц товара в одну единицу времени и полностью исчерпывается к моменту t_i прихода следующей поставки.

Для напряженного плана издержки по хранению равны:

$$s\int_0^T y(t)dt = s\sum_{i=1}^{n(T)}\frac{Q_{i-1}(t_i - t_{i-1})}{2} = s\sum_{i=1}^{n(T)}\frac{\mu(t_i - t_{i-1})^2}{2} = s\sum_{i=1}^{n(T)}\frac{\mu\Delta_i^2}{2} = \frac{\mu s}{2}\sum_{i=1}^{n(T)}\Delta_i^2,$$

где $\Delta_i = t_i - t_{i-1}$, $i = 1, 2, \ldots, n(T)$, $t_{n(T)} = T$. Ясно, что Δ_i, $i = 1, 2, \ldots, n(T)$, — произвольные неотрицательные числа, в сумме составляющие T. Следовательно, для минимизации издержек среди напряженных планов с фиксированным числом поставок достаточно решить задачу оптимизации:

$$\begin{cases} \Delta_1^2 + \Delta_2^2 + \ldots + \Delta_n^2 \to \min, \\ \Delta_1 + \Delta_2 + \ldots + \Delta_n = T, \\ \Delta_i \geq 0,\ i = 1,2,\ldots,n, \end{cases}$$

где $n = n(T)$.

Полученная задача оптимизации формально никак не связана с логистикой, она является чисто математической. Для ее решения целесообразно ввести новые переменные $\alpha_i = \Delta_i - T/n$, $i = 1, 2, \ldots, n$. Тогда:

$$\sum_{i=1}^{n}\alpha_i = \sum_{i=1}^{n}(\Delta_i - \frac{T}{n}) = \left(\sum_{i=1}^{n}\Delta_i\right) - n\frac{T}{n} = T - T = 0.$$

Поскольку $\Delta_i = T/n + \alpha_i$, то:

$$\Delta_i^2 = \frac{T^2}{n^2} + 2\frac{T}{n}\alpha_i + \alpha_i^2,$$

следовательно, с учетом предыдущего равенства имеем:

$$\sum_{i=1}^{n} \Delta_i^2 = n\frac{T^2}{n^2} + 2\frac{T}{n}\sum_{i=1}^{n}\alpha_i + \sum_{i=1}^{n}\alpha_i^2 = \frac{T^2}{n} + \sum_{i=1}^{n}\alpha_i^2.$$

Сумма квадратов всегда неотрицательна. Она достигает минимума, равного 0, когда все переменные равны 0, т.е. при $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$. Тогда:

$$\Delta_i = \frac{T}{n},\ i = 1,2,...,n.$$

При этих значениях Δ_i выполнены все ограничения оптимизационной задачи. Итак, утверждение 2 доказано.

Для плана с равными интервалами между поставками все партии товара имеют одинаковый объем. Для такого плана издержки по хранению равны:

$$s\int_0^T y(t)dt = \frac{\mu s}{2}\sum_{i=1}^{n(T)}\Delta_i^2 = \frac{\mu s T^2}{2n(T)}.$$

Средние издержки (на единицу времени) таковы:

$$f(T;y) = \frac{1}{T}\left\{gn(T) + \frac{\mu s T^2}{2n(T)}\right\} = g\frac{n(T)}{T} + \mu s\frac{T}{2n(T)}.$$

Итак, минимизация средних издержек — это задача дискретной оптимизации. На третьем этапе построения оптимального плана необходимо найти натуральное число $n(T)$ — самое выгодное число поставок.

Поскольку к моменту T запас товара должен быть израсходован, то общий объем поставок за время T должен совпадать с общим объемом спроса, следовательно, равняться μT. Справедливо балансовое соотношение (аналог закона Ломоносова-Лавуазье сохранения массы при химических реакциях):

$$Qn(T) = \mu T.$$

Из балансового соотношения следует, что:

$$\frac{n(T)}{T} = \frac{\mu}{Q}.$$

Средние издержки (на единицу времени) можно выразить как функцию размера партии Q:

$$f(T;y) = g\frac{n(T)}{T} + \mu s\frac{T}{2n(T)} = f_1(Q) = \frac{\mu g}{Q} + \frac{sQ}{2}. \qquad (1)$$

Задача состоит в минимизации $f_1(Q)$ по Q. При этом возможная величина поставки принимает дискретные значения, поскольку:

$$Q \in \left\{\frac{\mu T}{n}.\ n = 1,2,\ldots\right\}.$$

Изучим функцию $f_1(Q)$, определенную при $Q > 0$. При приближении к 0 она ведет себя как гипербола, при росте аргумента — как линейная функция. Производная имеет вид:

$$\frac{df_1(Q)}{dQ} = -\frac{\mu g}{Q^2} + \frac{s}{2}. \qquad (2)$$

Производная монотонно возрастает, поэтому рассматриваемая функция имеет единственный минимум в точке, в которой производная равна 0, т.е. при:

$$Q_0 = \sqrt{\frac{2\mu g}{s}}. \qquad (3)$$

Получена знаменитая в теории управления запасами «формула квадратного корня».

В литературе иногда без всяких комментариев рекомендуют использовать напряженный план, в котором размеры всех поставляемых партий равны Q_0. К сожалению, получаемый таким путем план почти всегда не является оптимальным, т.е. популярная рекомендация неверна или не вполне корректна. Дело в том, что почти всегда:

$$Q_0 \notin \left\{\frac{\mu T}{n}.\ n = 1,2,\ldots\right\}.$$

Как же найти оптимальный план? Всегда можно указать неотрицательное целое число n такое, что:

$$Q_1 = \frac{\mu T}{n+1} < Q_0 \le \frac{\mu T}{n} = Q_2. \qquad (4)$$

Утверждение 3. Решением задачи оптимизации:

$$f_1(Q) = \frac{\mu g}{Q} + \frac{sQ}{2} \to \min,$$
$$Q \in \left\{\frac{\mu T}{n}.\ n = 1,2,\ldots\right\}$$

является либо Q_1, либо Q_2.

Действительно, из всех возможных объемов партии:

$$Q \in \left\{\frac{\mu T}{n}.\ n = 1,2,\ldots\right\}$$

часть лежит правее Q_0, из них наименьшим является Q_2, а часть лежит левее Q_0, из них наибольшим является Q_1. Для построения оптимального плана обратим внимание на то, что производная (2) отрицательна левее Q_0 и положительна правее Q_0, следовательно, функция средних издержек $f_1(Q)$ убывает левее Q_0 и возрастает правее Q_0. Значит, минимум по:

$$Q \in \left\{\frac{\mu T}{n}.\ n = 1,2,\ldots\right\} \cap \{Q : Q \ge Q_0\}$$

достигается при $Q = Q_2$, а минимум по:

$$Q \in \left\{\frac{\mu T}{n}.\ n = 1,2,\ldots\right\} \cap \{Q : Q < Q_0\}$$

— при $Q = Q_1$. Последнее утверждение эквивалентно заключению утверждения 3.

Итак, алгоритм построения оптимального плана таков.

1. Найти Q_0 по формуле квадратного корня (3).

2. Найти n из условия (4).

3. Рассчитать $f_1(Q)$ по формуле (1) для $Q = Q_1$ и $Q = Q_2$, где Q_1 и Q_2 определены соотношением (4).

4. Наименьшее из двух чисел $f_1(Q_1)$ и $f_1(Q_2)$ является искомым минимумом, а то из Q_1 и Q_2, на котором достигается минимум — решением задачи оптимизации. Обозначим его Q_{opt}.

Итак, оптимальный план поставки — это напряженный план, в котором объемы всех поставок равны Q_{opt}.

Замечание. Если $f_1(Q_1) = f_1(Q_2)$, то решение задачи оптимизации состоит из двух точек Q_1 и Q_2. В этом частном случае существует два оптимальных плана.

Пример 1. На складе хранится некоторая продукция, пользующаяся равномерным спросом. За 1 день со склада извлекается 5 т продукции. Плата за хранение 1 т. продукции в день — 50 руб. Плата на доставку одной партии — 980 руб. Горизонт планирования — 10 дней. Найти оптимальный план поставок.

В рассматриваемом случае μ = 5 (т/день), s = 50 (руб./т·день), g = 980 (руб./партия), T = 10 (дней). По формуле (3) рассчитываем:

$$Q_0 = \sqrt{\frac{2\mu g}{s}} = \sqrt{\frac{2\times 5\times 980}{50}} = \sqrt{196} = 14.$$

Множество допустимых значений для Q имеет вид:

$$\left\{\frac{\mu T}{n},\ n = 1,2,\ldots\right\} = \left\{50;\ \frac{50}{2};\ \frac{50}{3};\ \frac{50}{4};\ldots\right\} = \{50;\ 25;\ 16,67;\ 12,5;\ldots\}$$

Следовательно, Q_1 = 12,5 и Q_2 = 16,67. Первое значение определяет напряженный план с четырьмя одинаковыми зубцами, а второе — с тремя. Поскольку:

$$f_1(Q) = \frac{5\times 980}{Q} + \frac{50Q}{2} = \frac{4900}{Q} + 25Q,$$

то

$$f_1(Q_1) = f_1(12,5) = \frac{4900}{12,5} + 25\times 12,5 = 392 + 312,5 = 704,5$$

и

$$f_1(Q_2) = f_1(50/3) = \frac{4900\times 3}{50} + 25\times\frac{50}{3} = 294 + 416,67 = 710,67.$$

Поскольку $f_1(Q_1) < f_1(Q_2)$, то $Q_{opt} = Q_1$ = 12,5. Итак, оптимальным является напряженный план с четырьмя зубцами.

Как уже отмечалось, часто рекомендуют применять план поставок с $Q=Q_0$. Каков при этом проигрыш по сравнению с оптимальным планом?

Для плана с $Q=Q_0$ интервал между поставками составляет $Q_0/\mu = 14/5 = 2{,}8$ дня. Следовательно, партии придут в моменты $t_0 = 0$; $t_1 = 2{,}8$; $t_2 = 5{,}6$; $t_3 = 8{,}4$. Следующая партия должна была бы прийти уже за пределами горизонта планирования $T=10$, в момент $t_4 = 11{,}2$. Таким образом, график уровня запаса на складе в пределах горизонта планирования состоит из трех полных зубцов и одного не полного. К моменту $T=10$ пройдет $10 - 8{,}4 = 1{,}6$ дня с момента последней поставки, значит, со склада будет извлечено $5 \cdot 1{,}6 = 8$ т продукции и останется $14 - 8 = 6$ т. План с $Q=Q_0$ не является напряженным, а потому не является оптимальным для горизонта планирования $T=10$.

Подсчитаем общие издержки в плане с $Q=Q_0$. Площадь под графиком уровня запаса на складе равна сумме площадей трех треугольников и трапеции. Площадь треугольника равна $(14\times2{,}8)/2 = 19{,}6$, трех треугольников — 58,8. Основания трапеции параллельны оси ординат и равны значениям уровня запаса в моменты времени $t_3 = 8{,}4$ и $T=10$, т.е. величинам 14 и 6 соответственно. Высота трапеции лежит на оси абсцисс и равна $10 - 8{,}4 = 1{,}6$, а потому площадь трапеции есть $[(14 + 6) \cdot 1{,}6]/2 = 16$. Следовательно, площадь под графиком равна $58{,}8 + 16 = 74{,}8$, а плата за хранение составляет $50 \cdot 74{,}8 = 3\,740$ руб.

За 10 дней доставлены 4 партии товара (в моменты $t_0 = 0$; $t_1 = 2{,}8$; $t_2 = 5{,}6$; $t_3 = 8{,}4$). Следовательно, затраты на доставку равны $4 \cdot 980 = 3\,920$ руб. Общие издержки за 10 дней составляют $3\,740 + 3\,920 = 7\,660$ руб., а средние издержки — 766 руб. Они больше средних издержек в оптимальном плане в $766/704{,}5 = 1{,}087$ раза, т.е. на 8,7 %.

Отметим, что:

$$f_1(Q_0) = \frac{4900}{Q_0} + 25Q_0 = \frac{4900}{14} + 25\times14 = 350 + 350 = 700,$$

т.е. меньше, чем в оптимальном плане. Таким образом, из-за дискретности множества допустимых значений средние издержки возросли на 4,5 руб., т.е. на 0,64 %.

При этом оптимальный размер партии (12,5 т) отличается от $Q_0 = 14$ т на 1,5 т, т.е. $Q_{opt}/Q_0 = 0{,}89$ — различие на 11 %. Достаточно большое различие объемов поставок привело к пренебрежимо малому изменению функции $f_1(Q)$. Это объясняется тем, что в точке Q_0 функция $f_1(Q)$ достигает минимума, а потому ее производная в этой точке равна 0.

Оба слагаемых в $f_1(Q_0)$ равны между собой. Случайно ли это? Покажем, что нет. Действительно,

$$\frac{\mu g}{Q_0}=\frac{\mu g}{\sqrt{\dfrac{2\mu g}{s}}}=\sqrt{\frac{\mu g s}{2}};\ \ \frac{sQ_0}{2}=\frac{s\sqrt{\dfrac{2\mu g}{s}}}{2}=\sqrt{\frac{\mu g s}{2}}.$$

Таким образом, составляющие средних издержек, порожденные различными причинами, уравниваются между собой.

Средние издержки в плане с Q=Q_0 равны $\sqrt{2\mu g s}$. Интервал между поставками при этом равен:

$$\frac{Q_0}{\mu}=\frac{\sqrt{\dfrac{2\mu g}{s}}}{\mu}=\sqrt{\frac{2g}{\mu s}}.$$

Издержки в течение одного интервала между поставками таковы:

$$\sqrt{2\mu g s}\times\sqrt{\frac{2g}{\mu s}}=2g,$$

при этом половина (т.е. g) приходится на оплату доставки партии, а половина — на хранение товара.

Асимптотически оптимальный план. Из проведенных рассуждений ясно, что напряженный план с Q=Q_0 является оптимальным тогда и только тогда, когда горизонт планирования T приходится на начало очередного зубца, т.е.:

$$T=n\frac{Q_0}{\mu}=n\sqrt{\frac{2g}{\mu s}},\ \ n=1,2,\ldots \tag{5}$$

Для всех остальных возможных горизонтов планирования T этот план не является оптимальным. Оптимальным будет напряженный план с другим размером поставки. Для дальнейшего весьма существенно, что при изменении горизонта планирования T оптимальный план меняется на всем интервале [0; T].

Как происходит это изменение? При малых T делается лишь одна поставка (при T = 0), график уровня запаса на складе состоит из одного зубца. При увеличении T размер зубца плавно увеличивается. В некоторый момент $T(1)$ происходит переход от одного зубца к двум. В этот момент оптимальны сразу два плана поставки — с одним зубцом и с двумя. При переходе к планам с двумя зубцами размер зубца скачком уменьшается. При дальнейшем увеличении горизонта планирования оптимальный план описывается графиком с двумя одинаковыми зубцами, размер которых плавно растет. Далее в момент $T(2)$ становится оптимальным план с тремя зубцами, размер которых в этот момент скачком уменьшается (в компенсацию за увеличение числа скачков).

Проблема состоит в том, что в реальной экономической ситуации выбор горизонта планирования T весьма субъективен. Возникает вопрос, какой план разумно использовать, если горизонт планирования не известен заранее. Проблема горизонта планирования возникает не только в логистике. Она является общей для любого перспективного планирования, поэтому весьма важна для стратегического менеджмента [5, 15]. Для решения проблемы горизонта планирования необходимо использование конкретной модели принятия решения, в рассматриваемом случае — классической модели управления запасами.

Ответ можно указать, если горизонт планирования является достаточно большим. Оказывается, можно использовать план, в котором все размеры поставок равны Q_0. Для него уровень запаса на складе описывается функцией $y_0(t)$, $0 \le t < +\infty$, состоящей из зубцов высоты Q_0. *Предлагается пользоваться планом, являющимся сужением этого плана на интервал* $[0; T)$. Другими словами, предлагается на интервале $[0; T)$ использовать начальный отрезок этого плана. Он состоит из некоторого количества треугольных зубцов, а последний участок графика, описываемый трапецией, соответствует тому, что последняя поставка для почти всех горизонтов планирования не будет израсходована до конца. Такой план иногда называют планом Вильсона [5].

Ясно, что этот план не будет оптимальным (для всех T, кроме заданных формулой (5)). Действительно, план Вильсона можно улучшить, уменьшив объем последней поставки. Однако у него есть то полезное качество, что при изменении горизонта планирования его начальный отрезок не меняется. Действительно, планы поставок для горизонтов планирования T_1 и T_2, определенные с помощью функции $y_0(t)$, $0 \le t < +\infty$, задающей уровень запасов на складе, совпадают на интервале $[0; \min\{T_1, T_2\})$.

Определение. Асимптотически оптимальным планом называется план поставок — функция $y:[0;+\infty) \to [0;+\infty)$ такая, что:

$$\lim_{T\to\infty} \frac{f(T; y_{opt}(T))}{f(T; y)} = 1,$$

где $y_{opt}(T)$ — оптимальный план на интервале $[0; T)$.

В соответствии с определениями и обозначениями, введенными в начале раздела, $f(T; y_{opt}(T))$ — средние издержки за время T для плана $y_{opt}(T)$, определенного на интервале $[0; T)$, а $f(T;y)$ — средние издержки за время T для плана $y:[0;+\infty)\to[0;+\infty)$.

Теорема 1. План $y = y_0$ является асимптотически оптимальным.

Таким образом, для достаточно больших горизонтов планирования T планы $y_0(t)$, $0\leq t<T$, все зубцы у которых имеют высоту Q_0, имеют издержки, приближающиеся к минимальным. Следовательно, эти планы Вильсона, являющиеся сужениями одной и той же функции $y:[0;+\infty)\to[0;+\infty)$ на интервалы $[0; T)$ при различных T, можно использовать одновременно при всех достаточно больших T.

Замечание. Согласно [5] решение проблемы горизонта планирования состоит в использовании асимптотически оптимальных планов, которые близки (по издержкам) к оптимальным планам сразу при всех достаточно больших T.

Доказательство. По определению оптимального плана:

$$\frac{f(T; y_{opt}(T))}{f(T;y)}\leq 1. \tag{6}$$

Найдем нижнюю границу для рассматриваемого отношения. При фиксированном горизонте планирования T можно указать неотрицательное целое число n такое, что:

$$\frac{nQ_0}{\mu}\leq T<\frac{(n+1)Q_0}{\mu}.$$

Так как $Tf(T; y_{opt}(T))$ и $\frac{nQ_0}{\mu}f\left(\frac{nQ_0}{\mu}; y_{opt}(T)\right)$ — общие издержки на интервалах $[0; T)$ и $[0; nQ_0/\mu)$ соответственно при использовании оптимального на [0; Т) плана, то, очевидно, поскольку второй интервала — часть первого (или совпадает с ним), первые издержки больше вторых, т.е.:

$$Tf(T; y_{opt}(T))\geq \frac{nQ_0}{\mu}f\left(\frac{nQ_0}{\mu}; y_{opt}(T)\right).$$

Далее, т.к. на интервале $(0;\ nQ_0/\mu)$, включающем целое число периодов плана y_0, оптимальным является начальный отрезок этого плана $y_0(nQ_0/\mu)$, то:

$$\frac{nQ_0}{\mu} f\left(\frac{nQ_0}{\mu}; y_{opt}(T)\right) \geq \frac{nQ_0}{\mu} f\left(\frac{nQ_0}{\mu}; y_0(T)\right).$$

В правой части последнего неравенства стоит $\frac{nQ_0}{\mu}\sqrt{2\mu gs}$ (здесь использована формула для минимального значения средних издержек $f(T;\ y)$ при T, кратном nQ_0/μ). Из проведенных рассуждений вытекает, что:

$$Tf(T;\ y_{opt}(T)) \geq \frac{nQ_0}{\mu}\sqrt{2\mu gs}. \tag{7}$$

Для общих издержек на интервалах $[0;\ T)$ и $[0;\ (n+1)Q_0/\mu)$ при использовании плана y_0, очевидно, справедливо следующее неравенство:

$$Tf(T;\ y_0(T)) \leq \frac{(n+1)Q_0}{\mu} f\left(\frac{(n+1)Q_0}{\mu}; y_0(T)\right).$$

Следовательно,

$$Tf(T;\ y_0\,(T)) \leq \frac{(n+1)Q_0}{\mu}\sqrt{2\mu gs}. \tag{8}$$

Из неравенств (7) и (8) вытекает, что:

$$\frac{f(T; y_{opt}(T))}{f(T; y_0)} \geq \frac{n}{n+1} = 1 - \frac{1}{n+1} \geq 1 - \frac{Q_0}{\mu T}.$$

Так как $\frac{Q_0}{\mu T} \to 0$ при $T \to \infty$, то, учитывая неравенство (6), из последнего неравенства выводим справедливость заключения теоремы 1. Таким образом, асимптотическая оптимальность плана y_0 доказана.

При небольшом T средние издержки в плане Вильсона могут существенно превышать средние издержки в оптимальном плане. Превышение вызвано скачками функции $f(T;\ y_0(T))$, связанными с переходами через моменты прихода очередных поставок (и увеличением общих издержек скачком на величину платы за доставку партии). Величину превышения средних издержек в плане Вильсона по сравнению с оптимальными планами можно рассчитать.

Пусть горизонт планирования $T = t_k + \varepsilon$, где t_k — момент прихода $(k + 1)$-й поставки в плане Вильсона, $\varepsilon > 0$. Тогда, как можно доказать:

$$\lim_{\varepsilon \to 0} \frac{f(T; y_0(T))}{f(T; y_{opt}(T))} = \lim_{\varepsilon \to 0} \frac{f(t_k + \varepsilon; y_0(t_k + \varepsilon)}{f(t_k + \varepsilon; y_{opt}(t_k + \varepsilon)} = 1 + \frac{1}{2k}.$$

Таким образом, затраты в плане Вильсона являются минимальными (относительно оптимального плана) при $T = t_k$, $k = 1, 2, \ldots$, где t_k — моменты прихода поставок. Напомним, что план Вильсона является оптимальным при указанных значениях T. Однако при T, бесконечно близком к t_k, но превосходящем t_k, затраты увеличиваются по сравнению с затратами в оптимальном плане в $\{1 + 1/(2k)\}$ раз. При дальнейшем возрастании T отношение издержек (средних или общих) в плане Вильсона к аналогичным издержкам в оптимальном плане постепенно уменьшается, приближаясь к 1 при приближении (снизу) к моменту t_{k+1} прихода следующей поставки. А там — новый скачок, но уже на меньшую величину $\{1 + 1/(2k + 2)\}$.

Сразу после прихода первой поставки отношение затрат составляет 1,5 (превышение на 50 %), после прихода второй — 1,25 (превышение на 25 %), третьей — 1,167 (превышение на 16,7 %), четвертой — 1,125 (превышение на 12,5 %), пятой — 1,1 (превышение на 10 %), и т.д. Таким образом, при небольших горизонтах планирования T превышение затрат может быть значительным, план Вильсона отнюдь не оптимальный. Но чем больше горизонт планирования, тем отклонение меньше. Уже после сотой поставки оно не превышает 0,5 %.

Влияние отклонений от оптимального объема партии. В реальных производственных и управленческих ситуациях часто приходится принимать решения об использовании объемов партии, отличных от оптимальной величины Q_0, рассчитанной по формуле квадратного корня (3). Например, при ограниченной емкости склада или для обеспечения полной загрузки транспортных средств большой вместимости. Это возможно также в ситуации, когда величина партии измеряется в целых числах (штучный товар) или даже в десятках, дюжинах, упаковках, ящиках, контейнерах и т.д., а величина Q_0 не удовлетворяет этому требованию и, следовательно, не может быть непосредственно использована в качестве объема поставки.

Поэтому необходимо уметь вычислять возрастание средних издержек при использовании напряженного плана с одинаковыми поставками объема Q, отличного от Q_0, по сравнению со средними издержками в оптимальном плане. Будем сравнивать средние издержки за целое число периодов. Как показано выше, они имеют вид:

$$f_1(Q) = \frac{\mu g}{Q} + \frac{sQ}{2},$$

где Q — объем партии. Тогда:

$$\frac{f_1(Q) - f_1(Q_0)}{f_1(Q_0)} = \frac{1}{2}\left(\frac{Q-Q_0}{Q}\right)\left(\frac{Q-Q_0}{Q_0}\right). \tag{9}$$

Это тождество нетрудно проверить с помощью простых алгебраических преобразований.

Пример 2. Пусть используется план с $Q = 0{,}9\ Q_0$. Тогда:

$$\frac{f_1(Q) - f_1(Q_0)}{f_1(Q_0)} = \frac{1}{2}\left(\frac{-0{,}1Q_0}{0{,}9Q_0}\right)\left(\frac{-0{,}1Q_0}{Q_0}\right) = \frac{0{,}01}{1{,}8} = 0{,}0056.$$

Таким образом, изменение объема партии на 10 % привело к увеличению средних издержек лишь на 0,56 %.

Пример 3. Пусть используемое значение объема поставки Q отличается от оптимального не более чем на 30 %. На сколько могут возрасти издержки?

Из формулы (9) вытекает, что максимальное возрастание издержек будет в случае $Q = 0{,}7\ Q_0$. Тогда:

$$\frac{f_1(Q) - f_1(Q_0)}{f_1(Q_0)} = \frac{1}{2}\left(\frac{-0{,}3Q_0}{0{,}7Q_0}\right)\left(\frac{-0{,}3Q_0}{Q_0}\right) = \frac{0{,}09}{1{,}4} = 0{,}0643.$$

Таким образом, издержки могут возрасти самое большее на 6,43 %.

На первый взгляд представляется удивительным, что сравнительно большое отклонение значения переменной Q от оптимального (на 10 %) приводит к пренебрежимо малому возрастанию значения оптимизируемой функции. Этот факт имеет большое прикладное значение. Из него следует, что область «почти оптимальных» значений параметра весьма обширна, следовательно, из нее можно выбирать для практического использования те или иные значения, исходя из иных принципов. Можно, например, минимизировать какую-либо иную целевую функцию, тем самым решая задачу многокритериальной оптимизации. Можно «вписаться» в действующую дискретную систему возможных значений параметров.

Важное замечание 1. Обширность области «почти оптимальных» значений параметра — общее свойство оптимальных решений, получаемых путем минимизации гладких функций. Действительно, пусть необходимо минимизировать некоторую функцию $g(x)$, трижды дифференцируемую. Пусть минимум достигается в точке x_0. Справедливо разложение Тейлора —Маклорена:

$$g(x) = g(x_0) + \frac{dg(x_0)}{dx}(x - x_0) + \frac{1}{2}\frac{d^2 g(x_0)}{dx^2}(x - x_0)^2 + O((x - x_0)^3).$$

Однако в x_0 выполнено необходимое условие экстремума (в данном случае — минимума):

$$\frac{dg(x_0)}{dx} = 0.$$

Следовательно, с точностью до бесконечно малых более высокого порядка (по сравнению с $(x\text{-}x_0)^2$) справедливо равенство:

$$g(x) - g(x_0) = \frac{1}{2}\frac{d^2 g(x_0)}{dx^2}(x - x_0)^2. \tag{10}$$

Это соотношение показывает, что приращение значений минимизируемой функции — бесконечно малая более высокого порядка по сравнению с приращением независимой переменной. Если:

$$x = x_0 + \varepsilon,$$

то

$$g(x) - g(x_0) = C\varepsilon^2,$$

где

$$C = \frac{1}{2}\frac{d^2 g(x_0)}{dx^2}^2 .$$

Вернемся к классической модели управления запасами. Для нее надо рассматривать $f_1(Q)$ в роли $g(x)$. С помощью соотношения (10) заключаем, что:

$$f_1(Q) - f_1(Q_0) = \frac{1}{2}\frac{d^2 f_1(Q_0)}{dQ^2}(Q - Q_0)^2$$

с точностью до бесконечно малых более высокого порядка. Вычислим вторую производную $f_1(Q)$. Поскольку:

$$\frac{df_1(Q)}{dQ} = \frac{d}{dQ}\left(\frac{\mu g}{Q} + \frac{sQ}{2}\right) = -\frac{\mu g}{Q^2} + \frac{s}{2},$$

то

$$\frac{d^2 f_1(Q)}{dQ^2} = \frac{d}{dQ}\left(-\frac{\mu g}{Q^2} + \frac{s}{2}\right) = \frac{2\mu g}{Q^3}.$$

Теперь заметим, что

$$\frac{2\mu g}{Q_0} = \frac{2\mu g}{\sqrt{\dfrac{2\mu g}{s}}} = \sqrt{2\mu g s} = f_1(Q_0).$$

Следовательно,

$$f_1(Q) - f_1(Q_0) = \frac{1}{2}\frac{f_1(Q_0)}{Q_0^2}(Q - Q_0)^2$$

с точностью до бесконечно малых более высокого порядка. Отличие этой формулы от точной формулы (9) состоит только в том, что Q в знаменателе одной из дробей заменено на Q_0.

Устойчивость выводов в математической модели. Вполне ясно, что рассматриваемая классическая модель управления запасами, как и любые иные экономико-математические модели конкретных экономических явлений и процессов, является лишь приближением к реальности. Приближение может быть более точным или менее точным, но никогда не может полностью уловить все черты реальности. Поэтому с целью повышения адекватности получаемых на основе экономико-математической модели выводов целесообразно изучить устойчивость этих выводов по отношению к допустимым отклонениям исходных данных и предпосылок модели [3, 5]. Выше изучено изменение средних издержек при малых отклонениях величины поставки.

Предположим теперь, что вместо истинных значений параметров μ, g, s нам известны лишь их приближенные значения $\mu^* = \mu + \Delta\mu$, $g^* = g + \Delta g$, $s^* = s + \Delta s$. Мы применяем план Вильсона, но с искаженным объемом партии:

$$Q^* = Q^*(\mu^*, g^*, s^*) = \sqrt{\frac{2\mu^* g^*}{s^*}}.$$

Это приводит к возрастанию средних издержек. Согласно формулам (9)–(10) возрастание пропорционально $(\Delta Q)^2$ (с точностью до бесконечно малых более высокого порядка). Здесь:

$$\Delta Q = Q^*(\mu^*, g^*, s^*) - Q_0(\mu, g, s).$$

Выделим в ΔQ главный линейный член:

$$\Delta Q = \frac{\partial Q}{\partial \mu}\Delta\mu + \frac{\partial Q}{\partial g}\Delta g + \frac{\partial Q}{\partial s}\Delta s = \sqrt{\frac{g}{2\mu s}}\Delta\mu + \sqrt{\frac{\mu}{2gs}}\Delta g - \sqrt{\frac{\mu g}{2s^3}}\Delta s \tag{11}$$

(с точностью до бесконечно малых более высокого порядка).

Величину Δμ можно определить по фактическим данным о спросе, оценив величину отклонения реального спроса от линейного приближения [5], например, с помощью математического аппарата линейного регрессионного анализа [3]. Для определения значений параметров g и s необходимо проведение специальных достаточно трудоемких исследований. К тому же существуют различные методики расчета этих параметров, результаты расчетов по которым не совпадают. Поэтому естественно оценить разумную точность определения g и s по известной точности определения μ. Для этого воспользуемся «принципом уравнивания погрешностей», предложенным в [5].

Важное замечание 2. Принцип уравнивания погрешностей состоит в том, что погрешности различной природы должны вносить примерно одинаковый вклад в общую погрешность математической модели. Так, определение рационального объема выборки в статистике интервальных данных основано на уравнивании влияния метрологической и статистической погрешностей [3]. Согласно подходу [5] выбор числа градаций в социологических анкетах целесообразно проводить на основе уравнивания погрешностей квантования и неопределенности в ответах респондентов. В классической модели управления запасами целесообразно уравнять влияние неточностей в определении параметров на отклонение целевой функции от оптимума.

Выберем Δg и Δs так, чтобы увеличение затрат, вызванное неточностью определения g и s, было таким же, как и вызванное неточностью определения μ. С точностью до бесконечно малых более высокого порядка это означает, что необходимо уравнять между собой три слагаемых в правой части (11). После сокращения общего множителя получаем, что согласно принципу уравнивания погрешностей должно быть справедливо соотношение:

$$\frac{|\Delta\mu|}{\mu} = \frac{|\Delta g|}{g} = \frac{|\Delta s|}{s}. \qquad (12)$$

Таким образом, относительные погрешности определения параметров модели должны совпадать.

В соотношении (12) используются истинные значения параметров, которые неизвестны. Поэтому целесообразно вначале вместо параметров использовать их грубые оценки, из (12) определить их примерную точность, затем провести исследования, уточняющие значения параметров. Эту процедуру повторяют до тех пор, пока не произойдет некоторое уравнивание относительных погрешностей определения параметров модели.

Модель с дефицитом. Классическая модель управления запасами может быть обобщена в различных направлениях. Одно из наиболее естественных обобщений — введение в модель возможности дефицита.

В рассматриваемой до сих пор модели предполагалось, что дефицит не допускается, т.е. некоторое количество товара на складе всегда есть. Но, может быть, выгоднее сэкономить на расходах по хранению запаса, допустив небольшой дефицит — потребность в товаре в некоторые интервалы времени может остаться неудовлетворенной?

Как подсчитать убытки от дефицита, в частности, от потери доверия потребителя? Будем считать, что если нет товара, то владеющая складом организация платит штраф — каждый день пропорционально нехватке. По приходе очередной поставки все накопленные требования сразу же удовлетворяются.

Сохраним все предположения и обозначения рассматриваемой до сих пор модели, кроме отсутствия дефицита. Неудовлетворенный спрос будем рассматривать как отрицательный запас. График изменения величины запаса на складе изображен на рис. 3.

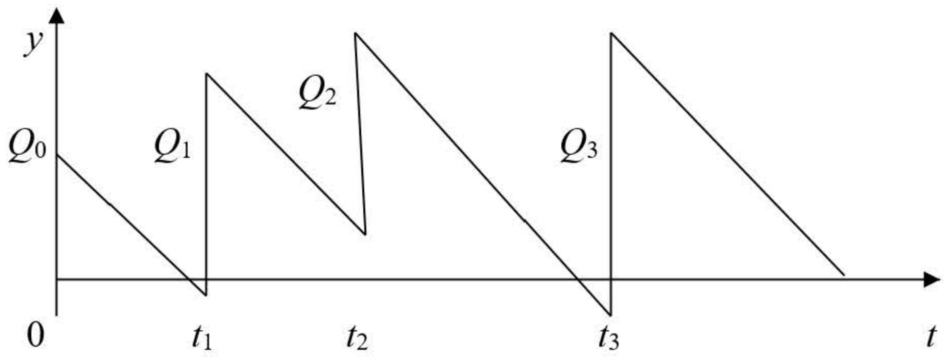


Рис. 3. График изменения величины запаса на складе
при возможности дефицита

Очевидно, рис.1 и рис.3 отличаются только тем, что на последнем рисунке зубцы графика могут опускаться ниже оси абсцисс, что соответствует сдвигу графика рис.1 как единого целого вниз вдоль оси ординат.

Пусть h — плата за нехватку единицы товара в единицу времени (например, в день). Тогда средние издержки за время T определяются формулой:

$$f_1(T,y) = f_1(y(t),\ 0 \le t \le T) = \frac{1}{T}\left\{ s\int_0^T y(t)\chi(y(t) \ge 0)dt + h\int_0^T |\,y(t)\,|\ \chi(y(t) < 0)dt + gn(T) \right\},$$

где $\chi(A)$ — индикатор множества A, т.е. $\chi(y(t){\ge}0) = 1$ при $y(t){\ge}0$ и $\chi(y(t){\ge}0){=}0$ при $y(t){<}0$, в то время как $\chi(y(t){<}0) = 1$ при $y(t){<}0$ и $\chi(y(t){<}0) = 0$ при $y(t){\ge}0$. Таким образом, площадь под частью графика уровня запаса, лежащей выше оси абсцисс, берется с множителем s, а площадь между осью абсцисс и частью графика $y(t)$, соответствующей отрицательным значениям запаса, берется с заметно большим по величине множителем h.

Для модели с дефицитом оптимальный план находится почти по той же схеме, что и для модели без дефицита. Сначала фиксируем моменты поставок и находим при этом условии оптимальные размеры поставок. Фактически речь идет о выборе уровня запаса Y в момент прихода очередной поставки (рис. 4).

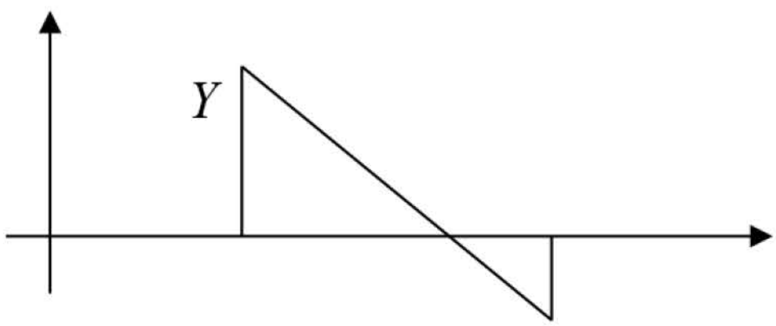


Рис. 4. Первый шаг построения оптимального плана в модели с дефицитом

Увеличивая или уменьшая Y, можно увеличивать или уменьшать площадь треугольника над осью абсцисс (учитываемую с коэффициентом s) и соответственно уменьшать или увеличивать площадь треугольника под осью абсцисс (учитываемую с коэффициентом h), добиваясь минимизации взвешенной суммы этих площадей. Все элементы прямоугольных треугольников на рис. 4 выражаются через Y, заданный интервал времени Δ между поставками и параметры модели. Минимизация соответствующего квадратного трехчлена дает оптимальное значение:

$$Y = \frac{h}{s+h}\mu\Delta.$$

При этом минимальная сумма затрат на хранение и издержек, вызванных дефицитом, равна:

$$\frac{\Delta^2\mu}{2}\frac{sh}{s+h}.$$

Второй шаг нахождения оптимального плана в модели с дефицитом полностью совпадает с аналогичным рассуждением в исходной модели. Фиксируется число поставок, и с помощью варьирования размеров интервалов между поставками минимизируется целевой функционал. Поскольку сумма квадратов некоторого числа переменных при заданной их сумме достигает минимума, когда все эти переменные равны между собой, то оптимальным планом является план, у которого все зубцы одинаковы, т.е. уровень запаса в момент прихода очередной поставки — всегда один и тот же. При этом все объемы поставок, за исключением начальной (нулевой), равны между собой:

$$Q = Q_1 = Q_2 = Q_3 = \ldots, Q_0 = \frac{h}{s+h}Q. \tag{13}$$

На третьем этапе среди указанного однопараметрического дискретного множества планов находим оптимальный. Как и для модели без дефицита, в качестве ориентира используется план с размером поставки, определяемой по формуле квадратного корня,

$$Q_0(\mu, g, s, h) = \sqrt{\frac{2\mu g(s+h)}{sh}}.$$

Для горизонтов планирования T, кратных $Q_0(\mu, g, s, h)/\mu$, оптимальным является план типа (13) с $Q = Q_0(\mu, g, s, h)$. Для всех остальных горизонтов планирования, как и в случае модели без дефицита, необходимо найти неотрицательное целое число n такое, что:

$$Q_1 = \frac{\mu T}{n+1} < Q_0(\mu, g, s, h) < \frac{\mu T}{n} = Q_2,$$

а затем, сравнив издержки для $Q = Q_1$ и $Q = Q_2$, объявить оптимальным то из этих двух значений, для которого издержки меньше.

Отметим, что модель без дефицита является предельным случаем для модели с дефицитом при безграничном возрастании платы за дефицит. В частности,

$$\lim_{h \to \infty} Q_0(\mu, g, s, h) = \sqrt{\frac{2\mu g}{s}}.$$

Как и в случае модели без дефицита, план с объемом поставки, определяемой по формуле квадратного корня, $Q = Q_0(\mu, g, s, h)$, является асимптотически оптимальным.

Система моделей на основе модели Вильсона. Классическая модель теории управления запасами, называемая также моделью Вильсона, допускает различные обобщения.

Одно из таких обобщений — модель с конечной скоростью поставки υ, т.е. модель, в которой за время Δt поставляется продукция объемом $\upsilon\Delta t$ (при наличии в то же время постоянного спроса с интенсивностью μ, причем считается, что $\upsilon>\mu$). Таким образом, в этой модели поставка происходит не мгновенно, а в течение некоторого интервала времени, причем объем поставляемой продукции линейно зависит от времени. Такие поставки будем называть линейными с интенсивностью υ.

Другое обобщение классической модели связано с обобщением функции от объема запаса, задающей плату за хранение. В исходной модели считалось, что расходы за хранение пропорциональны объему продукции на складе. Естественно считать, что эти расходы должны содержать постоянный член a, не зависящий от объема продукции на складе (расходы на содержание самого склада, оплату работников и т.д.). Однако оптимальный план при таком обобщении не изменится. Действительно, в формуле для издержек добавится постоянный член a, и положение минимума не изменится при его добавлении.

Однако в модели с дефицитом ситуация иная. Затраты на хранение возникают только при наличии товара на складе, и издержки этого вида вполне естественно разделить на постоянные и переменные (пропорциональные объему запаса на складе).

Аналогично издержки, вызванные дефицитом, вполне естественно разделить на постоянные (вызванные самим фактом дефицита) и переменные (пропорциональные величине дефицита).

В классической модели плата за доставку партии не зависит от объема партии. Здесь используются только постоянные издержки. Представляется вполне естественным ввести линейный член, соответствующий возрастанию платы за доставку в зависимости от величины партии (переменные издержки). (Ниже будет показано, что добавление этого члена не влияет на решение задачи оптимизации и вид оптимального плана.) Дальнейшее обобщение — введение скидок в зависимости от величины партии. Это приводит к выражению платы за доставку в виде квадратного трехчлена от объема партии.

Можно рассматривать одновременно несколько обобщений. В результате получаем систему моделей на основе классической модели управления запасами, состоящую из 36 моделей [16]. Каждая из них может быть описана набором четырех чисел ($a(1)$, $a(2)$, $a(3)$, $a(4)$). Каждое из этих чисел соответствует одному из рассмотренных выше видов обобщений исходной модели.

При этом $a(1) = 0$, если поставки мгновенные, и $a(1) = 1$, если поставки являются линейными с интенсивностью υ, причем $\upsilon>\mu$.

Если плата за хранение продукции объемом y в течение единицы времени равна sy, то $a(2) = 0$. Если же учтены постоянные (при наличии товара на складе) издержки, т.е. указанная плата равна $sy+a$, $a>0$, то $a(2) =1$.

Если плата за нехватку продукции объемом y в течение единицы времени бесконечна (т.е. дефицит не допускается), то $a(3) = 0$. Если эта плата равна hy (рассмотренная выше модель с дефицитом), то $a(3)=1$. Если же вводятся также постоянные издержки (плата за само наличие дефицита), т.е. плата за нехватку продукции объемом y в течение единицы времени равна $hy + b$, $b > 0$, то $a(3) = 2$.

Наконец, $a(4) = 0$, если плата за доставку партии продукции объемом Q равна g. Если учитываются переменные издержки, т.е. эта плата равна $g + g_1Q$, то $a(4) = 1$. Если же в модели учитываются скидки на объем партии, т.е. если плата за доставку партии продукции объемом Q равна $g + g_1Q + g_2Q^2$, то $a(4) = 2$.

Для $a(1)$ имеется два возможных значения, для $a(2)$ — тоже два, для $a(3)$ — три возможных значения, для $a(4)$ — тоже три. Всего имеется $2 \cdot 2 \cdot 3 \cdot 3 = 36$

возможных комбинаций, т.е. 36 возможных моделей. Классическая модель управления запасами описывается набором (0, 0, 0, 0), а модель с дефицитом — набором (0, 0, 1, 0).

Рассмотрим наиболее обобщенную модель рассматриваемой системы. Она описывается набором (1, 1, 2, 2). Можно показать, что для нее справедливы основные утверждения, касающиеся классической модели и модели с дефицитом. Однако «формула квадратного корня» имеет более сложный вид, а именно,

$$Q_0(\mu,\nu,s,a,h,b,g,g_1,g_2)=\sqrt{\frac{\mu g-\dfrac{(a-b)^2}{2(s+h)}\left(\dfrac{1}{1-\dfrac{\mu}{\nu}}\right)}{\dfrac{sh}{2(s+h)}\left(1-\dfrac{\mu}{\nu}\right)+\mu g_2}}.$$

В частности, план с $Q = Q_0(\mu,\nu,s,a,h,b,g,g_1,g_2)$ является асимптотически оптимальным.

Формула для $Q_0(\mu,\nu,s,a,h,b,g,g_1,g_2)$ позволяет обнаружить ряд любопытных эффектов. Так, в ней не участвует параметр g_1. Другими словами, при любом изменении этого параметра оптимальный объем поставки не меняется. Если запас пополняется весьма быстро по сравнению со спросом, т.е. $\upsilon \gg \mu$, то соответствующий множитель в «формуле квадратного корня» исчезает, и для моделей с $a(1) = 0$ получаем более простую формулу:

$$Q_0(\mu,+\infty,s,a,h,b,g,g_1,g_2)=\sqrt{\frac{\mu g-\dfrac{(a-b)^2}{2(s+h)}}{\dfrac{sh}{2(s+h)}+\mu g_2}}.$$

Дальнейшее упрощение получаем при $a = b$. Это равенство означает, что постоянные (в другой терминологии — фиксированные) платежи за хранение и в связи с дефицитом совпадают, например, равны 0. Если последнее утверждение справедливо, то:

$$Q_0(\mu,+\infty,s,0,h,0,g,g_1,g_2)=\sqrt{\frac{\mu g}{\dfrac{sh}{2(s+h)}+\mu g_2}}.$$

Предположим теперь, что при доставке партии отсутствуют скидки (или надбавки) за размер партии. Тогда «формула квадратного корня» упрощается дальше и приобретает вид:

$$Q_0(\mu,+\infty,s,0,h,0,g,g_1,0)=\sqrt{\frac{\dfrac{\mu g}{sh}}{\sqrt{\dfrac{}{2(s+h)}}}}.=\sqrt{\frac{2\mu g(s+h)}{sh}}.$$

Эта формула уже была получена выше при рассмотрении модели с дефицитом. При безграничном возрастании h получаем формулу Вильсона для классической модели управления запасами:

$$Q_0(\mu,+\infty,s,0,+\infty,0,g,g_1,0)=\sqrt{\frac{2\mu g}{s}}.$$

Новое в последних двух формулах — наличие в левой части параметра g_1, не участвующего в формировании объема партии.

Важное замечание 3. Модели конкретных экономических (и не только) процессов и явлений обычно не встречаются и не изучаются поодиночке. Обычно имеется совокупность моделей, объединенных в систему, переходящих друг в друга при тех или иных предельных переходах. Часто более простые модели используются для конкретных расчетов, в то время как более сложные применяются для изучения точности, достигаемой с помощью этих более простых моделей, согласно подходу, развитому в [3, 5].

О практическом применении классической модели управления запасами. Для отработки методики практического использования классической модели управления запасами был проведен эксперимент на снабженческо-сбытовой базе, а именно, на Реутовской химбазе (Московская область). Собраны и обработаны данные по одному из товаров, распространяемых этой организацией в большом объеме, — по кальцинированной соде. В качестве исходной информации о спросе использовались данные об ежедневном отпуске кальцинированной соды потребителям, зафиксированные на карточках складского учета. Рассчитана величина затрат на хранение как соответствующая доля общей суммы издержек по содержанию базы, а также расходы на доставку новых партий. Для определения расходов на хранение запасов использованы данные о заработной плате складского персонала (включая основную и дополнительную заработная плата, начисления на зарплату), расходах на содержание охраны, эксплуатацию складских зданий и сооружений, расходах по текущему ремонту, по таре, на приемку, хранение, упаковку и реализацию товаров, о величине амортизационных отчислений и др. Для расчета расходов на доставку новых партий товара использованы данные о расходах по завозу, о плате за пользование вагонами и контейнерами сверх установленных норм, расходах на содержание и эксплуатацию подъемно-транспортных механизмов, о заработной плате работников, занятых в процессе доставки товара, канцелярских, почтовых и телеграфных расходах и др.

Полезным оказалось вытекающее из «принципа уравнивания погрешностей» соотношение (12). Интенсивность спроса μ и погрешность определения этого параметра найдены методом наименьших квадратов. Это дало возможность установить величину относительной точности определения параметров модели, вытекающих из величин погрешностей исходных данных для спроса. Параметры классической модели управления запасами g и s оценивались двумя способами — по методике Всесоюзного института материально-технического снабжения и по методике Центрального экономико-математического института АН СССР. Для каждой из методик с помощью соотношения (12) были определены абсолютные погрешности определения параметров g и s. Оказалось, что для каждой из методик интервалы $(s - \Delta s, s + \Delta s)$ и $(g - \Delta g, g + \Delta g)$ таковы, что числа, рассчитанные по альтернативной методике, попадают внутрь этих интервалов. Это означает, что для определения параметров g и s можно пользоваться любой из указанных методик (в пределах точности расчетов, заданной наблюдаемыми колебаниями спроса).

Вызванное отклонениями параметров модели в допустимых пределах максимальное относительное увеличение суммарных затрат на доставку и хранение продукции не превосходило 26 % (колебания по кварталам 22,5–25,95 %). Фактические издержки почти в 3 раза превышали оптимальные (в зависимости от квартала фактические издержки составляли 260–349 % от оптимального уровня). Следовательно, внедрение модели Вильсона в практику управления запасами на Реутовской химбазе дает возможность снизить издержки, связанные с доставкой и хранением кальцинированной соды не менее чем в 2 раза [17].

Таким образом, несмотря на то, что параметры модели определены неточно и отклонения значений параметров (от тех значений, по которым рассчитывается оптимальный план поставок) приводят к некоторому увеличению затрат по сравнению с затратами в оптимальном плане, использование рассматриваемой модели для реального управления запасами конкретной продукции может дать значительный экономический эффект. Аналогичным является положение со многими другими моделями управления запасами. Это утверждение подтверждает и зарубежный опыт, проанализированный в монографии [5].

Двухуровневая модель управления запасами. Создание любой информационной (в более старой терминологии — автоматизированной) системы управления материально-техническим снабжением (в другой терминологии — процессами логистики), базирующейся на комплексе экономико-математических моделей, должно включать в себя разработку (в качестве блоков) моделей деятельности отдельных баз (складов). Поэтому большое внимание уделяется проблеме построения оптимальной политики управления запасами на базе (складе). Оптимизационную экономико-математическую теорию удается развивать в основном для однопродуктовых моделей.

Двухуровневая модель управления запасами — это однопродуктовая модель работы склада, в которой заявки потребителей удовлетворяются мгновенно. При отсутствии продукта заявки учитываются. Как только запас на складе опускается до уровня $R < 0$, мгновенно поступает партия товара величиной Q и запас на складе оказывается равным $R+Q>0$. Как и в рассмотренном выше варианте классической модели Вильсона с дефицитом, издержки складываются из издержек по хранению, издержек от дефицита и издержек по доставке. Средние издержки за время T имеют вид:

$$f_1(T,y) = f_1(y(t),\ 0 \le t \le T) = \frac{1}{T}\left\{ s\int_0^T y(t)\chi(y(t) \ge 0)dt + h\int_0^T |\,y(t)\,|\ \chi(y(t) < 0)dt + gn(T) \right\},$$

где $y(t)$ — уровень запаса на складе, $\chi(A)$ — индикатор множества A, т.е. $\chi(y(t)\ge 0) = 1$ при $y(t)\ge 0$ и $\chi(y(t)\ge 0)=0$ при $y(t)<0$, в то время как $\chi(y(t)<0) = 1$ при $y(t)<0$ и $\chi(y(t)<0) = 0$ при $y(t)\ge 0$, параметры модели s, h, g имеют тот же смысл, что и выше. Оптимизация состоит в определении значений нижнего уровня R и верхнего уровня $R+Q$, минимизирующих средние издержки.

В 1950-х гг. американский исследователь К. Эрроу (в будущем — нобелевский лауреат по экономике) с сотрудниками показал, что в ряде случаев оптимальная политика управления запасами — это политика, основанная на двухуровневой модели. Этот принципиально важный теоретический результат стимулировал развитие исследований свойств двухуровневой модели. Однако окончательная теория была построена только в конце 1970-х гг. [5].

Важными являются характеристики потока заявок. Пусть $\tau(T)$ — число заявок за время T. Эта величина предполагается случайной. С прикладной точки зрения вполне естественно предположить, что математическое ожидание $M\tau(T)$ конечно. Накопленный спрос за время T имеет вид:

$$X(T) = X_1 + X_2 + \ldots + X_{\tau(T)},$$

где X_j — величина j-ой заявки. Предполагается, что $X_1, X_2, \ldots, X_n, \ldots$ — последовательность независимых одинаково распределенных случайных величин с математическим ожиданием MX_1. Таким образом, накопленный спрос за время T является суммой случайного числа случайных слагаемых. Накопленный спрос определяет уровень запаса на складе, поэтому математический аппарат изучения двухуровневой модели — это предельная теория сумм случайного числа случайных слагаемых.

При некоторых условиях регулярности (выполняющихся для реальных систем управления запасами) в [5] найдены оптимальные (для горизонта планирования T) значения нижнего и верхнего уровней:

$$R_0(T) = -\sqrt{\frac{2gsM\tau(T)MX_1}{Th(s+h)}},$$

$$Q_0(T) = \sqrt{\frac{2g(s+h)M\tau(T)MX_1}{Tsh}}.$$

Часто можно принять, что число поступающих заявок обладает некоторой равномерностью. Например, вполне естественно принять, что:

$$\lim_{T\to\infty}\frac{M\tau(T)}{T} = \lambda$$

при некотором λ. Здесь λ — параметр, описывающий предельную (т.е. асимптотическую при $T \to +\infty$) интенсивность спроса. Тогда асимптотически оптимальные уровни имеют вид:

$$R_0 = -\sqrt{\frac{2gs\lambda MX_1}{h(s+h)}},$$

$$Q_0 = \sqrt{\frac{2g(s+h)\lambda MX_1}{sh}}.$$

Отметим, что асимптотическое распределение уровня запаса на складе — равномерное на отрезке $[R, R+Q]$.

Модель планирования размеров поставок на базу (склад). В двухуровневой модели накопленный спрос в любой момент времени является случайной величиной. Это не всегда соответствует экономической реальности. Достаточно часто в соответствии с заключенными договорами размеры поставок на базу и объемы запрашиваемой потребителями продукции определены до начала года (с разбивкой по кварталам или по месяцам) и затем не меняются. Однако поставщик имеет право отгружать продукцию, а потребители — забирать ее в течение всего квартала (или месяца).

Опишем соответствующую однопродуктовую модель [18]. Пусть интервал планирования разбит на m периодов, не обязательно одинаковых по продолжительности. В течение каждого периода приходит на базу одна поставка. В i-й период ее величина равна H_i, а момент поступления — случайная величина $\tau(i)$ с функцией распределения $G(i,t)$, $0 \leq t \leq 1$, где t — отношение времени, прошедшего с начала i-го периода, к продолжительности его, $i = 1, 2, \ldots, m$.

В i-й период имеется $n(i)$ потребителей, получающих с базы строго определенное количество продукта, $c(1,i)$, $c(2,i)$, …, $c(n(i),i)$ соответственно. Моменты поступления требований от потребителей — случайные величины $\delta(i,j)$, $j = 1, 2, \ldots, n(i)$, $i = 1, 2, \ldots, m$, с функциями распределения $F(i,j,t)$, $0 \leq t \leq 1$, где t — отношение времени, прошедшего после начала соответствующего периода, к продолжительности этого периода. Если в момент прихода требования на базе имеется достаточное количество продукта, то он отпускается мгновенно. Если продукта нет, то потребителю придется ждать очередной поставки. Если продукта недостаточно, то весь оставшийся товар отпускается сейчас же, а оставшуюся часть приходится ждать.

В течение i-го периода, $i = 1, 2, \ldots, m$, все моменты поступления товара и требований $\tau(i)$, $\delta(i,j)$, $j = 1, 2, \ldots, n(i)$, предполагаются независимыми в совокупности. Потери, как обычно, складываются из издержек по хранению и от дефицита (расходы на доставку партий заданы заранее, т.е. постоянны, а потому их можно не включать в минимизируемый функционал). Издержки по хранению предполагаются пропорциональными времени хранения и величине запаса с коэффициентами пропорциональности $s(i)$, $i = 1, 2, \ldots, m$. Издержки от дефицита складываются из потерь у каждого из потребителей; они пропорциональны величине и длительности дефицита с коэффициентами пропорциональности $h(i,j)$, $j = 1, 2, \ldots, n(i)$, $i = 1, 2, \ldots, m$.

Пусть $x(0)$ — начальный запас, $x(i)$ — количество продукта на базе в конце i-го периода, $i = 1, 2, \ldots, m$. Пусть $S(i) = \{s(i), c(j,i), \mathrm{h}(i,j), G(i,t), F(i,j,t), 0 \le t \le 1, j = 1, 2, \ldots, n(i)\}$ — исходные данные модели в i-й период. Как легко видеть, математическое ожидание издержек за i-й период зависит только от $x(i-1)$, $x(i)$ и $S(i)$. Для краткости обозначим его через $f(x(i-1), x(i), S(i))$. Тогда математическое ожидание издержек за m периодов равно:

$$Z(m) = f(x(0), x(1), S(1)) + f(x(1), x(2), S(2)) + \ldots$$
$$+ f(x(i-1), x(i), S(i)) + \ldots + f(x(m-1), x(m), S(m)).$$

Необходимо минимизировать $Z(m) = Z(x(0), x(1), \ldots, x(i), \ldots, x(m))$ по совокупности переменных. Таким образом, необходимо найти оптимальные значения уровней запаса на складе в начале и в конце периодов. Это эквивалентно определению оптимальных размеров поставок по периодам и начального запаса. Ограничения рассматриваемой оптимизационной задачи выписаны в [5, 18].

Вначале была сделана попытка рассматривать задачу минимизации $Z(m)$ как задачу динамического программирования и решать ее типовыми методами. Однако вычислительных мощностей оказалось недостаточно для выполнения расчетов. Тогда нам удалось показать, что функция $(m + 1)$-го переменного $Z(m)$ в действительности является суммой $(m + 1)$ функции одного переменного. Действительно,

$$f(x(i-1), x(i), S(i)) = f_1\,(x(i-1), x(i), S(i)) + f_2\,(x(i-1), x(i), S(i)),$$

где $f_1(x(i-1), x(i), S(i))$ — математическое ожидание затрат, произведенных до прихода очередной поставки, $f_2(x(i-1), x(i), S(i))$ — математическое ожидание затрат после поступления поставки.

Ясно, что $f_1(x(i-1), x(i), S(i))$ определяется запасом на начало периода и спросом до прихода поставки, но не зависит от запаса на конец периода, т.е. от $x(i)$. Таким образом, можно записать:

$$f_1(x(i-1), x(i), S(i)) \equiv f_1(x(i-1), S(i)).$$

Пусть H_i — объем поставки на склад в i-й период. Сразу же после прихода поставки запас y на складе равен:

$$y(\tau(i)) = x(i-1) + H_i - \xi(\tau(i)) = x(i) + \sum_{1 \le j \le n(i)} c(j,i) - \xi(\tau(i)),$$

где $\xi(\tau(i))$ — накопленный с начала периода спрос. Поскольку $\xi(\tau(i))$ не зависит от $x(i–1)$, то и $f_2(x(i–1), x(i), S(i))$ не зависит от $x(i–1)$. Итак,

$$f_2(x(i-1), x(i), S(i)) \equiv f_2(x(i), S(i)).$$

Следовательно, минимизируемая функция имеет вид:

$$Z(m) = f_1(x(0), S(1)) + \sum_{1 \le i \le m-1} \{f_2(x(i), S(i)) + f_1(x(i), S(i+1))\} + f_2(x(m), S(m)).$$

При этом ограничения наложены на каждую переменную $x(i)$ по отдельности [5, 18]. Ясно, что задача минимизации $Z(m)$ распадается на $m + 1$ задачу минимизации функций одной переменной:

$$\begin{aligned} &f_1(x(0), S(1)) \to \min, \\ &f_2(x(i), S(i)) + f_1(x(i), S(i+1)) \to \min, \\ &f_2(x(m), S(m)) \to \min \end{aligned} \tag{14}$$

(ограничения не указаны). Следовательно, $x(k)$ зависит только от исходных данных смежных периодов $S(k)$ и $S(k + 1)$ и остается неизменным при любом изменении $S(i)$, $i \ne k$, $i \ne k+1$. Из указанного разложения задачи многомерной оптимизации на ряд задач одномерной оптимизации вытекает также, что при планировании на $m(1)$ и $m(2)$ периодов совпадают оптимальные значения начального запаса и поставок за первые $\min\{m(1), m(2)\} - 1$ периодов. В частном случае стационарного режима $S(i) = S$, $i = 1, 2, \ldots, m$, оптимальный план имеет вид $\{a, b, b, \ldots, b, \ldots, b, c\}$, где a — решение первой из указанных в (14) задач, b — решение второй задачи и c — третьей.

Переход к задачам (14) не только позволяет решить исходную задачу минимизации (напомним, что для минимизации задачи в исходной форме не хватало вычислительных мощностей), но также получить весьма важный для экономической интерпретации вывод о независимости оптимальных значений поставок и начального запаса от горизонта планирования m.

Важное замечание 4. Рассмотренная модель дает хороший пример пользы математического анализа оптимизационной задачи при принятии решений. Такой анализ позволяет решать задачу не стандартными методами, требующими больших вычислительных ресурсов, а с помощью специально разработанных алгоритмов, учитывающих специфику задачи и позволяющих весьма существенно сократить вычисления. Плата за экономию вычислительных ресурсов — необходимость квалифицированного труда специалистов по экономико-математическим методам и прикладной математике.

В настоящее время логистика — одна из экономических дисциплин, весьма развитая как в теоретическом, так и в практическом отношении. В ней рассматривается масса конкретных моделей управления запасами. Из перспективных направлений назовем использование случайных множеств в моделях логистики. Моделирование задач логистики с целью нахождение оптимальных решений выше было продемонстрировано на примерах системы моделей, исходящих из классической модели Вильсона, двухуровневой модели, модели оптимизации объемов поставок на базу (склад). Этот перечень не охватывает даже малой доли математических моделей логистических процессов. Поскольку в реальных процессах движения материальных финансовых и информационных потоков всегда присутствует неопределенность, большое значение для логистики имеют вероятностно-статистические модели [5].

8.4. Статистические модели в истории

В разделе рассмотрены статистические модели и методы анализа текстовой информации, находящейся в исторических хрониках (т. е. любых повествованиях об исторических событиях, расположенных в порядке их следования по годам). Цель анализа исторических данных — установление моментов времени, когда происходили те или иные события, т.е. построение научно обоснованной хронологии. Рассматриваемые статистические модели и методы нацелены на выявление хроник, рассказывающих об одних и тех же событиях (так называемых «дубликатов»).

Постановка задачи. Если вдумываться в факты, сообщаемые в посвященных истории сочинениях, то возникают недоуменные вопросы. Пожалуй, основной из них — как могли погибнуть великие древние цивилизации и наступить «темные века» раннего средневековья, когда достижения предыдущих столетий были практически полностью забыты.

Другой интересный вопрос — как осознать появление «Эпох Возрождения» (Италия, Египет), когда через полторы тысячи лет повторяются основные черты эпохи — язык, одежда, литературный стиль. Например, нам говорят, что сочинения на латыни достигли высокого литературного уровня в Древнем Риме. Читаем — действительно, это так. Затем в средние века уровень владения языком упал, латынь стала «варварской». Никаких литературных опусов, корявые перечни и записки. И вдруг — снова расцвет классической латыни — в века итальянского Возрождения. Сопоставим с нашей историей. Фактически современная русская культура начинается около двухсот лет назад — с Пушкина, первого поэта, и Карамзина, первого историка. Все, что было до этого, уже трудно воспринимать. Фактически нужен перевод на современный литературный язык. А если отступить еще на 100–200 лет вглубь времен, то чтение превращается в расшифровку текста в буквальном смысле слова. И не только у нас. Язык Шекспира за 400 лет тоже достаточно сильно оторвался от современного английского языка. Можно ли поверить, что мы вдруг перейдем на язык V в. н.э.? Или хотя бы на язык «Слова о полку Игореве»? Можно ли поверить, что мы будем одеваться так, как в Древней Руси? Ответ очевиден. Но как же тогда быть с итальянским Возрождением?

Для того, чтобы поставить ряд недоуменных вопросов, нет необходимости быть историком. Достаточно тщательно сопоставить страницы распространенных учебников. Это сделал, например, чемпион мира по шахматам Гарри Каспаров. Его размышления помещены в качестве предисловия к монографии [19].

На недоуменные вопросы должны быть даны ответы. Историческая наука продолжает развиваться. И один из видов ее интеллектуальных инструментов — статистические модели и методы.

Откуда мы знаем, когда происходили древние события? Всем рассказывали в школе о событиях, которые произошли тысячи лет назад. А откуда такая уверенность, что, например, битва при Марафоне между греками и персами была именно в 490 г. до н.э.?

Очевидно, есть три основные вида источников знаний о древней истории: старинные тексты, остатки материальной культуры и — психологически самое важное! — сложившаяся традиция. Обсудим их.

Древние тексты требуют критического анализа. Их не всегда легко перевести на современный язык. Более того, многие тексты не удается расшифровать. Они отнюдь не всегда точны и беспристрастны. Например, они отражают точку зрения победившей стороны. И совершенно ясно, что современник Марафонской битвы не может ее дату отсчитывать от Рождества Христова, которое — по традиционной хронологии — будет позже примерно на 500 лет! Очевидно, эта дата — 490 г. до н.э. — поставлена теми, кто составлял глобальную хронологию, а не современниками. А составители могли ошибаться. Могли и фальсифицировать даты, если это было кому-то выгодно. Это касается и даты создания самого «древнего текста».

На археологических остатках тоже не написано, к какому времени они относятся. Можно указать верхний слой раскопок, средний, нижний, то есть выяснить, что из найденного в данном месте относится к более раннему времени, а что к более позднему. Уже при сопоставлении с раскопками в других местах возникают проблемы: похоже или не похоже, можно отнести к одному и тому же времени или нет. Тут часто решает традиция.

Бесспорно, что конкретную дату добытых археологами предметов в подавляющем большинстве случаев установить невозможно. Это утверждение пытались поставить под сомнение, ссылаясь на различные естественнонаучные методы датировки. Наиболее известным таким методом является радиоуглеродный анализ. К сожалению, его погрешности велики (плюс-минус 1 000 лет!). Выяснилось, что все известные естественнонаучные методы датировки практически бесполезны на том отрезке истории человечества, о котором сохранились письменные источники (подробнее см. [20, с. 33–39]).

Сложившаяся традиция — огромная сила. Вспомним, как она появилась. Около 200 лет назад император России дал задание одному из лучших литераторов подготовить изложение истории России для широкого читателя. А широкий читатель тогда — несколько тысяч дворян и лиц духовного звания, с вкраплением отдельных разночинцев. Раньше исторические тексты были, как мы сказали бы теперь, секретными документами, хотя время от времени и выпускались сочинения от имени Татищева и Ломоносова, почему-то после смерти авторов.

И появилась «История» Карамзина. И одновременно была создана сеть гимназий во всех губернских городах, другие учебные заведения. Во всех преподавалась российская история — по учебникам на основе сочинений Карамзина. Так и пошло. Иногда менялись акценты, особенно в 20-е гг. XX в., но фактическая основа не менялась.

Традиция преподавания всеобщей и российской истории, с одной стороны, молода — ей не более 200 лет. С другой стороны, за это время сменилось примерно 8 поколений учеников, а этого вполне достаточно, чтобы традиционная хронология стала выглядеть единственно возможной.

Пример возможности ошибочной датировки. Недавно автор этой книги убедился, насколько трудно восстановить правильную последовательность событий. Надо было перенести свой архив с старого компьютера на новый. Всего-то около тысячи файлов. Среди них были «гнезда», посвященные подготовке одного и того же документа (статьи или отчета) — наборы файлов под сходными названиями и со схожим содержанием, отличающиеся лишь некоторыми деталями — исправлениями и дополнениями. К сожалению, я не всегда уничтожал промежуточные версии. Впрочем, почему к сожалению? Иногда окончательную версию приходилось сокращать, и ряд нужных соображений оставался только в промежуточном тексте.

Обратите внимание, что после копирования файлов на новый компьютер даты создания файлов изменились — теперь в машинных описаниях файлов хранились даты их появления на новом компьютере, а не первоначального создания на на старом компьютере. И вот пришел день, когда понадобилось обратиться к одному из «гнезд», чтобы отредактировать статью в соответствии с рецензией, пришедшей из редакции журнала. Просмотрел «гнездо» — нашел два текста с одинаковым названием. Объемы близки. В свойствах файлов значится, что они созданы одновременно — в тот самый день, когда я копировал их на новый компьютер. Информация о дате последнего изменения файлов помочь не могла — при переносе файлов между двумя различными платформами они подвергаются специальной конвертации, и тем самым изменяются в тот же день и час, когда и переносятся. Какой же текст окончательный, а какой предварительный? Так и не смог определить. Стал работать с одним из них. И, видимо, ошибся. Думаю так, поскольку в компьютерном тексте были опечатки, исправленные в статье, побывавшей в редакции.

Таким образом, я не смог понять, какой из двух компьютерных текстов окончательный, а какой является лишь заготовкой, — при всех имеющихся в программном обеспечении средствах для определения даты создания. Насколько же труднее было сравнивать тексты хроник пятьсот лет назад!

Предположим теперь, что кто-то захочет составить собрание моих сочинений на основе памяти компьютера. Перед ним сразу встанет проблема «гнезд», в которых находятся последовательные версии одного и того же документа, имеющие одну и ту же дату создания и изменения. Лучше всего было бы найти окончательную версию и включить ее в собрание сочинений, а остальные проигнорировать. Но это может быть нелегким делом, ведь я сам, автор документа, не могу указать окончательную версию, по крайней мере быстро. Возникает желание опубликовать все версии. Кстати, именно так я поступил при переносе архива. Ведь я не стал проводить отбор, а отложил проблему и перенес все файлы.

Предположим теперь, что публикатор упорядочит файлы не по «гнездам», а по какому-либо иному признаку — по названиям, по объему или как-либо еще. И вот мы получаем сборник текстов, некоторые из которых близки по смыслу. А именно, близки те, что «произошли» из одного «гнезда». Но читатель-то не знает, какие тексты имеют общее происхождение, а какие — нет! Поэтому он, скорее всего, подумает, что у автора много раз возрождался интерес к одной и той же теме, автор, так сказать, переживал «период возрождения».

Таким образом, проблемы построения правильной хронологии легко промоделировать на примере типовых проблем рядового пользователя компьютера.

Сложившаяся традиция: истоки. Объективный анализ первоисточников показывает, что ныне принятая версия всеобщей хронологии исторических событий была сформулирована сравнительно недавно — лишь в начале XVII в. Это — период «смуты» на Руси. А известная всем нам со школы история России подготовлена немецкими специалистами при российском императорском дворе еще на полтораста лет позже — в XVIII в. Недаром она была представлена Карамзину для изложения популярным языком лишь двести лет назад. За подробностями отошлем к многочисленным публикациям группы А.Т. Фоменко (см., например, солидные монографии [19–25]).

Конечно, ныне принятая хронология готовилась долго, столетиями. Примерно за триста лет, в XIV–XVI вв. западноевропейскими хронологами была проведена огромная работа, в основных чертах завершенная И. Скалигером (1540–1609) и Д. Петавиусом (1583–1652). Необходимо было сопоставить между собой многочисленные хроники, написанные на разных языках, относящиеся к различным государствам, пользующимися своими собственными системами отсчета времени (от начала очередного царствования, от основания Рима, от первых олимпийских игр, и т.д.). Результат (назовем его хронологией Скалигера) — в наших учебниках.

Могли ли быть допущены ошибки при построении хронологии Скалигера? Конечно, могли, причем, по крайней мер, по двум различным причинам. Первая состоит в ошибочной датировке хроник, в неумении различить хроники-дубликаты, говорящие об одних и тех же событиях. Вторая — в сознательном внесении искажений с целью идеологического обоснования тех или иных положений. Например, работа по хронологии шла под патронатом католической церкви, которой было выгодно «удревнить» историю Италии, тем самым поставить ее выше «молодой» Руси. Это — типичный пример информационной войны, которая в средневековье велась не менее интенсивно, чем сейчас, хотя и затрагивала, прежде всего, элиту.

Сложившаяся традиция: критика. Хронология Скалигера сразу же стала подвергаться критике. Одна из причин — противоречия между данными истории и астрономии. В исторических сочинениях под определенными датами описываются астрономические явления — затмения, расположение планет среди созвездий и т.п. Астрономия — точная наука, и ее методами можно рассчитать возможные даты тех событий, о которых идет речь в исторической хронике. Иногда эти даты несовместимы. Тогда естественным является желание исправить хронологию, перенеся событие в тот момент времени, когда оно астрономически возможно. А за этим событием «тянутся» все с ним связанные.

Наиболее известными критиками хронологии Скалигера являются великий физик и математик Исаак Ньютон (1642–1727) и выдающийся русский ученый-энциклопедист Николай Александрович Морозов (1854–1946), почетный академик АН СССР. В семитомном издании «Христос» (первоначальное название — «История человеческой культуры в естественнонаучном освещении»), выпущенном в 1924–1932 гг., Н.А. Морозов выдвинул и частично обосновал гипотезу о том, что хронология Скалигера искусственно растянута, удлинена по сравнению с подлинной историей. Он обнаружил «повторы в истории», указал на древние хроники, описывающие одни и те же события, но датированные при конструировании скалигеровской хронологии разными эпохами, которые считаются сегодня отделенными друг от друга сотнями и тысячами лет.

Работы И. Ньютона и Н.А. Морозова (и многих иных, менее нам известных — де Арсилла, Ж. Гардуина, Р. Балдауфа, Э. Джонсона и др.) никто не смог опровергнуть. О них предпочли забыть. Их замалчивают, а если человек слишком известен, как Исаак Ньютон — вскользь говорят как о заблуждениях великого ума.

Во второй половине XX в. появилась возможность применить современные статистические методы, основанные на использовании мощных компьютеров. В 1970-е гг. были получены основные результаты нового направления исторической науки — статистической хронологии. В течение следующего десятилетия, в 1980-е гг., специалисты по прикладной статистике неоднократно и подробно обсуждали новые математические методы анализа летописей и других исторических источников, разработанные группой академика РАН А.Т. Фоменко. Затем в 1990-е гг. опубликована серия более чем из 10 монографий, для широкого круга читателей описывающая результаты применения этих методов (некоторые из этих монографий указаны в списке литературы).

В трудах научного коллектива под руководством академика РАН А.Т. Фоменко восстановлены основные черты реальной хронологии. Можно, конечно, называть новую хронологию гипотезой, но, во всяком случае, эта гипотеза более обоснована, чем альтернативная, даваемая в стандартных учебниках. Новая компьютерная математико-статистическая хронология всеобщей и российской истории, построенная группой академика РАН А.Т. Фоменко, оказалась полезной и для обсуждения современных экономических и политических проблем взаимоотношений России и Запада на пороге XXI века.

Компьютерный анализ исторических текстов. Группа А.Т. Фоменко разработала и применила новые статистические методы анализа исторических текстов (хроник). Эти методы основаны на интенсивном использовании компьютерных технологий, которых, конечно, не было ни у И. Ньютона, ни у Н.А. Морозова. Именно из-за необходимости проводить обширные вычисления И. Ньютон и Н.А. Морозов не смогли проанализировать весь хронологический материал. Они были вынуждены ограничиться отдельными расчетами и на их основе формулировать свои предположения и выводы.

Обработка текстов проводилась объективными (формальными) методами — любые другие исследователи, воспользовавшись теми же методиками, получат точно такие же результаты. Тем самым их работа носит строгий научный характер. Она восстанавливает классическое понимание хронологии как раздела прикладной математики. И именно современные компьютерные технологии, позволяющие проделать огромный объем расчетов за достаточно короткое время, позволили применить эти объективные методы и получить результаты.

Объект изучения — существующие в настоящее время исторические хроники, описывающие события год за годом. Типичная хроника — древнерусская летопись. Современный учебник истории — это тоже хроника. Каждую из его глав можно рассматривать как отдельную хронику. Одна из наиболее известных хроник — та, что составляет хронологическую канву Библии.

Каждую хронику можно разбить на фрагменты — более короткие хроники. Основная используемая группой А.Т. Фоменко идея в хронологии состоит в том, что некоторые фрагменты, привязанные в хронологии Скалигера к различным эпохам, на самом деле описывают одни и те же события. Их вслед за А.Т. Фоменко будем называть дубликатами. Коротко рассматриваемую идею можно сформулировать так: в хронологии есть дубликаты.

О существовании дубликатов в истории Древнего Рима (в классическом изложении, т.е. по Скалигеру) писал Н.А. Морозов. Он их указывал явно. Так что новизны в идее поиска дубликатов нет. Достижение группы А.Т. Фоменко состоит, во-первых, в том, что были предложены статистические (т.е. формально-математические) методы поиска дубликатов и, во-вторых, в том, что с помощью компьютеров был обсчитан весь массив имеющихся хроник.

Основная математико-статистическая идея группы А.Т. Фоменко состоит в формальном введении того или иного расстояния (меры похожести) в формальном же пространстве, описывающем возможные варианты фрагментов исторических хроник (использованные ими расстояния будут рассмотрены ниже).

Поясним эту математическую фразу. Сначала конструируется некоторое математическое пространство, в котором лежат математические образы фрагментов хроник. Это — отнюдь не элементарная операция, поскольку А.Т. Фоменко стремится учесть возможные варианты фрагментов, ошибки переписчиков, например, пропуск в хронике того или иного правителя, и т.п. В результате образ фрагмента — это не точка, а скорее облако точек, сконцентрированное в определенном месте пространства. Затем вводится расстояние (в математическом смысле) или показатель близости, похожести (или различия) между образами фрагментов. Поясним: предлагается способ расчета некоторого показателя $f(a,b)$ для любых двух образов фрагментов a и b, показывающего степень их похожести: если этот показатель мал, то фрагменты a и b похожи, если велик, то существенно отличаются. В нескольких методиках показатель $f(a,b)$ не является коммутативным: $f(a,b)$ не совпадает с $f(b,a)$. Чтобы не усложнять изложение, будем говорить о расстояниях (как известно, большинство авторов считает, что расстояние коммутативно: $f(a,b) = f(b,a)$ для любых a,b).

Следующий шаг — интенсивное применение компьютеров для сплошной обработки всего массива образов фрагментов хроник, сформированного на предыдущем шаге. Цель обработки массива — выделение пар фрагментов хроник, расстояние между которыми меньше некоторого порогового числа. Такие пары рассматриваются как дубликаты, повествующие об одних и тех же событиях. Итогом компьютерного анализа является составление списка дубликатов.

Пороговое значение определялось по «обучающей выборке» — набору фрагментов хроник, про которые точно известно, когда они говорят об одних и тех же событиях, а когда — о различных. Например, как пишут Г.В. Носовский и А.Т. Фоменко, обучающую выборку можно сформировать из фрагментов западноевропейских хроник о событиях после 1700 г. Численные эксперименты показали, что значения используемых расстояний для пар дубликатов на несколько порядков меньше значений для независимых фрагментов. Это позволяет надежно выделять дубликаты и в том массиве образов фрагментов хроник, который не входит в обучающую выборку.

Дополнительным подтверждением правильности выделения пар дубликатов служит то, что различные методики группы А.Т. Фоменко (использующие различные пространства, образы фрагментов хроник, расстояния) дают одни и те же результаты. С точки зрения общей теории устойчивости [3, 5] это говорит о том, что дубликаты — объективная реальность, они действительно присутствуют в массиве фрагментов хроник, не зависят от субъективизма исследователя. Вот если бы разные методы давали разные множества пар дубликатов, были бы все основания усомниться в их объективном существовании. Выбор метода — в распоряжении исследователя, какой метод выбрал — такой и результат получил.

Итак, дубликаты выделены. У фрагмента *a* может быть не один дубликат *b*, а еще и дубликаты *c, d*, ... Поскольку совершенно невероятно, чтобы история повторялась, все дубликаты *a, b, c, d, ...* соответствуют одним и тем же реальным событиям. Основная гипотеза А.Т. Фоменко состоит в том, что эти реальные события соответствуют последнему по времени дубликату. Остальные дубликаты получены сдвигами некоторых из реальных событий последнего по времени дубликата вглубь времен, другими словами, влево по оси времени. После выделения пар дубликатов фрагментов хроник проводится анализ связей между дубликатами и их группами с целью выделения «костяка», из которого путем дублирования получаются все остальные цепочки хроник.

Итак, реальные события конкретного фрагмента хроники раздваиваются, растраиваются и т.д., уходя в древность. При этом конкретная личность получает отражения, дубликаты в прошлых веках. Как справлялись с этим затруднением хронисты Скалигер и Пентавиус и многие иные? Блестяще справлялись, с выдумкой.

Старались не дублировать имена. И это удавалось. У каждого человека и тогда, и даже сейчас — много имен. Например, в детстве человека звали Шуриком, сейчас он для близких — Саша, официальное имя — Александр, коллеги и студенты величают Александром Ивановичем, в официальной обстановке — Профессор Такой-то. Пять имен у него. И это еще не все. Он может быть и Саней, и Аликом, и Старшим Лейтенантом Таким-то. А ведь бывают еще псевдонимы, прозвища. В древности на Руси было личное имя, и было церковное, даваемое при крещении. При вступлении на престол меняли имя. Так что у хронистов был большой выбор для именования лиц-дубликатов.

Как быть с событиями-дубликатами? Иногда одни события «уезжали» в далекое прошлое, другие оставались на своем месте. Но бывало и так, что одни и те же события описывались несколько раз. Тогда выручали подробности, для каждого дубликата свои. С совершенно неправдоподобной точностью цитируют слова великих греков и римлян по тому или иному поводу, как будто этих лиц постоянно сопровождали стенографистки или они носили с собой диктофоны. Впрочем, не следует подходить к работе старых хронистов со своими мерками. Нам важна точность в описании исторических событий. А их, возможно, больше интересовала занимательность и учет интересов заказчика (ведь рукопись надо выгодно продать).

Перед группой А.Т. Фоменко встала задача восстановления реальной истории, «сбора» дубликаты воедино. При решении этой задачи полезными оказались астрономические данные, лингвистические соображения, анализ предметов материальной культуры и другие методы, не связанные впрямую с компьютерным анализом фрагментов хроник.

Сформулируем два основных научных результата группы А.Т. Фоменко. Они получены путем статистического анализа исторических документов с помощью современных компьютерных технологий.

1. **Большинство исторических хроник**, известных нам в настоящее время**, дублируют друг друга**. В частности, хроники, описывающие так называемые «Древний Рим» и «Средневековье», говорят об одних и тех же событиях.

2. В исторических хрониках, известных нам в настоящее время, рассказывается о событиях, отстоящих от современности не более чем на 1000 лет.

Составлены списки хроник, рассказывающих об одних и тех же событиях. Выяснено, что хроники, описывающие историю «древних времен» и «средних веков», а также хроники китайской истории и истории различных европейских государств рассказывают не о разных, а об одних и тех же событиях. Выявлены так называемые «параллелизмы» в истории — то есть рассказы в разных хрониках об одних и тех же событиях, ошибочно считаемых сейчас за рассказы о разных событиях. Предпринята попытка новой датировки исторических событий и восстановления информации о подлинной истории человеческого общества на основе новых данных.

С точки зрения такой научно-практической дисциплины, как статистические методы, исторические хроники и образы их фрагментов — это частные случаи объектов нечисловой природы. Поэтому разработанные группой А.Т. Фоменко компьютерно-статистические методы (они описаны, например, в [19, с. 196–226; 20, с. 654–717; 26]) относятся к нечисловой статистике (статистике объектов нечисловой природы), как об этом и было заявлено в обзоре 1990 г. [27] по этой тематике.

Основные идеи нечисловой статистики связаны с введением «расстояния» или показателя близости (различия) в соответствующем пространстве объектов нечисловой природы и дальнейшем использовании этого показателя в различных алгоритмах. В работах группы А.Т. Фоменко применяется несколько (по крайней мере, 7) различных методик для введения расстояний. Что самое вдохновляющее — оказалось, что выводы, полученные по разным методикам, согласуются между собой. Как уже отмечалось, в соответствии с общей теорией устойчивости в социально-экономических моделях [3, 5] эта согласованность выводов свидетельствует о том, что полученные результаты отражают реальность, а не субъективный выбор модели и метода исследователем.

Математико-статистические методы, разработанные и использованные группой А.Т. Фоменко, многократно подробно обсуждались в 1980-х гг. на семинарах и конференциях по прикладной математической статистике. Среди них были и такие представительные, как Международные Вильнюсские конференции по теории вероятностей и математической статистике (1985, 1989).

Результаты реконструкции мировой истории на основании новых статистических методов представлены в монографиях [19–26]. Существует сайт, на котором в электронном виде представлены книги и другие публикации по «Новой хронологии».

Методы группы А.Т. Фоменко. Рассмотрим некоторые статистические методы анализа хроник, примененные группой А.Т. Фоменко. Поскольку реализовать их на практике можно лишь с помощью современной электронной техники, то их можно назвать компьютерно-статистическими.

Метод корреляции максимумов. Любую хронику, например, летопись, можно разбить на последовательные куски, соответствующие отдельным годам. Вполне естественно, что объемы текстов, посвященные тем или иным годам, различаются. В одни годы происходило много событий, «достойных» быть отмеченными в хронике — смена правителя, династический брак, рождение наследника, война, боярский заговор, эпидемия, солнечное затмение и т.п. В другие годы такие события могли вообще отсутствовать, и запись за год оказывалась короткой. В качестве объекта компьютерного анализа будем рассматривать объем текста хроники, приходящийся на тот или иной год.

Пусть T — натуральное число, обозначающее номер рассматриваемого года. Отсчет может идти от начала хроники или в какой-либо условной системе отсчета, например, «от сотворения мира», это не влияет на результат расчетов. С формальной точки зрения хронику можно описать функцией $V(T)$ — объемом записей за год T, измеренным, например, количеством использованных символов. Если строки или страницы примерно одинаковы по объему, то объем можно измерять в строках или в страницах. Особенно это удобно, если хроника опубликована.

Как те же события будут отражаться в других хрониках? Стоит обсудить, что это за «другие хроники». Это либо летописи, написанные примерно в то же время, но в других местах и странах. Другой вариант — сводки, составленные позже, вплоть до нынешних учебников. Не исключено, что и рассматриваемая хроника является подобной сводкой, тогда и для нее есть исходный текст или тексты. Как соотносятся функции объема записей $V(T)$ для различных хроник, повествующих об одном и том же? Конечно, они не будут совпадать. В сводке выпадут малозначащие (пол мнению составителя) события, объем записей $V(T)$ уменьшится. Основная гипотеза А.Т. Фоменко формулируется так: **от тех лет, которым первоначально было посвящено больше текстов, больше текстов и останется**.

Каждый местный летописец добавит что-то свое, касающееся его княжества, города или монастыря, и сократит что-то из чужого. Предполагаем, однако, что всплески (локальные максимумы) функции объема записей $V(T)$ останутся примерно на тех же местах. Надо отметить, что при большом внимании летописца к чисто местным проблемам, например, к сбору зерна на полях его монастыря и смене игуменов, целесообразно сначала отредактировать местную хронику, оставив в ней лишь информацию о событиях более широкого масштаба, отраженных и в других хрониках.

Надо ввести показатель похожести функций объема записей для двух хроник. Пусть наши данные имеют следующий вид. Рассматриваются две хроники за одно и то же количество лет, для определенности, за k лет. Пусть $V_1 = (V_1(T_1), V_1(T_2), ..., V_1(T_k))$ — вектор, составленный из объемов записей в первой хронике за эти годы, а $V_2 = (V_2(T_1), V_2(T_2), ... , V_2(T_k))$ — аналогичный вектор для второй хроники. Вычислим коэффициент корреляции между векторами V_1 и V_2. Очевидно, он будет близок к 1, если функции объемов записей для двух хроник похожи, и будет заметно меньше 1 в противном случае. Предыдущая фраза — это гипотеза, которая подлежит проверке.

Сначала надо уточнить два обстоятельства. Коэффициентов корреляции много. Наиболее известен обычный линейный коэффициент (Пирсона). Также часто рекомендуют использовать непараметрические коэффициенты корреляции Кендалла и Спирмена. Непараметрические коэффициенты нацелены на проверку одинаковой упорядоченности координат векторов. Одинаковая упорядоченность координат этих векторов означает, что если в первой хронике текста за один год содержится больше, чем за другой, то и во второй хронике больше говорится о первом годе, чем о втором. Это именно то, что мы предполагаем. Поэтому для проверки коррелированности функций объемов записей больше подходят непараметрические коэффициенты корреляции, чем обычный линейный коэффициент корреляции.

Второе обстоятельство связано с тем, что никакой вероятностной модели здесь нет. Мы отнюдь не утверждаем, что независимые хроники (относящиеся к различным временам и странам) независимы в статистическом смысле. Методы математической статистики, основанные на теории вероятностей, здесь применять нецелесообразно. Метод корреляции максимумов А.Т. Фоменко относится к анализу данных — разделу прикладной статистики, не опирающемуся на вероятностные модели.

Этот метод основан на обширном компьютерном эксперименте. Коэффициенты корреляции вычислялись для различных пар хроник (одинаковой длины) — заведомо зависимых и заведомо независимых (не только для хронологии Скалигера, но и с точки зрения группы А.Т. Фоменко), а также тех, о которых не было априорного мнения.

Результаты компьютерного вычислительного эксперимента замечательны. Произошло превосходное разделение. Все заведомо зависимые пары сосредоточились в окрестности 1. Все заведомо независимые пары легли значительно левее. Посередине — пусто. Никакого перемешивания!

Весьма интересны результаты для тех пар, о которых не было априорного мнения. Некоторое количество из них оказалось среди заведомо зависимых пар, весьма близко к 1. Это и есть искомые дубликаты. С точки зрения хронологии Скалигера эти хроники никак не связаны, они повествуют о событиях, которые произошли в разных странах и в разные годы. Но формальный анализ показывает их поразительную похожесть. Значит, в них разными словами говорится об одном и том же.

Как ясно из сказанного выше, сравнивались не хроники целиком, а куски хроник длиной k лет. Какое k брать? При малых k (до нескольких десятков) нельзя ожидать такой четкой картины разделения зависимых и независимых пар, которая описана только что. Поэтому целесообразно брать k возможно большим. Тогда и дубликаты получаются длинными, их легче интерпретировать. Однако чем больше k, тем меньше хроник такой длины можно найти. Расчеты группой А.Т. Фоменко проводились при разных k, но наиболее перспективным оказался интервал около k = 150 (лет). Надо отметить, что группа А.Т. Фоменко использовала несколько иной коэффициент близости функций объема текстов по годам, чем описанный выше. Их коэффициент определен, например, в [19, с. 199].

Несмотря на сравнительную простоту идеи, реализовать ее можно было лишь на основе интенсивного использования современных компьютеров. Ни И. Ньютон, ни Н.А. Морозов никогда не смогли бы провести нужного объема вычислений. Добавим, что компьютерно-статистические методы далеко не исчерпали своих возможностей при анализе нарративной (то есть повествовательной) исторической информации — хроник. Так, в рассматриваемом методе напрашивается проведение целой серии дополнительных расчетов, связанных с использованием того или иного коэффициента корреляции или похожести функций объема годовых записей. Не ясно, дадут ли эти расчеты новые результаты в компьютерно-статистической хронологии, но бесспорно, что они дадут новые аргументы в ее поддержку и позволят отработать новые инструменты анализа исторической информации.

Метод династий. Этот метод разработан еще Н.А. Морозовым. Его суть состоит в выделении «династии» — последовательности следующих друг за другом k правителей, k порядка 15, и рассмотрении k-мерного вектора длительностей их правления. Если для двух династий эти вектора совпадают — имеем идеальный дубликат.

Ситуация осложняется тем, что организация властных структур всегда страдала некоторой хаотичностью. Часто власть делилась между соправителями. Кроме того, следует различать формальное царствование и реальное. Царем могли провозгласить младенца, а до его совершеннолетия в таком случае действовал реальный правитель (регент), не имеющий династических прав и не претендующий на формальное признание. Кроме того, коронация, как правило, отстоит от реального взятия власти.

Свою лепту в искажение исторических текстов мог внести и летописец. Он мог перепутать следующих друг за другом правителей, особенно при их кратком пребывании у власти. Он мог сдвинуть годы правления, учитывая или не учитывая период борьбы за власть (между реальными соправителями), время регентства, отсчитывая правление от взятия власти или от коронации, в паре «отец-сын» беря правление сына от реальной передачи власти или от смерти отца, и т.д.

В результате одна реальная династия могла отразиться в хрониках разными способами. Вместо одного вектора в k-мерном пространстве получаем «облачко» таких векторов. В этом «облачке — все возможные описания данной династии, как реально существующие в исторических документах, так и те, что могли бы возникнуть. Если за конкретный вектор как описание династии можно критиковать исследователя — дескать, в такой-то хронике она описана по-другому, то в «облачке» это другое описание, скорее всего, уже содержится. Именно «облачко» — адекватное описание реальной династии в математических терминах.

Очевидно, если «облачка», построенные по считающимся в хронологии Скалигера независимым данным, пересекаются, то династии, скорее всего, являются дубликатами. Если «облачка» далеки друг от друга, они не могут быть дубликатами.

Дальнейшее статистику очевидно — надо ввести расстояние между «облачками», уточнить термины предыдущего абзаца «пересекаются» и «далеко», а затем провести компьютерный эксперимент, вычисляя расстояния между парами династий.

С точки зрения статистика задача напоминает кластер-анализ. Требуется определить расстояние между кластерами — «облачками» в многомерном пространстве. Есть целый ряд методов. Согласно методу ближнего соседа, в качестве расстояния берется минимальное из расстояний между двумя точками, одна из которых лежит в первом кластере, а другая — во втором. Согласно этому методу, расстояние от династии до нее самой равно 0. В методе дальнего соседа, наоборот, в качестве расстояния берется максимальное из расстояний между двумя точками, одна из которых лежит в первом кластере, а другая — во втором. Согласно этому методу, расстояние от династии до нее самой равно диаметру «облачка». В методе средней связи в качестве расстояния берется среднее из расстояний между двумя точками, одна из которых лежит в первом кластере, а другая — во втором. Поскольку средних величин бесконечно много, то можно рассматривать разные алгоритмы вычисления расстояния, основанные на среднем арифметическом, среднем геометрическом, медиане и т.д. Сам А.Т. Фоменко предложил и использовал несколько иной метод. Так что существует большое поле для экспериментов с «облачками»-династиями.

Каков же результат компьютерных экспериментов А.Т. Фоменко? Как и следовало ожидать, он напоминает результат применения метода корреляции максимумов. Явно зависимые династии оказались близкими, явно независимые династии — далекими. Но обнаружились пары династий-дубликатов. Эти династии в хронологии Скалигера считались никак не связанными, далекими друг от друга в пространстве и во времени, но они оказались весьма близкими в терминах того расстояния, что использовалось в эксперименте. При этом все расстояния между «облачками» разделились на две группы, по величине различающиеся на несколько порядков: близкие к 0 расстояния между зависимыми династиями и дубликатами и достаточно большие расстояния между независимыми династиями. В промежутке — пустота.

Метод затухания частот. Этот метод подробно описан в книге [20, с. 654–717]. Согласно нему «при правильном хронологическом порядке фрагментов частота употребления имен персонажей данного поколения должна в среднем уменьшаться, «затухать» при переходе к описанию все более отдаленных от него во времени персонажей». Таким образом, метод датировки опирается на использование имен исторических персонажей.

Из сформулированного принципа вытекает, что целесообразно ввести понятие «глава-поколение» — фрагмент хроники, описывающий события приблизительно одного поколения. Рассмотрим группу имен, впервые появившихся в рассматриваемой «главе-поколении». Обязательное условие — раньше они никогда не появлялись, но встречаются в «главе-поколении», соответствующей моменту $T(0)$. Очевидно, некоторые из них будут встречаться и дальше, летописцы будут ссылаться на царей и героев предыдущих эпох, но все реже и реже. Если построить график частоты их встречаемости как функции времени, то левее точки $T(0)$ частота равна 0 (по определению), максимума она достигает в момент $T(0)$ (в те два-три десятилетия, что соответствуют интересующему нас поколению), а потом монотонно убывает.

Если же монотонность нарушается, «старая» группа имен дает заметный всплеск, то есть мы имеем дело с «возрождением» — в данном случае «возрождением» множества имен исторических лиц, то, очевидно, нам попался дубликат, повторное описание прежних событий. Метод, основанный на затухании частот, позволяет также датировать фрагмент текста, оптимизируя на компьютере возможное место для данного фрагмента так, чтобы функции встречаемости как можно меньше отличались от идеальных.

При более тщательном анализе используется понятие «возраст имени» — число глав-поколений с момента первого появления такого же имени в тексте. Оказалось, что имена можно разделить на два класса — «вечные» (около 1/4 имен) и «обычные». В соответствии с принципом затухания частот возраст «обычных» имен сохраняет приблизительно постоянное значение (со случайным разбросом вокруг него). Возраст же «вечных» имен постоянно растет со временем (по линейному закону). Следовательно, средний возраст имен стоит описывать линейной регрессией. Дубликат будет выражаться резким поднятием части графика на одну и ту же величину. Как выражается А.Т. Фоменко, возникает характерная «полка», высота которой соответствует расстоянию между дубликатами. Обнаружить такие «полки» можно чисто формально, как отклонения графика от линейной зависимости. Для этого используют алгоритмы обнаружения разладки, те же, что и для обнаружения разладки технологического процесса в рамках статистических методов управления качеством продукции.

Рассматриваемый сюжет широко развернут в публикациях группы А.Т. Фоменко. В частности, разработаны методы анализа матриц связей для хронологических списков имен.

Метод анкет-кодов. Анкет-кодом А.Т. Фоменко называет формализованную биографию упомянутого в хронике деятеля. Точнее говоря, речь идет об «исторической литературной биографии», которая может не совпадать с реальной биографией (о которой мы ничего не знаем). Анкет-код состоит из 34 пунктов, ответ на каждый из них, как правило, формализован и состоит в выборе одного из нескольких вариантов ответов.

Приведем начало анкет-кода.

1. Пол.

2. Длительность жизни.

3. Длительность правления. Конец правления практически всегда однозначно зафиксирован, начало правления допускает иногда несколько вариантов, которые отмечаются как равноправные.

4. Социальное положение и занимаемый пост (код *а* — царь, император, король, великий князь, то есть самостоятельный правитель; код *б* — полководец; политик, общественный деятель; код *в* — ученый, писатель, актер, философ и т.д.; код *г* — религиозный вождь, патриарх, папа, епископ и т.д.).

5. Смерть правителя (*а* — естественная (то есть не насильственная) смерть в мирной обстановке; *б* — убит на поле боя противниками или смертельно ранен; *в* — убит в результате заговора, вне войны; *г* — убит в результате заговора во время войны; *д* — специальные, экзотические обстоятельства смерти).

6. Стихийные бедствия во время правления (*а* — голод; *б* — наводнения; *в* — повальные болезни; г — землетрясения; *д* — извержения вулканов; при этом отмечаются также длительность бедствий и год или годы правления, когда они имели место).

7. Астрономические явления во время правления (*а* — есть; *б* — нет; если есть, то какие именно: *в* — затмения, *г* — кометы; *д* — «вспышки звезд»).

8. Войны во время правления (*а* — есть; *б* — нет).

9. Число войн.

И еще 25 пунктов анкеты, которая напоминает анкеты социолога при массовом опросе, маркетолога при изучении предпочтений потребителя, формализованную историю болезни, используемую в научных медицинских исследованиях, и т.п.

Не надо думать, что этих данных слишком мало для идентификации личности. Напомним, что в десятимиллионной Москве для идентификации жителя (с указанием его домашнего адреса и телефона) вполне достаточно всего лишь фамилии, имени, отчества (хотя многие из них широко распространены), года и места рождения.

Анкет-коды правителей при объединении дают анкет-код династии. Для нахождения дубликатов А.Т. Фоменко разработал методику, основанная на принципе «малых искажений». В соответствии с ней, если анкет-коды двух династий «мало отличаются» друг от друга, то они изображают одну и ту же реальную династию. Если же два анкет-кода изображают разные династии, то анкет-коды «далеки» друг от друга.

С точки зрения специалиста по статистике объектов нечисловой природы, анкет-код — это вектор с разнотипными признаками. В пространствах подобных векторов можно разными способами ввести расстояния или показатели различия, позволяющие формализовать понятия «мало отличаются» и «далеки».

Восстановление подлинной истории. Мы разобрали ряд компьютерных методов поисков дубликатов, использованных группой А.Т. Фоменко. Самое интересное начинается тогда, когда дубликаты уже выявлены. Процесс их сопоставления вряд ли можно формализовать, но он очень интересен. Обычно обнаруживаются сходные события и персонажи, «уехавшие» из основного периода в дубликаты. Поколения историков постарались наполнить колоритными подробностями сухие строки хроник, и теперь приходится очищать события от этой шелухи. В работах группы А.Т. Фоменко в ход идут самые разные соображения. И астрономические, использованные еще И. Ньютоном и Н.А. Морозовым. И лингвистические. И связанные с анализом остатков древней материальной культуры (например, старинного оружия). Весьма полезными оказываются древние географические карты и предметы искусства прежних веков (прежде всего картины).

В результате удается во многом восстановить истинную картину исторических событий. Этому и посвящены книги группы А.Т. Фоменко 1990-х гг. О компьютерных методах обнаружения дубликатов, интересных в основном специалистам, там говорится сравнительно мало, но вполне достаточно для понимания сути дела. Первый этап восстановления истинной хронологии — выявление дубликатов — по существу закончен. Чего нельзя сказать о втором — выявлении истинной всемирно-исторической картины.

Сложности касаются последней тысячи лет. Про предыдущее все ясно — у нас нет никаких письменных сведений о событиях ранее IX в. н.э. А вот потом письменные свидетельства сочетаются с наличием хронологических сдвигов. Какие-то события сдвинуты в хрониках на 300 лет вниз, какие-то, возможно, наоборот, приближены к нам.

При анализе исторических хроник приходится сталкиваться с различными сложностями. В качестве примера кратко обсудим легенду о «призвании варягов». Сколько написано об этом! Одни выясняли, откуда варяги родом — из Швеции или из Пруссии, другие утверждали, что сама мысль о призвании иноземцев унижает Россию. И до сих пор идут дискуссии. И только А.Т. Фоменко посмотрел на исходную рукопись. Выяснилось, что лист про призвание варягов **вклеен**. Подделка. Как, видимо, и вся Радзивиловская рукопись. Интересно также, что все это было, вообще говоря, известно историкам, но они молчали. Был разработан даже специальный «Эзопов» язык. Например, что значит фраза: «список второй половины XVII в. летописи XII в.»? Значит, текст написан во второй половине XVII в. Что в нем от летописи XII в. — неизвестно. Зная про события тех времен, например, про «исправление богослужебных книг», можно с полным основанием предположить, что и древняя летопись была исправлена. Нельзя думать, что переписчики копировали «буква в букву». Даже сейчас редакторы газет, журналов и книг считают возможным делать с текстом практически беззащитного автора все, что им угодно. Например, публиковать только вставной сюжет, отбросив основное содержание. Так, в книгах известного популяризатора науки Я.И. Перельмана часто встречаешь сюжеты, которые он физически не мог написать, поскольку в его время (он умер в блокадном Ленинграде) не было соответствующих реалий, например, он не мог знать названий и характеристик советских космических кораблей. Кто-то написал и вставил соответствующие разделы. Кто — неизвестно. Если так поступают сейчас, то можно представить себе, что творилось несколько веков назад...

Компьютерно-статистическая хронология. Теория группы академика А.Т. Фоменко — новая компьютерно-статистическая хронология мировой истории — является подлинным триумфом современных статистических методов и информационных технологий. С помощью компьютеров, позволяющих за несколько часов провести такое количество вычислений, на которое ранее ушли бы годы труда десятков, а то и сотен людей, появилась возможность разработать объективные, строго научные методы исследования столь, казалось бы, далеких от точных наук данных, как исторические хроники. Появилась возможность выявлять не только ошибочно датированные, но и сфальсифицированные исторические тексты, восстанавливать их истинную датировку, выявлять, какую эпоху они описывают. В целом можно сказать, что новые методики исследования текстовой информации, разработанные группой академика А.Т. Фоменко, позволяют проводить анализ текстов вне зависимости от субъективного восприятия их исследователем, то есть использовать строго научный подход в этой области.

Новая хронология позволяет понять многое в борьбе идей в современном научном и массовом сознании. Во-первых, становится ясной глубинная причина настороженного отношения Запада к России. Это — устойчивая реакция на его многовековое вассальное положение по отношению к Руси. Не сумев победить в открытой борьбе, он добился реванша «дипломатическими» методами, посеяв смуту в высшем руководстве и выдвинув на первое место своих ставленников.

Во-вторых, искусственное удлинение западной истории создает чувство неполноценности у граждан «молодой» России. На это же направлено и искусственное введение веков «татаро-монгольского ига» в нашу историю. Цель — заставить нас смотреть снизу вверх на Запад с его якобы всегда более высокой культурой, продвинутой наукой, богатой историей, эффективной промышленностью и т.д.

Века прозападной пропаганды дали результат. Российские потребители с гораздо большим энтузиазмом покупают западные товары, чем отечественные. Распространенный эксперимент — поменять упаковки, скажем, российских и французских духов. По мнению покупателей, товар во французской упаковке всегда лучше, независимо от того, что в ней содержится.

Значение результатов группы А.Т. Фоменко не ограничивается историей. При преподавании экономических дисциплин совершенно необходимо использовать современные достижения исторической науки — новую компьютерно-статистическую хронологию [19–26], построенную на основе интенсивного применения статистических методов. Она проливает новый свет не только на экономическую историю, но и на борьбу идей в современном научном и массовом сознании. Становятся яснее причины отличия России от стран Запада и необходимость разработки экономической теории, адекватной российской социальной психологии и экономической практике.

Работы по новой хронологии лишний раз доказывают, что с развитием статистических методов и информационных технологий перед человечеством открылись новые, грандиозные возможности.

8.5. Вероятностно-статистическое моделирование помех, создаваемых электровозами

Движение электровозов создает помехи, влияющие на проводные линии связи. Создание достаточно эффективных и в то же время экономичных средств защиты проводных линий связи от мешающих влияний, создаваемых тяговыми сетями переменного тока, предполагает в качестве подготовительного этапа разработку математических моделей помех, создаваемых электровозами.

Как показано в [28, с.59], m-ю гармонику ($m \geq 7$) создаваемой электровозом помехи можно описать двумерным случайным вектором:

$$(\xi_1, \xi_2) = (r\cos\varphi,\ r\sin\varphi)\,, \tag{1}$$

где r — амплитуда и φ — фаза помехи, причем r и φ независимы (как случайные величины), распределение φ равномерно на $[0;2\pi]$. Как нетрудно подсчитать,

$$M(\xi_1) = M(\xi_2) = M(\xi_1\xi_2) = M(\xi_1^3) = M(\xi_2^3) = 0,\ D(\xi_1) = D(\xi_2) = \frac{1}{2}M(r^2)\,. \tag{2}$$

В случае n электровозов рассматриваемая гармоника суммарной помехи описывается вектором (α_n, β_n), причем из физических соображений:

$$(\alpha_n, \beta_n) = \sum_{i=1}^{n} r_i(\cos\varphi_i, \sin\varphi_i)\,. \tag{3}$$

Каждый вектор в правой части (3) соответствует электровозу с тем же номером. Будем считать, что случайные величины $\{r_i, \varphi_i, i = 1,2,\ldots,n\}$ независимы в совокупности, а все фазы φ_i распределены равномерно на $[0;2\pi]$. Из (2) и (3) вытекает, что

$$M(\alpha_n^2 + \beta_n^2) = \sum_{i=1}^{n} M(r_i^2)\,, \tag{4}$$

если все слагаемые в правой части (4) существуют. Последнее будем считать выполненным. Правую часть (4) обозначим R_n^2.

Из многомерной центральной предельной теоремы (см. раздел 14.2) вытекает следующее утверждение.

Лемма 1. Пусть выполнены условия теоремы Линдеберга — Феллера, в данном случае:

$$\lim_{n\to\infty} R_n = +\infty,\ \lim_{n\to\infty}\left(\frac{1}{R_n^2}\max_{1\leq i\leq n} M(r_i^2)\right) = 0\,. \tag{5}$$

Тогда для любых чисел *x*, *y*:

$$\lim_{n\to\infty} P\left\{\frac{\sqrt{2}}{R_n}\alpha_n < x;\ \frac{\sqrt{2}}{R_n}\beta_n < y\right\} = \Phi(x)\Phi(y)\,, \tag{6}$$

где $\Phi(z)$ — функция нормального распределения с математическим ожиданием 0 и дисперсией 1.

Нам понадобится следующее утверждение.

Лемма 2. Пусть ξ и η — независимые случайные величины, имеющие стандартное нормальное распределение $\Phi(z)$. Пусть выполнены условия леммы 1. Пусть функция $f : R^2 \to R^1$ интегрируема по Риману по любому квадрату. Тогда:

$$\lim_{n\to\infty} P\left\{f\left(\frac{\sqrt{2}}{R_n}\alpha_n,\ \frac{\sqrt{2}}{R_n}\beta_n\right) < u\right\} = P\{f(\xi,\eta) < u\} \tag{7}$$

при всех *u*, при которых правая часть (7) непрерывна по *u*.

Доказательство вытекает из леммы 1 и результатов о наследовании сходимости (см. раздел 14.3).

В методиках расчетов, связанных с защитой проводных линий связи, используется математическое ожидание амплитуды суммарной помехи (α_n, β_n). Следовательно, надо рассмотреть функцию:

$$f(z,w) = \sqrt{z^2 + w^2}\,.$$

Из леммы 2 следует, что распределение нормированной амплитуды суммарной помехи:

$$\gamma_n = \frac{\sqrt{2}}{R_n}\sqrt{\alpha_n^2 + \beta_n^2}$$

сходится к распределению случайной величины:

$$\delta = \sqrt{\xi^2 + \eta^2}\,.$$

Случайная величина δ, как известно (см., например, [29, с. 262]), имеет плотность:

$$h(x) = x\exp\left\{-\frac{x^2}{2}\right\}, \ x \geq 0. \tag{8}$$

Следовательно,

$$M(\delta) = \int_0^{+\infty} x^2 \exp\left\{-\frac{x^2}{2}\right\}dx = \sqrt{\frac{\pi}{2}} \approx 1{,}26.$$

При проектировании защиты проводных линий связи нормативные ограничения накладываются на математическое ожидание суммарной помехи. Поэтому для практического применения полезны следующие результаты.

Теорема 1. Пусть предельные соотношения (5), (6) имеют место и, кроме того,

$$\overline{\lim_{n\to\infty}} \frac{1}{R_n^4}\sum_{i=1}^{n} M(r_i^4) < +\infty. \tag{9}$$

Тогда

$$\lim_{n\to\infty} M(\gamma_n) = M(\delta).$$

Следствие. Для математического ожидания $M[|(\alpha_n, \beta_n)|]$ амплитуды суммарной помехи имеет место предельное соотношение:

$$\lim_{n\to\infty} \frac{M[|(\alpha_n, \beta_n)|]}{\sqrt{\frac{\pi}{4}}R_n} = 1.$$

Отметим, что соотношения (5) и (9) выполнены, например, в случае, когда существуют числа $0 < r < R$ такие, что $r < r_{in} < R$ при всех n, $i = 1, 2, \ldots, n$, т.е. амплитуды электровозов ограничены сверху и снизу. Это условие является весьма естественным с прикладной точки зрения.

Доказательство теоремы 1. При любом $a > 0$ справедливо равенство:

$$M(\gamma_n) - M(\delta) = \left\{\int_{|x|<a} xdP(\gamma_n < x) - \int_{|x|<a} xh(x)dx\right\} + \left\{\int_{|x|\geq a} xdP(\gamma_n < x) - \int_{|x|\geq a} xh(x)dx\right\}. \tag{10}$$

Каждую из скобок в представлении (10) будем обрабатывать своим способом. Поскольку

$$\int_{-a}^{a} x dF(x) = xF(x)\Big|_{-a}^{a} - \int_{-a}^{a} F(x)dx\,, \qquad (11)$$

то, подставляя в (11) вместо $F(x)$ соответственно $P(\gamma_n < x)$ и $P(\delta < x)$, заключаем с помощью леммы 2, что при фиксированном значении a первая скобка в (10) стремится к 0 при росте n.

Вторую скобку оценим следующим образом:

$$\left| \int_{|x|\geq a} x dP(\gamma_n < x) - \int_{|x|\geq a} x h(x)dx \right| \leq \int_{|x|\geq a} |\, x \,|\, dP(\gamma_n < x) + \int_{|x|\geq a} |\, x \,|\, h(x)dx\,. \qquad (12)$$

Из явного вида функции $h(x)$ (см. (8)) следует, что второе слагаемое в (11) стремится к 0 при росте n. Для оценки первого слагаемого в (11) воспользуемся соотношением:

$$\int_{a}^{+\infty} x dF(x) = \int_{a}^{+\infty} x d(F(x) - 1) = x(F(x) - 1)\Big|_{a}^{+\infty} - \int_{a}^{+\infty} (F(x) - 1)dx\,,$$

в котором положим $F(x) = P(\gamma_n < x)$. Оценим сначала порядок убывания по x величины:

$$1 - P(\gamma_n < x) = 1 - P\left(\frac{2}{R_n^2}(\alpha_n^2 + \beta_n^2) < x^2\right).$$

С помощью несложных, но достаточно продолжительных вычислений можно найти $M[(\alpha_n^2 + \beta_n^2)^2]$ и на основе неравенства (9) доказать, что существует константа C такая, что при всех n:

$$M\left[\left(\frac{2}{R_n^2}(\alpha_n^2 + \beta_n^2)\right)^2\right] < C\,.$$

Воспользовавшись неравенством Чебышева:

$$P\{|\,\xi\,| \geq b\} \leq \frac{M(\xi^2)}{b^2}$$

для любой случайной величины ξ и любого положительного числа b, заключаем, что:

$$P\left\{\frac{2}{R_n^2}(\alpha_n^2+\beta_n^2)\geq x^2\right\}=1-P(\gamma_n<x)\leq\frac{C}{x^4}.$$

Следовательно,

$$0\leq\int_a^{+\infty}xdP(\gamma_n<x)<\int_a^{+\infty}(1-P(\gamma_n<x))dx\leq\int_a^{+\infty}\frac{C}{x^4}dx=\frac{C}{3a^3}. \qquad (13)$$

Из соотношений (11), (8) и (12) следует, что первое слагаемое в правой части равенства (10) при фиксированном a стремится к 0 при росте n, в то время как вторая скобка стремится к 0 при росте a равномерно по n. Для завершения доказательства теоремы 1 по фиксированному $\varepsilon > 0$ найдем число a такое, что при всех n вторая скобка меньше $\varepsilon/2$, а затем по этому a найдем n_0 такое, что при $n > n_0$ первая скобка также меньше $\varepsilon/2$. Тогда $|M(\gamma_n)-M(\delta)|<\varepsilon$ при $n > n_0$. Теорема 1 доказана.

Рассмотрим важный частный случай. Пусть амплитуды всех электровозов детерминированы, одинаковы и равны 1, т.е. $r_i = 1$, $i = 1, 2, \ldots, n$. Тогда по следствию из теоремы 1 математическое ожидание амплитуды суммарной помехи приблизительно равно $0{,}886\sqrt{n}$, в то время как формула (44) на с. 60 электротехнической работы [28] дает $\sqrt{n}$. Таким образом, согласно развитой выше теории значение амплитуды суммарной помехи на 11,4 % меньше, чем согласно мнению специалистов в этой области.

Чем же объясняется различие в 11,4 % Дело в том, что «закон сложения высших гармоник в тяговой сети при двух электровозах в зоне питания» (формула (42) на с. 60 работы [28]) является приближенным и выполняется с точностью порядка 1 % (там же, с. 59). Можно было бы ожидать, что при рассмотрении n электровозов, а не двух, ошибка будет порядка $0{,}01n$, однако в силу того, что рассматриваемый «закон» выполняется тем точнее, чем больше различаются амплитуды слагаемых, ошибка имеет порядок лишь $0{,}114\sqrt{n}$, что при больших n существенно меньше, чем $0{,}01n$.

Итак, формула (44) в работе [28] завышает среднюю амплитуду суммарной помехи в $1/0{,}886\approx 1{,}13$ раза, что приводит к соответствующему увеличению средств, выделяемых на защиту проводных линий связи. Следовательно, внедрение развитой в настоящем разделе теории позволяет сэкономить достаточно большие материальные и финансовые ресурсы.

Однако приведенные выше результаты являются предельными, ими можно пользоваться при «достаточно большом» числе электровозов n. Для практики же наиболее интересны значения n в пределах от 2 до 10. Возникает вопрос о скорости сходимости в леммах 1–2 и особенно в теореме 1.

Пусть сначала $n = 2$. Нетрудно показать, что в случае детерминированных амплитуд r_1 и r_2:

$$M[|(\alpha_n,\beta_n)|] = R_2 M\left(\sqrt{1+\gamma\cos\varphi}\right), \qquad (14)$$

где $R_2 = \sqrt{r_1^2 + r_2^2}$, фаза φ — равномерно распределенная на $[0;2\pi]$ случайная величина, а

$$\gamma = \frac{2r_1r_2}{r_1^2 + r_2^2}.$$

Легко видеть, что при $\gamma = 1$ (т.е. при $r_1 = r_2$) математическое ожидание в правой части (13) рано 0,897. При $\gamma < 1$ соответствующий определенный интеграл приводится к эллиптическому интегралу и в явном виде не берется. Приведем разложение в ряд по степеням параметра γ:

$$M\left(\sqrt{1+\gamma\cos\varphi}\right) = \sum_{k\ge 0}\binom{0,5}{2k}2^{-2k}\binom{2k}{k}\gamma^{2k} = 1 - \frac{1}{16}\gamma^2 - \frac{15}{1024}\gamma^4 - \frac{105}{16384}\gamma^6 - \dots$$

Вычисленное при $n = 2$ значение 0,897 весьма близко к предельному (при $n \to \infty$) значению 0,886. Поэтому можно ожидать быстрой сходимости, по крайней мере при малоразличающихся значениях амплитуд.

Обсудим применение теоретических методов изучения скорости сходимости. Асимптотическое разложение характеристической функции случайного вектора (α_n,β_n) начинается с члена порядка n^{-1} (в силу того, что третьи моменты координат вектора равны 0 согласно (2)). Путем обращения этого разложения можно найти асимптотическое разложение для плотности (аналог разложения Эджворта), которое также начинается с члена порядка n^{-1}. Последний шаг — подобно доказательству теоремы 1 может быть сделан переход к интересующий нас разности:

$$M(\gamma_n) - M(\delta) = O\left(\frac{1}{n}\right). \qquad (15)$$

Обосновать описанную цепочку утверждений, приводящую к (14), и получить численные оценки скорости сходимости можно, оценивая остаточные члены во всех описанных выше преобразованиях. К сожалению, не удалось обнаружить публикаций, в которых имелись бы нужные оценки. Это объясняется тем, что рассматриваемая ситуация является весьма частной с точки зрения общей теории. Мы не стали строить микротеорию для обоснования правдоподобных рассуждений, приведших к (14), поскольку нас интересует точность приближения предельным распределением при $2 \le n \le 10$, а, как показывает опыт анализа распределения двухвыборочной односторонней статистики Н.В. Смирнова (разд. 14.6, подраздел «Точные формулы и асимптотика»), константы, которые можно таким путем получить для уточнения (14), сильно завышены. Поэтому обратились к численным методам, а именно, к вычислительному эксперименту методом статистических испытаний (Монте-Карло), т.е. с использованием датчика псевдослучайных чисел. В соответствии с рекомендациями [30] использовался алгоритм перемешивания Макларена — Марсальи (М-алгоритм).

В табл. 1 приведены значения оценок величины

$$\rho = \frac{M[|(\alpha_n, \beta_n)|]}{\sqrt{\sum_{1 \le i \le n} r_i^2}} - 0{,}886 \tag{16}$$

для пяти наборов амплитуд [31].

Таблица 1

Оценка скорости сходимости в теореме 1 методом Монте-Карло

№ эксперимента	Число электровозов n				
	1	2	3	4	5
1	1 –	1 0,009	1 0,018	1 0,010	1 0,013
2	0,40 –	0,41 0,016	0,72 0,029	0,64 0,021	1,18 0,022
3	0,40 –	0,40 0,010	0,41 0,022	0,86 0,030	0,80 0,016
4	0,40 –	0,53 0,031	2,52 0,095	0,72 0,077	0,85 0,067
5	0,40 –	0,41 0,014	0,85 0,041	0,56 0,017	0,82 0,017

№ эксперимента	Число электровозов n				
	1	**2**	**3**	**4**	**5**
1	1 0,013	1 0,001	1 0,001	1 0,010	1 0,010
2	0,66 0,022	1,60 0,025	1,70 0,018	0,87 0,008	2,60 0,007
3	0,86 0,012	0,85 0,009	0,90 0,007	2,97 0,040	0,72 0,035
4	3,00 0,025	0,65 0,029	1,46 0,017	0,63 0,023	0,84 0,022
5	0,58 0,007	1,63 0,025	1,64 0,011	0,87 0,009	1,00 0,012

Примечание. Для каждого из пяти экспериментов верхняя строка — значения амплитуд, нижняя — значения ρ из (15), оцененные по 10 000 испытаниям. В каждом наборе из 10 амплитуд сначала моделируется ситуация с первыми двумя амплитудами, потом — с первыми тремя, затем — с первыми четырьмя и т.д.

Математические ожидания в (15) оценивались как средние арифметические 10 000 наблюдений случайной величины $|(\alpha_n, \beta_n)|$. Дисперсия результата одного испытания согласно лемме 2, аналогу теоремы 1 для вторых моментов и соотношению (8) оценивается как $D(\delta)/2 \approx 0{,}22$, а потому среднеквадратическое отклонение табличных значений оценивается как 0,0015 (в предположении идеальности датчика псевдослучайных чисел).

Данные табл. 1 показывают, что наиболее быстрая сходимость в теореме 1 имеет место в случае равенства амплитуд. Напротив, если значение одной из амплитуд безгранично возрастает, а значения остальных фиксированы, то $\rho \to 0{,}114$. С целью изучения сходимости в случае различных амплитуд были взяты четыре независимые выборки значений амплитуд из распределения, заданного гистограммой для 9-й гармоники на рис. 17 в монографии Р.Н. Карякина [28, с. 40]. Объемы выборок — 10 наблюдений. Выборки приведены в табл. 1. В каждом случае рассчитаны оценки ρ для первых m элементов выборки, m = 2, 3, …, 10. Таким образом, табличные значения одной строки зависимы между собой, но строки независимы. Программирование и расчеты осуществлены С.Ю. Адамовым.

Из табл. 1 видно, что существенные различия между амплитудами приводят к возрастанию ρ, однако при росте числа электровозов n это возрастание становится все менее существенным. Так, для $n \geq 6$ максимальное табличное значение равно 0,040, т.е. при $n \geq 6$ имеем основания полагать, что:

$$M[|(\alpha_n, \beta_n)|] \leq 0{,}926 \sqrt{\sum_{1 \leq i \leq n} r_i^2} \tag{17}$$

для подавляющего большинства реальных ситуаций. Замена 1 на 0,926 в методиках защиты проводных линий связи сулит значительный экономический эффект. Кроме (16), данные табл. 1 позволяют дать ряд других полезных практических рекомендаций.

Замечание. Предположение о равномерности распределения фазы φ использовалось выше лишь при выводе соотношений (2) и (13). Поэтому все результаты настоящего раздела, кроме (13) справедливы для любых распределений фаз, удовлетворяющих соотношениям (2). В частности, для различных электровозов распределения могут быть различными.

Литература

1. *Неуймин, Я.Г.* Модели в науке и технике. История, теория, практика / Я.Г. Неуймин. — Ленинград : Наука, 1984. — 190 с.

2. *Моисеев, Н.Н.* Математические задачи системного анализа / Н.Н. Моисеев. — Москва : Наука, 1981. — 488 с.

3. *Орлов, А.И.* Эконометрика / А.И. Орлов. — Москва : Экзамен, 2002. — 576 с.

4. *Нейлор, Т.* Машинные имитационные эксперименты с моделями экономических систем / Т. Нейлор. — Москва : Мир, 1975. — 500 с.

5. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

6. *Гнеденко, Б.В.* Математика и контроль качества продукции / Б.В. Гнеденко. — Москва : Знание, 1978. — 64 с.

7. *Лэйард, Р.* Макроэкономика / Р. Лэйард. — Москва : Джон Уайли энд Санз, 1994. — 160 с.

8. *Нейман, Дж. фон.* Теория игр и экономическое поведение / Дж. фон Нейман, О. Моргенштейн. — Москва : Наука, 1970.

9. *Лотов, А.В.* Введение в экономико-математическое моделирование / А.В. Лотов. — Москва : Наука, 1984.

10. *Ланкастер, К.* Математическая экономика / К. Ланкастер. — Москва : Советское радио, 1972.

11. Население России 2000. Восьмой ежегодный демографический доклад / под редакцией А.Г. Вишневского. — Москва : Университет, 2000. — 176 с.

12. Экология / под редакцией С.А. Боголюбова. — Москва : Знание, 1999. — 288 с.

13. *Гундаров, И.А.* Пробуждение: пути преодоления демографической катастрофы в России / И.А. Гундаров. — Москва : Беловодье, 2001. — 352 с.

14. Предположительная численность населения Российской Федерации до 2016 года : (статистический бюллетень). — Москва : Госкомитет России по статистике, 2000. — 149 с.

15. Менеджмент / под редакцией Ж.В. Прокофьевой. — Москва : Знание, 2000. — 288 с.

16. *Орлов, А.И.* Математические модели в экономике. Модель Вильсона управления запасами / А.И. Орлов, Т.А. Конюхова. — Москва : Изд-во Московского государственного института электроники и математики (технического ун-та), 1994. — 31 с.

17. *Душкесас, Р.Ф.* Проблемы устойчивости в классической модели управления запасами : дипломная работа / Р.Ф. Душкесас ; МИНХ им. Г.В. Плеханова. — Москва : МИНХ им. Г.В. Плеханова, факультет экономической кибернетики, 1977. — 70 с.

18. *Орлов, А.И.* Некоторые модели планирования оптимальных размеров поставок и начального запаса / А.И. Орлов, Э.Э. Пейсахович // Экономика и математические методы. — 1975. — Т. XI. — №. 4. — С. 681–694.

19. *Носовский, Г.В.* Введение в новую хронологию (Какой сейчас век?) / Г.В. Носовский, А.Т. Фоменко. — Москва : КРАФТ+ЛЕАН, 1999. — 768 с.

20. *Носовский, Г.В.* Империя. Русь, Турция, Китай, Европа, Египет. Новая математическая хронология древности / Г.В. Носовский, А.Т. Фоменко. — Москва : Факториал, 1996. — 752 с.

21. *Носовский, Г.В.* Новая хронология и концепция древней истории Руси, Англии и Рима (Том 1. Русь. Том 2. Англия и Рим) / Г.В. Носовский, А.Т. Фоменко. — Москва : Учебно-научный центр довузовского образования МГУ, 1995. — 382 + 290 с.

22. *Носовский, Г.В.* Русь и Рим. Правильно ли мы понимаем историю Европы и Азии? (Том 1. Том 2) / Г.В. Носовский, А.Т. Фоменко. — Москва : Олимп, 1997. — 608 + 624 с.

23. *Носовский. Г.В.* Новая хронология Руси / Г.В. Носовский, А.Т. Фоменко. — Москва : Факториал, 1997. — 256 с.

24. *Носовский, Г.В.* Новая хронология Руси, Англии и Рима / Г.В. Носовский, А.Т. Фоменко. — Москва : Анвик, 1999. — 544 с.

25. *Носовский, Г.В.* Русь-Орда на страницах библейских книг / Г.В. Носовский, А.Т. Фоменко. — Москва : Анвик, 1998. — 432 с.

26. *Фоменко, А.Т.* Методы математического анализа исторических текстов: приложения к хронологии / А.Т. Фоменко. — Москва : Наука, 1996. — 467 с.

27. *Орлов, А.И.* Статистика объектов нечисловой природы / А.И. Орлов // Заводская лаборатория. — 1990. — Т. 56. — № 3. — С. 76–83.

28. *Карякин, Р.Н.* Резонанс в тяговых сетях и его демпфирование / Р.Н. Карякин. — Москва : Высшая школа, 1961. — 231 с.

29. *Крамер, Г.* Математические методы статистики / Г. Крамер. — 2-е изд. — Москва : Мир, 1975. — 648 с.

30. *Тюрин, Ю.Н.* О проверке датчиков псевдослучайных чисел / Ю.Н. Тюрин, В.Э. Фигурнов // Заводская лаборатория. — 1990. — Т. 56. — № 3. — С. 89–92.

31. *Карякин, Р.Н.* Вероятностная теория высших гармоник помех, создаваемых электровозами / Р.Н. Карякин, А.И. Орлов, С.Ю. Адамов // Прикладной многомерный статистический анализ. Ученые записки по статистике. — Т. 33. — Москва : Наука, 1978. — С. 376–380.

32. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в цифровой экономике (на примере управления качеством) / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 169. — С. 216–242.

33. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в организации производства в эпоху цифровой экономики / А.И. Орлов // Инновации в менеджменте. — 2021. — № 2 (28). — С. 36–45.

34. *Орлов, А.И.* Отечественная научная школа в области организационно-экономического моделирования, эконометрики и статистики / А.И. Орлов // Контроллинг. — 2019. — № 73. — С. 28–35.

35. *Орлов, А.И.* Прикладная статистика — состояние и перспективы / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 119. — С. 44–74.

36. *Орлов, А.И.* Статистические и экспертные методы в задачах экономики и управления наукой / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 166. — С. 1–35.

37. *Орлов, А.И.* О развитии теории принятия решений и экспертных оценок / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 167. — С. 177–198.

38. *Орлов, А.И.* Организационно-экономическое моделирование при решении задач управления хозяйственными единицами / А.И. Орлов // Научный журнал КубГАУ. — 2013. — № 87. — С. 637–664.

39. *Орлов, А.И.* Методология моделирования процессов управления в социально-экономических системах / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 101. — С. 166–196.

40. *Куликова, С.Ю.* Организационно-экономическое моделирование при решении задач контроллинга / С.Ю. Куликова, В.С. Муравьева, А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 118. — С. 486 — 506.

41. *Орлов, А.И.* О некоторых подходах к экономико-математическому моделированию малого бизнеса / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 4 (108). — С. 1–28.

42. *Баев, Г.О.* Проблемы управления малыми производственными предприятиями на ранних стадиях жизненного цикла / Г.О. Баев, А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 118. — С. 275 — 304.

43. Демография : учебное пособие / А.И. Щербаков, М.Г. Мдинарадзе, А.Д. Назаров, Е.А. Назарова ; под общей редакцией А.И. Щербакова. — Москва : ИНФРА-М, 2017 — 216 с.

44. *Орлов, А.И.* Оптимальный план управления запасами нельзя найти на основе формулы квадратного корня / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 106. — С. 270–300

45. *Орлов, А.И.* Асимптотика квантования, выбор числа градаций в социологических анкетах и двухуровневая модель управления запасами / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 123. — С. 660–687.

46. *Орлов, А.И.* Статистические методы в истории / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 115. — С. 227–262.

47. *Орлов, А.И.* Новая хронология всеобщей и российской истории — основа государственно-патриотического мировоззрения / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 120. — С. 60–85.

48. *Глазьев, С.Ю.* Новая хронология Фоменко и борьба с кризисом. Интервью Максиму Калашникову 3 октября 2020 г. — URL: https://glazev.ru/articles/165-interv-ju/84410-novaja-khronologija (дата обращения: 14.08.2021).

49. *Орлов, А.И.* Вероятностно-статистическое моделирование помех, создаваемых электровозами / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 106. — С. 225–238.

Контрольные вопросы и задачи

1. Назовите основные виды переменных, используемых в математических моделях.

2. Какие виды математических моделей обычно выделяют?

3. Почему среднюю ожидаемую продолжительность предстоящей жизни считают наиболее адекватной характеристикой здоровья и уровня жизни населения?

4. Какие выводы о возможном изменении численности населения России в течение ближайших 50 лет можно сделать на основе рассмотренных в разделе 8.2 демографических моделей?

5. На складе хранится некоторая продукция, пользующаяся равномерным спросом. За 1 день со склада извлекается 0,5 *т* продукции, плата за хранение 1 *т* продукции в день — 2 тыс. руб., плата за доставку одной партии — 50 тыс. руб. Планирование производится на 21 день. На сколько процентов затраты в плане Вильсона (объем партии определяется по формуле квадратного корня) превышают затраты в оптимальном плане?

6. Оцените увеличение затрат в плане Вильсона (объем партии определяется по формуле квадратного корня) по сравнению с оптимальным планом за целое число периодов, если размер партии отличается от оптимального не более чем на 5 %.

7. Каким образом концепция асимптотически оптимального плана позволяет решить проблему горизонта планирования?

8. В чем состоит основной вклад математики при разработке модели планирования оптимальных размеров поставок и начального запаса?

9. В чем состоит основная идея компьютерно-статистического подхода к построению хронологии?

10. В каком случае два исторических периода считают «дубликатами»?

11. Какие методики обнаружения «дубликатов» Вы знаете?

12. Расскажите о вероятностно-статистической модели, предназначенной для изучения помех, создаваемых электровозами.

Темы докладов, рефератов, исследовательских работ

1. Классификация математических моделей.

2. Математическое моделирование и «математическая экономика».

3. «Точки роста» в теории и практике математического моделирования.

4. Роль нечисловых переменных в современных математических моделях решений.

5. Средние величины для основных характеристик (смертности, рождаемости, продолжительности жизни) в демографических моделях.

6. Экономико-математическое моделирование работы промышленного предприятия.

7. Применение теории массового обслуживания при моделировании работы организаций сферы массовых услуг (телефонных сетей, магазинов и др.).

8. Специфика решения задачи оптимизации при анализе классической модели работы склада.

9. В каком смысле оптимальна двухуровневая модель управления запасами?

10. Современная логистика в системе организационно-экономических методов управления организацией.

11. Реконструкция исторических событий в соответствии с компьютерно-статистической хронологией.

12. Методологическое значение компьютерно-статистической хронологии.

13. Статистическое моделирование с использованием датчиков псевдослучайных чисел (метод Монте-Карло).

14. Методы проверки качества датчиков псевдослучайных чисел.

ГЛАВА 9. СТАТИСТИЧЕСКИЕ МОДЕЛИ ДИНАМИКИ

9.1. Метод ЖОК оценки результатов взаимовлияний факторов

Основные идеи метода компьютерного моделирования ЖОК. При вероятностно-статистическом моделировании технических, социально-экономических, медицинских и иных явлений и процессов исследователь часто сталкивается с тем, что различные факторы постоянно влияют друг на друга. Как правило, для каждого из рассматриваемых факторов можно выделить «непосредственное окружение», которое оказывает на него влияние на него в конкретный момент. На него же этот фактор оказывает некоторое обратное влияние. Дальше начинается самое интересное — волны влияний, порожденные разными факторами, распространяются по всей совокупности факторов, частично усиливают друг друга, частично погашают, порождая в каждый момент времени новые волны.

Разработан компьютерный метод, называемый далее ЖОК, предназначенный для оценки результатов влияния описывающих ситуацию факторов на итоговые показатели и друг на друга. Метод ЖОК позволяет получать выводы, полезные для управления различными структурами на микро — и макроуровнях, от бригад и предприятий до государства в целом. Этот метод использует модель многомерного временного ряда, в которой коэффициенты непосредственного влияния факторов друг на друга и начальные условия задаются экспертами. Опишем основные составляющие этого метода.

Сначала экспертным путем определяется список факторов, которые необходимо учитывать при анализе конкретной ситуации. В качестве примера рассмотрим здесь типовое промышленное предприятие. Для него такими факторами могут являться, на наш взгляд, устойчивость развития, уровень рентабельности, оценка состояния основных и оборотных фондов, положение на рынке, кадровый потенциал, финансовое положение, технологический уровень, технический уровень и качество продукции, степень учета экологических требований, уровень сертификации, научно-технический потенциал и степень его использования, положение в социальной сфере, развитость профсоюзного движения, оценка отношений с конкурентами и властями и т.д. Основная часть перечисленных факторов носит качественный характер.

Далее определяются необходимые для работы модели начальные уровни факторов, соответствующие начальному состоянию изучаемого объекта (проводится оцифровка нечисловых переменных). Они оцениваются экспертами на шкале от (– 1) до (+ 1) с шагом 0,1. В методе ЖОК степень привлечения экспертов может быть различна. От использования одного эксперта, хорошо знающего ситуацию и на основе своих знаний и интуиции указывающего необходимые параметры и связи, до подключения к работе комиссии экспертов, коллективно оценивающих указанные параметры и связи, с использованием той или иной схемы сбора и анализа экспертных мнений.

Затем экспертами составляется блок-схема непосредственных влияний факторов друг на друга и оценивается степень непосредственных влияний с помощью такой же шкалы от (– 1) до (+ 1) с шагом 0,1, Получается модель в виде взвешенного ориентированного графа с начальными данными в вершинах. Она несколько напоминает хорошо известную экономистам схему межотраслевого баланса В. Леонтьева, но в отличие от нее использует не только количественные, но, в основном, качественные факторы. Затем просчитываются итерации (опосредованные влияния второго, третьего и т.д. уровней, соответствующие второму, третьему и т.д. моментам времени) вплоть до получения стабильного состояния. Результат работы модели — конечные уровни факторов.

Модель позволяет просчитать развитие экономической структуры при различных сценариях. Обычно одновременно используют три типа сценариев: «Прогноз», «Поиск», «Оптимизация».

Сценарий «Прогноз» показывает результат при отсутствии управляющих воздействий. Он демонстрирует, как будет развиваться ситуация, если в нее не вмешиваться. Исходные данные для сценария «Прогноз» — начальные значения факторов и матрица непосредственных взаимовлияний факторов.

В сценариях типа «Поиск» вводится новое понятие — управляющие факторы. В сценариях этого типа анализируются результаты изменений при наличии тех или иных конкретных воздействий на управляющие факторы. Обычно специалист, работающий с системой ЖОК, имеет целью увеличение значений тех или иных факторов при «удержании» некоторых иных в заданных пределах. Однако от него не требуется сообщать свои цели компьютерной системе. В сценариях типа «Поиск» осуществляется эвристический процесс оптимизации, а также анализ поведения системы при тех или иных воздействиях на начальные значения факторов.

В сценариях типа «Оптимизация» кроме списка управляющих факторов задаются целевые факторы и условия на них, которых необходимо добиться. Обычно это — условия выхода на определенные уровни, например, рентабельность должна быть не менее 0,5, а социальная напряженность — не более 0,3. С помощью оптимизационных алгоритмов находится наилучшее управление, позволяющее достигнуть цели или максимально к ней приблизиться. Однако найденные компьютером рекомендации могут включать слишком резкие изменения тех или иных начальных параметров, поэтому результаты расчетов скорее указывают на перспективные варианты изменения управляющих параметров, чем непосредственно задают план действий. С помощью сценариев типа «Поиск» можно на основе этих результатов найти практически реализуемые рекомендации.

Система ЖОК позволяет проследить динамику изменения значений факторов вплоть до их стабилизации, которая обычно наступает через 15–25 итераций (интервалов времени). Такая быстрая сходимость вначале кажется неожиданной. Возможно, сам факт стабилизации является самым важным методологическим выводом из экспериментов с моделью ЖОК. После первоначальных всплесков замкнутая экономическая система стабилизируется, хотя бы и на весьма низком уровне производства и потребления.

При этом с помощью оцененных экспертами коэффициентов важности факторов (с учетом знака) можно отслеживать общую оценку экономической ситуации.

Система ЖОК является человеко-машинной. Для эффективной работы специалиста желательно, чтобы общее число факторов, используемых в конкретной модели, не превышало 20, а число непосредственных взаимосвязей — 40, хотя эти ограничения несущественны для математического обеспечения компьютерной системы ЖОК. Однако они существенны для обеспечения наглядности при построении, обсуждении и совершенствовании модели, для того чтобы факторы и связи между ними можно было изобразить на листе бумаги или экране компьютера в виде блок-схемы.

Система ЖОК с успехом использовалась для анализа ряда конкретных экономических ситуаций. Так, по заказу Минфина РФ она применялась для анализа взаимовлияний факторов, определяющих динамику налогооблагаемой базы и сбора подоходного налога с физических лиц, налога на имущество, налогов и сборов за пользование природными ресурсами и др. Построенная серия моделей обладала некоторыми общими чертами. Прогноз, исходящий из современного экономического положения, во всех случаях указывал на дальнейшее ухудшение ситуации. Активное вмешательство государства в экономику приводило к значительному улучшению показателей, в то время как управление с помощью чисто экономических (монетаристских) методов не позволяло улучшить исходное положение. Полученные результаты подтверждают известную концепцию пяти нобелевских лауреатов по экономике (К. Эрроу, В. Леонтьев и др.), разрабатываемую совместно с Отделением экономики Российской академии наук (Д.С. Львов, С.Ю. Глазьев и др.), о необходимости активного регулирования государством экономических процессов [1, 2].

Другие примеры применения системы ЖОК касались оптимизации экономической стороны деятельности промышленного предприятия или организации в иной сфере, экономических взаимоотношений отраслей народного хозяйства. А также макроэкономического моделирования, в ходе которого удалось вскрыть две неточности в основной схеме известного учебника [3], а затем исправить их, включив дополнительные блоки в соответствующую модель.

9.2. Система моделей налогообложения

Пример применения метода ЖОК: система моделей налогообложения. Подоходный налог на физических лиц — один из основных в системе налогообложения, действующей в Российской Федерации. Построена система моделей для изучения влияния различных факторов на налогооблагаемую базу подоходного налога на физических лиц. Построение модели предполагает выявление факторов, влияющих на налогооблагаемую базу, и взаимосвязей между этими факторами, численную оценку взаимовлияний факторов. После формулировки сценариев развития ситуаций последовало проведение расчетов и анализ полученных результатов. Расчеты проводились с помощью оригинального математического и программного обеспечения, разработанного в Институте высоких статистических технологий и эконометрики МГТУ им. Н.Э. Баумана.

Очевидными факторами, влияющими на налогооблагаемую базу подоходного налога, являются:

1) объем производства, соответственно, численность занятых работников и объем ФОТ (фонда оплаты труда);

2) размеры полной, скрытой, частичной безработицы;

3) объем доходов, не облагаемых налогами: от натурального приусадебного (садово-огородного) и домашнего хозяйства, частных услуг в области мелкого ремонта. репетиторства и др.;

4) инфляция, частично компенсируемая повышением заработной платы;

5) уклонение от уплаты налога отдельными физическими и юридическими лицами;

6) бартер, «черный нал» (миновавший кассу) и другие явления теневой экономики;

7) объем неплатежей в целом и невыплаты заработной платы в частности.

В учебном пособии [4, с. 245–265] перечислено более 50 видов полного или частичного освобождения граждан от уплаты подоходного налога. Перечень этот не полон: например, лицо, купившее квартиру, освобождается от уплаты определенной части подоходного налога (государство частично компенсирует стоимость квартиры).

Надо иметь в виду, что Росстат дает не реальное, а рассчитанное распределение населения по величине получаемого дохода (на основе логарифмически нормального закона). Поэтому по данным Росстата нельзя рассчитать влияние изменения ставок подоходного налога на суммарный сбор.

В различных странах сложилась различная доля оплаты труда в стоимости изделия. В США в 1988 г. ВНП составил 4 862 млрд долл., в то время как личный доход — 4 063 млрд, а индивидуальные налоги — 590 млрд (14,5 % от личных доходов, 12,13 % от ВНП). В России в 1995 г. подоходный налог составил 2 770,9 млрд руб., или 2,17 %: от всех налоговых сборов, или примерно 0,26 % от ВВП, т.е. доля этого налога примерно в 50 раз меньше, чем в США.

После предварительного анализа налоговой ситуации в России была построена пробная модель динамики экономики, включающая 13 факторов и 32 взаимосвязи между ними. В настоящем разделе она не рассматривается, поскольку включает в себя единый фактор налогообложения, без отдельного выделения подоходного налога с физических лиц.

Далее был рассмотрен список из 41 фактора, которые влияют на налогооблагаемую базу подоходного налога с физических лиц. После анализа из них были отобраны 30 факторов. На их основе построена модель НФЛ-30 (НФЛ — по первым буквам: «Налог на Физических Лиц»). Она включает эти 30 факторов и 465 взаимосвязей между ними. Расчет по модели НФЛ-30 показал, что современная ситуация в России ведет к уходу экономики «в тень», росту сокрытия налогов и коррупции, падению доверия к власти и готовности платить налоги, падению уровня жизни населения и очень сильному сокращению налогооблагаемой базы.

Однако модель НФЛ-30 является слишком сложной и трудной для непосредственного анализа. В частности, трудно было найти рациональное управление, обеспечивающее рост налогооблагаемой базы подоходного налога. Поэтому модель НФЛ-30 была упрощена в основном за счет исключения слабых и/или дублирующих взаимосвязей с переносом внимания на непосредственные взаимодействия. Кроме того, три фактора были исключены, а один добавлен.

В результате построена модель НФЛ-28 с 28 факторами и 64 взаимосвязями. Расчет по этой модели показал наличие больших возможностей у государства в целом и у государственных налоговых органов (ГНО) в частности по расширению налогооблагаемой базы подоходного налога.

Характер взаимодействий факторов в модели НФЛ-28 описывается с помощью ориентированного графа (орграфа), дугам которого приписаны весовые коэффициенты от (– 1) до (+ 1). От конкретного фактора дуги ведут к тем факторам, на которые этот фактор непосредственно влияет. Влияние может быть как положительным, так и отрицательным, возрастание абсолютной величины означает увеличение степени влияния.

Таким образом, построение модели НФЛ-28 изучения роста налогооблагаемой базы подоходного налога с физических лиц, как и других моделей рассматриваемого типа, состоит из ряда операций, осуществляемых экспертами, а именно:

- выделения и обоснования системы факторов, включаемых в модель;
- оценки важности факторов (по десятибалльной шкале) с учетом знаков — по отношению к задаче расширения налогооблагаемой базы подоходного налога с физических лиц;
- построения ориентированного графа непосредственного влияния факторов друг на друга (важно избегать связей, соответствующих опосредованным влияниям, а также по возможности сократить число циклов);
- приписывания дугам весов из интервала [– 1; + 1], отражающих степень влияния (матрица весов не должна содержать слишком много ненулевых элементов).

Затем с помощью имеющегося диалогового комплекса ЖОК проводятся расчеты по ряду сценариев. Результаты интерпретации итогов расчетов позволяют проанализировать свойства модели.

При описании конкретных взаимовлияний факторов в моделях рассматриваемого типа используется следующая шкала соответствия между лингвистической и числовой шкалами:

- очень сильно возрастает (0,9; 1,0);
- значительно возрастает (0,7; 0,8);
- существенно возрастает (0,5; 0,6);
- умеренно возрастает (0,3; 0,4);
- очень слабо возрастает (0,1; 0,2);
- очень слабо убывает (– 0,1; – 0,2);
- умеренно убывает (– 0,3; – 0,4);
- существенно убывает (– 0,5; – 0,6);
- значительно убывает (– 0,7; – 0,8);
- очень сильно убывает (– 0,9; – 1,0).

Однако и модель НФЛ-28 достаточно сложна для анализа. Поэтому были построены еще две модели — НФЛ-18 и НФЛ-19, множества факторов которых различны.

Основой для дальнейшего анализа служат две модели, обозначаемые как НФЛ-18 и НФЛ-19, т.е. модели подоходного налога на физических лиц с использованием 18 и 19 факторов соответственно. Модель НФЛ-18 с 18 факторами и 31 взаимосвязью межу ними основана на гипотезе об активном участии государства в экономических процессах (с помощью законотворчества, госзаказов и др.). В то время как модель НФЛ-19 с 19 факторами и 31 взаимодействием предполагает лишь косвенное вмешательство государства путем установления ставок таможенных сборов, борьбы с криминальным миром и др. Эта модель построена в основном на чисто экономических взаимодействиях, без непосредственного регулирующего влияния государства.

Начальные значения факторов — это приращения (изменения, «дифференциалы») включенных в модели экономических величин. Они выбраны исходя из оценки экономического положения России в июне 1999 г. Конечные значения факторов определяются рассматриваемым сценарием. Они также интерпретируются как приращения включенных в модели экономических величин.

В каждой из моделей НФЛ-18 и НФЛ-19 рассмотрено четыре сценария:

- прогнозируется динамика налогооблагаемой базы подоходного налога при отсутствии управляющих воздействий (сценарий обозначается как Пассивный-1);

- прогнозируется динамика налогооблагаемой базы подоходного налога при заданных управляющих воздействиях (сценарий обозначается как Активный-1);

- ищутся управляющие воздействия, позволяющие добиться прироста не менее 0,7 налогооблагаемой базы подоходного налога (сценарий обозначается как Цель-1);

- ищутся управляющие воздействия, позволяющие добиться прироста не менее 0,5 налогооблагаемой базы подоходного налога и прироста не менее 0,3 уровня жизни населения (сценарий обозначается как Цель-2).

Таким образом, в сценариях Цель-1 и Цель-2 выделяются целевые факторы, значения которых требуется максимизировать. В сценарии Цель-1 это один фактор — налогооблагаемая база подоходного налога, в сценарии Цель-2 добавляется еще один — уровень жизни населения. В сценарии Активный-1 выделяется другое подмножество факторов — управляющие, путем воздействия на которые специалист, работающий с моделью, может попытаться изменить неблагоприятные тенденции и улучшить ситуацию. В сценариях Цель-1 и Цель-2 управляющие факторы используются несколько по-иному — выбор их значений осуществляет не специалист, а компьютер в соответствии с алгоритмом оптимизации. Выделяются также наблюдаемые факторы, наиболее важные для анализа начальной и конечной экономической ситуации в каждом из сценариев. Во всех сценариях начальные значения факторов соответствуют современному состоянию экономики России.

Используется также количественная оценка экономической ситуации, представляющая собой взвешенную сумму значений факторов. Для ее расчета каждому используемому фактору приписано число от (– 10) до 10 — его важность, при этом желательное направление изменения фактора определяет знак указанного числа: если желателен рост, то ставится плюс, если желательно убывание — минус.

Основная часть подраздела начинается с описания факторов, включенных в модели НФЛ-18 и НФЛ-19, и обоснования необходимости их включения в модели. Имеется 12 факторов, включенных в обе модели, поэтому всего используется 25 факторов. Затем приводятся схемы (графы) взаимного влияния факторов и матрицы взаимного влияния факторов в моделях НФЛ-18 и НФЛ-19.

Результаты проведенной работы показывают, что использованный эконометрический аппарат на основе ориентированных графов позволяет получать качественные выводы, полезные для выбора стратегии управления процессами налогообложения.

Построение моделей НФЛ-18 и НФЛ-19. Анализ экономической ситуации, нацеленный на изучение динамики налогооблагаемой базы подоходного налога, приводит к выделению следующих блоков факторов: макроэкономические показатели; показатели доходов и уровня жизни населения; участие государства в экономической жизни; факторы, относящиеся непосредственно к сфере налогов; характеристики теневой экономики и криминального мира. Далее выделяются и обосновываются факторы, используемые в моделях НФЛ-18 и НФЛ-19. Поскольку компьютерные модели нацелена на изучение влияния изменений (приращений, дифференциалов) одних факторов на изменения других, то в названиях факторов имеются термины «рост», «прирост» и др. Для компьютерных моделей, имеющих качественный характер, факторы берутся в обобщенном виде, что отмечается при их описании. Модели, охватывающие столь обширную область экономической жизни, не могут не опираться на факторы в обобщенном виде, иначе они станут необозримыми для восприятия и бесполезными для применения.

Приведем сначала описание факторов, используемых в модели НФЛ-18.

Таблица 1

НФЛ-18

№ п/п	Фактор
1.	**Прирост налогооблагаемой базы подоходного налога** — прирост тех доходов населения, с которых, согласно действующему законодательству, реально может быть взят подоходный налог в денежной форме
2.	**Прирост ВВП** (валового внутреннего продукта) — прирост совокупного объема продукции и услуг, произведенных на территории России, базовый макроэкономический показатель, характеризующий положение и динамику экономики России
3.	**Увеличение роли государства в экономике** связано с созданием эффективно работающего аппарата государственного управления. В том числе налоговых органов; с ростом уровня налогового законодательства (т.е. с ростом согласованности и применимости законов, их обоснованности, соответствия национальным традициям, доступности для понимания масс налогоплательщиков). Например, в части различных льгот по уплате подоходного налога, роста доли зарплаты в цене товара (услуги); управлением финансовой сферой, борьбой с криминальным миром, коррупцией, теневой экономикой, эффективной организацией производством товаров и услуг в государственном секторе, повышением уровня жизни населения через систему трансфертов и др.
4.	**Рост государственных заказов,** в частности, расходов на покупку услуг населения путем организации, расширения или восстановления государственных предприятий, строительства дорог и проведения тех или иных «общественных работ». Согласно современной макроэкономической теории — необходимая мера в период спада
5.	**Рост кредитования отечественных товаропроизводителей** (за счет средств населения, находящихся в банках) — необходимое условие повышения производства отечественных товаров и услуг с целью удовлетворения спроса на них. Кредиты должны быть предназначены как для оптимизации оборотных средств предприятий, исключения бартера и неплатежей, так и для капиталовложений (инвестиций) с целью перехода на современные технологии
6.	**Повышение спроса на отечественную продукцию,** обеспеченное соответствующим повышением ее выпуска, приводит к увеличению объема средств отечественных производителей, в частности, к увеличению фонда оплаты труда (ФОТ), а потому и поступлений подоходного налога. Речь идет о реализованном, т.е. оплаченном спросе
7.	**Рост уровня работы государственных налоговых органов (ГНО)** предполагает совершенствование налогового законодательства, рост налоговой грамотности населения, индивидуальную работу с каждым налогоплательщиком, рост реальности и неотвратимости санкций за неуплату налогов, широкую информированность общества о них
8.	**Рост качества работы банковской системы** предполагает, своевременное осуществление платежей, существенное снижение необоснованно завышенных плат за банковские услуги, за перевод средств со счета на счет, за пользование банкоматами и др.
9.	**Рост доверия населения к государственной власти** и готовности платить налоги становится реальной силой как непосредственно в экономической жизни, так и в борьбе против криминального мира и теневой экономики — основного внутреннего врага государства

№ п/п	Фактор
10.	**Рост уровня жизни населения** повышает желание и возможность уплаты подоходного налога, а его снижение действует в обратном направлении. Понятие «уровень жизни» многогранно, помимо дохода за последний промежуток времени включает в себя и использование предыдущих накоплений, и способы распоряжения доходом и личной собственностью [5]
11.	**Прирост объема** (начисленных) **выплат из ФОТ** предприятий и организаций — основной источник поступлений подоходного налога. Влияние неплатежей сказывается в уменьшении реальных выплат из ФОТ и соответствующего уменьшения поступлений подоходного налога. Влияние изменения объемов других источников доходов, облагаемых подоходным налогом [4], учитывается с помощью связей с показателями уровня жизни и инфляции
12.	**Рост сбережений (накоплений)** населения, находящихся в банках и могущих быть использованными для кредитования отечественных товаропроизводителей — важный показатель доверия к государству, а потому и готовности платить налоги. Он противопоставляется, с одной стороны, практическому отсутствию сбережений у части населения, с другой стороны, хранению сбережений вне банков («в чулке»)
13.	**Прирост уровня занятости** населения на официально зарегистрированных работах — понятие, противоположное понятию прироста безработицы. Однако при обсуждении безработицы не всегда ясно, является ли отказ от работы вынужденным или добровольным (домохозяйки с детьми, лица старшего возраста). Имеются расхождения (в несколько раз) при оценке уровня безработицы, например, по числу зарегистрированных на биржах труда и расчетами по правилам МОТ (Международной Организации Труда)
14.	**Инфляция** — прирост общего уровня цен (при падении цен он может быть отрицательным), измеряемый по специальным методикам
15.	**Прирост неплатежей** (работникам и организациям) дезорганизует экономическую жизнь в целом и налоговую сферу в частности, приводит к приросту бартера, к выдаче заработной платы в натуральной форме, в результате — к сокращению налоговых поступлений (в т.ч. подоходного налога) и поступлений во внебюджетные фонды
16.	**Прирост доходов в домохозяйствах, остающихся вне сферы ГНО** — это прирост доходов в натуральной форме (сельскохозяйственная продукция, выращенная на собственном приусадебном или садово-огородном участке, даче для семейного употребления, домашняя работа (ремонт, изготовление мебели, одежды, пищи)). А также прирост доходов от мелких услуг типа косметических ремонтных работ в квартирах, индивидуального пошива одежды, ремонта обуви, репетиторства
17.	**Прирост сокрытия доходов и уклонения от уплаты налогов** частично связан с нарушениями нормального хода экономической жизни, в том числе с вынужденными неплатежами и бартером, ведущими к выплате заработной платы натурой и выдаче собственности организаций в пользование работникам, с налоговой неграмотностью населения. Но в основном имеет криминальный характер, в частности, порожденный распространением «черного нала»
18.	**Рост криминального мира**, **теневой экономики** и коррупции , т.е. сил, противостоящих государственной власти, приводит к «скрытию» от ГНО части экономики (т.н. «теневой») вместе со всеми налогами, которые обязаны были бы платить «спрятавшиеся» организации и лица

№ п/п	Фактор
В модели НФЛ-19 отсутствуют факторы 3, 4, 6, 8, 9, 15, но включены следующие семь факторов:	
19.	**Усиление борьбы государства с криминалом в экономике,** т.е. лишь одна, но наиболее важная часть из сферы **усиления роли государства в экономике** согласно описанному выше фактору 3
20.	**Улучшение финансового положения предприятий,** т.е. увеличение находящихся в распоряжении предприятий денежных средств, позволяет увеличить выпуск продукции и услуг (ВВП), осуществлять инвестиции, уменьшает неплатежи и бартер, увеличивает уровень жизни населения, его денежные доходы, а потому и сбор подоходного налога
21.	**Улучшение финансового положения бюджетной сферы** увеличивает уровень жизни работников бюджетной сферы, его денежные доходы, а потому и сбор подоходного налога
22.	**Улучшение внешнеэкономической ситуации** — повышение цен на отечественные товары и услуги, расширение иностранных инвестиций в отечественные предприятия улучшают экономическую ситуацию в России, повышают уровень жизни населения и увеличивают сбор подоходного налога
23.	**Рост курса доллара США** приводит к инфляции и уменьшает жизненный уровень населения, но увеличивает долю рынка, принадлежащую отечественным товаропроизводителям, и в итоге улучшает их финансовое положение
24.	**Повышение таможенных сборов на импортную продукцию** увеличивает объем доходов государства, а потому улучшает финансовое положение бюджетной сферы и части населения, увеличивает долю рынка, принадлежащую отечественным товаропроизводителям, и в итоге улучшает их финансовое положение, но одновременно приводит к инфляции и уменьшает жизненный уровень населения
25.	**Повышение таможенных сборов на экспортную продукцию** увеличивает объем доходов государства, а потому улучшает финансовое положение бюджетной сферы и части населения, но одновременно ухудшает финансовое положения отечественных товаропроизводителей

Следующий шаг — составление схем взаимодействия факторов в рассматриваемых моделях. На основе опроса экспертов получены схемы 1 и 2, представленные ниже.

Матрицы коэффициентов взаимного влияния факторов в моделях НФЛ-18 и НФЛ-19. Степень слияния факторов друг на друга можно оценить с помощью элементов матрицы влияния. Перечислим эти элементы в соответствии со схемой влияния факторов. Слева указан влияющий фактор, справа — фактор, на который оказывается влияние. Конкретные значения получены с помощью экспертов.

Начнем с модели НФЛ-18.

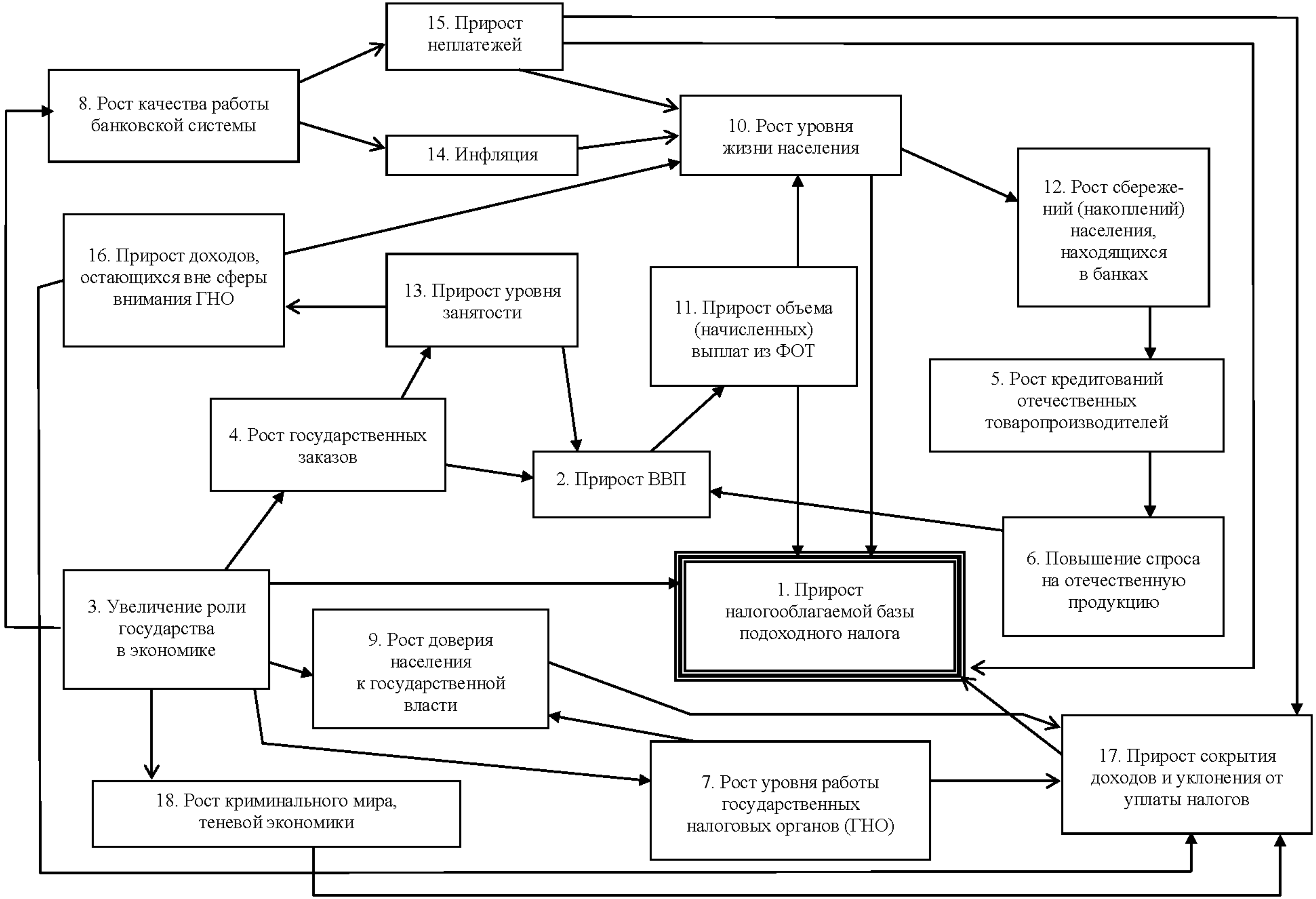


Схема 1. Модель НФЛ-18

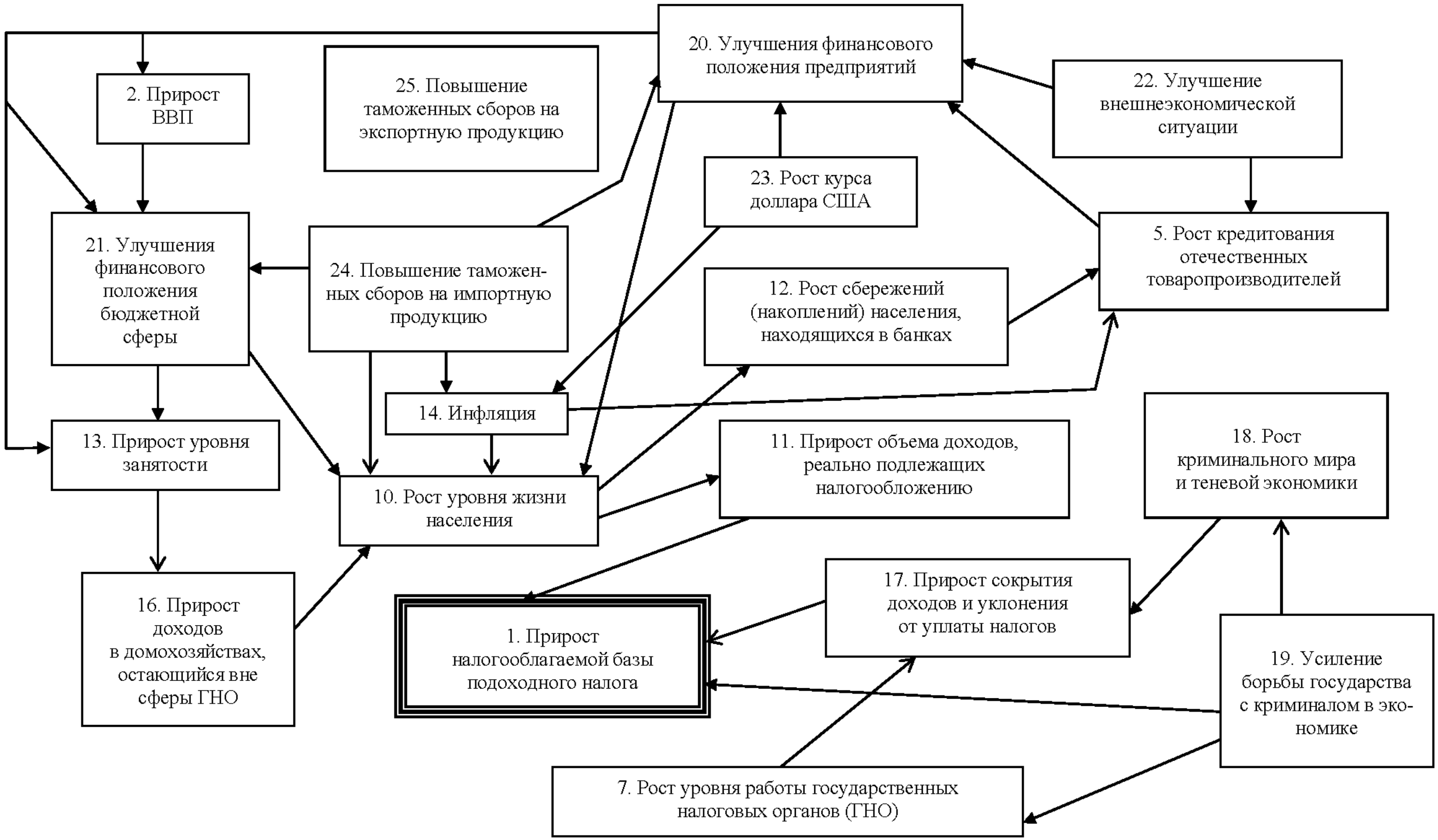


Схема 2. Модель НФЛ-19

Таблица 2

НФЛ-18

№ п/п	Фактор
2-11	Прирост ВВП оказывает значительное положительное влияние на прирост объема (начисленных) выплат из ФОТ, оцениваемое величиной 0,7
3-1	Увеличение роли государства в экономике умеренно увеличивает налогооблагаемую базу подоходного налога, в частности, с помощью законодательных мер (установлением льгот и др.), коэффициент влияния принимаем равным 0,3
3-4	Увеличение роли государства в экономике в значительной мере проявляется в увеличении государственных заказов, влияние оцениваем числом 0,7
3-7	Увеличение роли государства в экономике очень сильно проявляется в росте уровня работы государственных налоговых органов (ГНС), степень связи оцениваем как 0,9
3-8	Увеличение роли государства в экономике означает и значительное повышение качества работы банковской системы, что обеспечивается работой законодательных и исполнительных государственных органов, прежде всего Центрального банка и МВД. Степень влияния оценивается числом 0,7
3-9	Увеличение роли государства в экономике приводит к существенному возрастанию доверия населения к государственной власти, а потому и готовности платить налоги. Возрастает поддержка государственной власти народом, сплочение вокруг нее. Коэффициент влияния принимаем равным 0,6
3-18	Увеличение роли государства в экономике приводит к значительному уменьшению криминального мира и теневой экономики, как за счет непосредственной борьбы государства с ними, так и за счет «поворота» народных масс от «мафии» к государству. Степень влияния оценивается числом (– 0,7)
4-2	Рост государственных заказов приводит к значительному приросту ВВП, и степень влияния оценивается как 0,7
4-13	Рост государственных заказов ведет к существенному росту уровня занятости. Коэффициент влияния равен 0,5
5-6	Рост кредитования отечественных товаропроизводителей ведет к значительному увеличению выпуска конкурентоспособной и пользующейся спросом отечественной продукции. Степень влияния оценивается числом 0,7
6-2	Повышение спроса на отечественную продукцию ведет к очень сильному росту ВВП с коэффициентом влияния 0,9
7-9	Рост качества работы государственных налоговых органов (ГНО) вызывает существенный рост доверия населения к государству и готовности платить налоги. Коэффициент влияния принимаем равным 0,5
7-17	Рост качества работы государственных налоговых органов (ГНО) значительно уменьшает сокрытие доходов и уклонение от уплаты налогов. Степень влияния оцениваем числом (– 0,7)
8-14	Рост качества работы банковской системы с помощью различных финансовых инструментов позволяет существенно сократить инфляцию. Коэффициент влияния принимаем равным (– 0,5)

№ п/п	Фактор
8-15	Рост качества работы банковской системы позволяет существенно сократить неплатежи. В частности, с помощью системы взаимозачетов и за счет ускорения и целевого использования выделенных федеральным центром средств. Коэффициент влияния принимаем равным (– 0,5)
9-17	Рост доверия к государственной власти и готовности платить налоги ведет к существенному сокращению сокрытия доходов и уклонения от уплаты налогов с коэффициентом влияния (– 0,5)
10-1	Рост уровня жизни населения приводит к существенному расширению налогооблагаемой базы подоходного налога. Как из-за общего повышения денежных доходов населения, подлежащих налогообложению, так и за счет сокращения льгот для малообеспеченных граждан, доля которых уменьшается при общем росте уровня жизни. Степень влияния оцениваем числом 0,6
10-12	Рост уровня жизни населения при прочих равных условиях ведет к существенному увеличению сбережений (накоплений) граждан, находящихся в банках, с коэффициентом влияния 0,6
11-1	Прирост объема (начисленных) выплат из ФОТ ведет к очень сильному приросту налогооблагаемой базы подоходного налога (однако однозначной связи нет из-за неплатежей и льгот) с коэффициентом влияния 0,9
11-10	Прирост объема (начисленных) выплат из ФОТ к значительному росту уровня жизни населения (однако полного соответствия нет из-за неплатежей, доходов вне ГНО, трансфертов и льгот). Степень влияния оцениваем числом 0,7
12-5	Рост сбережений (накоплений) населения, находящихся в банках, дает возможность существенного роста кредитования отечественных товаропроизводителей, с коэффициентом влияния 0,6
13-2	Прирост уровня занятости, т.е. сокращение безработицы, ведет к умеренному возрастанию ВВП с коэффициентом влияния 0,4
13-16	Прирост уровня занятости, т.е. сокращение безработицы, ведет к умеренному сокращению доходов, остающихся вне сферы влияния ГНС, в частности, доходов от личного натурального хозяйства и доходов от личных услуг между домохозяйствами. Степень влияния оцениваем числом (– 0,4)
14-10	Инфляция значительно снижает уровень жизни населения, толкает его в сторону сокрытия доходов, уклонения от уплаты налогов, получения доходов способами, остающимися вне влияния ГНС. Степень влияния оцениваем числом (– 0,7)
15-1	Прирост неплатежей существенно сокращает налогооблагаемую базу подоходного налога с коэффициентом влияния (– 0,5)
15-10	Прирост неплатежей существенно снижает уровень жизни населения. Степень влияния оценивается числом (– 0,6)
15-17	Прирост неплатежей приводит к умеренному росту сокрытия доходов и уклонения от уплаты налогов с коэффициентом влияния 0,4
16-10,	Прирост доходов, остающихся вне сферы внимания ГНС, приводит к умеренному росту уровня жизни населения с коэффициентов влияния 0,3

№ п/п	Фактор
16-17	Прирост доходов, остающихся вне сферы внимания ГНО, влечет значительный рост сокрытия доходов и уклонения от уплаты налогов с коэффициентом влияния 0,7
17-1	Прирост сокрытия доходов и уклонения от уплаты налогов значительно снижает налогооблагаемую базу подоходного налога с коэффициентом влияния (– 0,7)
18-17	Рост криминального мира и теневой экономики очень сильно влияет на прирост сокрытия доходов и уклонения от уплаты налогов с коэффициентом влияния 0,9

Перейдем теперь к модели НФЛ-19.

Таблица 3

НФЛ-19

№ п/п	Фактор
2-21	Прирост ВВП влечет улучшение финансового положения государства, увеличение объема бюджета, а потому существенно улучшает финансовое положение бюджетной сферы с коэффициентом 0,6
5-20	Рост кредитования отечественных товаропроизводителей ведет к значительному увеличению выпуска конкурентоспособной и пользующейся спросом отечественной продукции, к. значительному улучшению финансового положения предприятий. Степень влияния оценивается числом 0,7
7-17	Рост качества работы государственных налоговых органов (ГНО) значительно уменьшает сокрытие доходов и уклонение от уплаты налогов. Степень влияния оцениваем числом (– 0,7)
10-12	Рост уровня жизни населения при прочих равных условиях ведет к существенному увеличению сбережений граждан, находящихся в банках, с коэффициентом влияния 0,6
11-1	Прирост объема (начисленных) выплат из ФОТ ведет к очень сильному приросту налогооблагаемой базы подоходного налога (однако однозначной связи нет из-за неплатежей и льгот) с коэффициентом влияния 0,9
11-10	Прирост объема (начисленных) выплат из ФОТ к значительному росту уровня жизни населения (однако полного соответствия нет из-за неплатежей, доходов вне ГНС, трансфертов и льгот). Степень влияния оцениваем числом 0,7
12-5	Рост сбережений (накоплений) населения, находящихся в банках, дает возможность существенного роста кредитования отечественных товаропроизводителей, с коэффициентом влияния 0,6
13-16.	Прирост уровня занятости, т.е. сокращение безработицы, ведет к умеренному сокращению доходов, остающихся вне сферы влияния ГНС, в частности, доходов от личного натурального хозяйства и доходов от личных услуг между домохозяйствами. Степень влияния оцениваем числом (– 0,4)
14-5	Инфляция существенно снижает объем кредитования отечественных товаропроизводителей, поскольку заставляет банки сокращать сроки кредитования и направляет кредитные потоки в сторону торговых предприятий. Степень влияния оцениваем числом (– 0,6)

№ п/п	Фактор
14-10	Инфляция значительно снижает уровень жизни населении. Толкает его в сторону сокрытия доходов, уклонения от уплаты налогов, получения доходов способами, остающимися вне влияния ГНС. Степень влияния оцениваем числом (– 0,7)
16-10	Прирост доходов, остающихся вне сферы внимания ГНО, приводит к значительному росту уровня жизни населения с коэффициентов влияния 0,7
17-1	Прирост сокрытия доходов и уклонения от уплаты налогов значительно снижает налогооблагаемую базу подоходного налога с коэффициентом влияния (– 0,7)
18-17	Рост криминального мира и теневой экономики очень сильно влияет на прирост сокрытия доходов и уклонения от уплаты налогов с коэффициентом влияния 0,9
19-1	Усиление борьбы государства с криминалом в экономике, в том числе в виде законотворчества, умеренно увеличивает ВВП, переводя часть «теневой» экономики в доступную учету область и защищая официально признанную предпринимательскую деятельность (коэффициент 0,3)
19-7	Усиление борьбы государства с криминалом в экономике очень сильно проявляется в росте уровня работы государственных налоговых органов (ГНС), степень связи оцениваем как 0,9
19-18	Усиление борьбы государства с криминалом в экономике приводит к значительному уменьшению криминального мира и теневой экономики, как за счет непосредственной борьбы государства с ними, так и за счет «поворота» народных масс от «мафии» к государству. Степень влияния оценивается числом (– 0,7)
20-2	Улучшение финансового положения предприятий приводит к значительному возрастанию ВВП в результате ликвидации неплатежей, роста капиталовложений (инвестиций), что ведет к росту производства товаров и услуг (коэффициент 0,7)
20-11	Улучшение финансового положения предприятий приводит к значительному возрастанию объема выплат из ФОТ (коэффициент 0,8)
20-13	Улучшение финансового положения предприятий приводит к существенному возрастанию занятости. Степень влияния оценивается числом 0,5
20-21	Улучшение финансового положения предприятий приводит к умеренному улучшению финансового положения бюджетной сферы. В основном за счет увеличения сбора налогов и расширения заказов предприятий организациям бюджетной сферы (коэффициент 0,3)
21-11	Улучшение финансового положения бюджетной сферы приводит к значительному возрастанию объема выплат из ФОТ (коэффициент 0,8)
21-13	Улучшение финансового положения бюджетной сферы приводит к существенному возрастанию занятости. Степень влияния оценивается числом 0,5
22-5	Улучшение внешнеэкономической ситуации дает возможность умеренного возрастания кредитования отечественных товаропроизводителей, как за счет выручки экспортеров, так и путем прямых иностранных инвестиций (коэффициент 0,3)
22-20	Улучшение внешнеэкономической ситуации приводит к умеренному улучшению финансового положения предприятий, в основном за счет продажи продукции на экспорт (коэффициент 0,4)

№ п/п	Фактор
23-14	Рост курса доллара США приводит к существенному росту цен (инфляции). Степень влияния оценивается числом 0,5
23-20	Рост курса доллара США существенно улучшает финансовое положение предприятий. Увеличивая их долю отечественного рынка и облегчая (путем снижения издержек) выход на зарубежный рынок. Степень влияния оценивается числом 0,5
24-10	Повышение таможенных сборов на импортную продукцию приводит к существенной инфляции (коэффициент 0,5)
24-14.	Повышение таможенных сборов на импортную продукцию приводит к повышению ее цены, а потому к существенному снижению уровня жизни населения (коэффициент 0,5)
24-20	Повышение таможенных сборов на импортную продукцию существенно улучшает финансовое положение предприятий, избавляя их от иностранных конкурентов внутри страны (коэффициент 0,5)
24-21	Повышение таможенных сборов на импортную продукцию значительно улучшает финансовое положение бюджетной сферы за счет сборов, поступающих в бюджет государства. Степень влияния оценивается как 0,8
25-20	Повышение таможенных сборов на экспортную продукцию приводит к умеренному ухудшению финансового положения предприятий за счет фактического изъятия в бюджет части средств за проданную ими продукцию. Степень влияния оценивается числом (– 0,3)
25-21	Повышение таможенных сборов на экспортную продукцию значительно улучшает финансовое положение бюджетной сферы за счет сборов, поступающих в бюджет государства. Степень влияния оценивается как 0,8

Основные результаты моделирования. Были рассмотрены сценарии прогноза (Пасс-1), активного воздействия (Акт-1), оптимального управления (Цель-1 и Цель-2). Сценарий Цель-1 состоял в нахождении управления, обеспечивающего значительное увеличение налогооблагаемой базы (не менее чем до 0,7), а сценарий Цель-2 нацелен на решение двухкритериальной задачи — обеспечить существенный (не менее 0,5) прирост налогооблагаемой базы при умеренном (не менее 0,3) приросте уровня жизни.

Основные результаты сведены в сводные таблицы 4 и 5 для НФЛ-18 и НФЛ-19 соответственно, в которых для наблюдаемых (т.е. основных) факторов указаны начальные состояния и приведены итоговые значения для каждого из четырех сценариев. Аналогичные сведения даны для суммарной оценки экономического состояния.

Таблица 4

Сводная таблица для НФЛ-18

Наблюдаемые факторы	Начальное состояние	Пасс-1	Акт-1	Цель-1	Цель-2
Прирост налогооблагаемой базы подоходного налога	0	– 0,76	1,71	2,45	2,45
Прирост ВВП	0,1	– 0,2	0,71	1,62	1,62
Рост доверия населения к государственной власти	– 0,5	– 0,34	0,58	0,21	0,21
Рост уровня жизни населения	– 0,5	– 0,5	0,12	0,75	0,75
Прирост уровня занятости	– 0,3	– 0,41	0,01	0,43	0,43
Инфляция	0,3	0,51	0,34	0,16	0,16
Суммарная оценка	– 0,14	– 0,33	0,48	0,67	0,67

Таблица 5

Сводная таблица для НФЛ-19

Наблюдаемые факторы	Начальное состояние	Пасс-1	Акт-1	Цель-1	Цель-2
Прирост налогооблагаемой базы подоходного налога	0	– 2,56	– 1,4	– 0,06	– 2,43
Прирост ВВП	0,1	– 0,91	– 0,72	-0,82	– 0,82
Рост уровня жизни населения	– 0,5	– 1,93	– 1,81	– 1,52	– 1,52
Прирост уровня занятости	– 0,3	– 1,6	– 1,21	– 0,69	– 0,69
Инфляция	0,3	0,6	0,9	1,1	1,1
Суммарная оценка	– 0,2	– 0,92	– 0,58	– 0,27	– 0,97

Рассмотрим сначала модель НФЛ-18, которая предполагает активное участие государственных органов в регулировании экономических факторов. Естественное (т.е. без вмешательства с помощью управляющих факторов) развитие ситуации описывается сценарием Пассивный-1, Ситуация ухудшается по всем факторам, кроме доверия населения к государственной власти. Налогооблагаемая база подоходного налога значительно убывает (от нулевого начального значения приходим к значению (– 0,76)). Рост ВВП (начальное значение 0,1) меняется на спад, хотя и слабый (– 0,2). Уровень жизни населения продолжает падать с той же скоростью (– 0,5). Падение занятости усиливается (с –0,3 до – 0,41). Инфляция растет (с 0,3 до 0,51). Хотя рост доверия населения к государственной власти несколько растет (с – 0,5 до – 0,34), но коэффициент остается отрицательным, так что более правильно сказать так, скорость нарастания отрицательного отношения населения к государственной власти несколько снизилась. Вполне естественно, что уменьшилась и изначально отрицательная суммарная оценка экономической ситуации — с (– 0,14) до (– 0,33). Общий вывод таков: если ничего не делать, то от плохой исходной ситуации страна придет к гораздо худшей.

Необходимы активные действия. Возможность резко изменить ситуацию к лучшему демонстрирует сценарий «Активный-1». В результате предложенных специалистами управляющих воздействий удается существенно улучшить почти все экономические показатели. Налогооблагаемая база подоходного налога очень сильно растет(от нулевого начального значения приходим к значению 1,71)). Валовой внутренний продукт значительно возрастает (от начального значения 0,1 до 0,71). Падение доверия населения к государственной власти (коэффициент – 0,5) сменяется заметным ростом (коэффициент 0,58, т.е. в целом коэффициент увеличился на 1,08). Падение уровня жизни населения (коэффициент – 0,5) сменилось слабым ростом (коэффициент 0,12). Падение занятости (коэффициент – 0,3) прекратилось (коэффициент 0,01). Единственный показатель, по которому ситуация несколько ухудшилась — это инфляция (рост с 0,3 до 0,34), однако это ухудшение весьма незначительно по сравнению с впечатляющим ростом по другим показателям. Вполне естественно, что резко выросла и изначально отрицательная суммарная оценка экономической ситуации — с (– 0,14) до 0,48. Итак, сценарий «Активный-1» демонстрирует большие возможности улучшения экономической ситуации вообще и налоговой ситуации в частности с помощью целенаправленных управляющих воздействий государственных органов.

Если в сценарии «Активный-1» система управляющих воздействий формировалась специалистами, оптимальность этой системы оставалась под вопросом, то в сценариях «Цель-1» и «Цель-2» система управляющих воздействий строилась с помощью компьютерной оптимизации в соответствии с заданными значениями целевых факторов. Поэтому вполне естественно, что по целевым факторам в результате оптимизации удалось еще больше улучшить ситуацию, чем в сценарии «Активный-1». При задании «получить коэффициент не менее 0,7» (сценарий «Цель-1») или «не менее 0,5» (сценарий «Цель-2») для налогооблагаемой базы подоходного налога удалось получить значение 2,45, заметно большее, чем в сценарии «Активный-1», т.е. 1,71, При задании «получить коэффициент не менее 0,3» для уровня жизни населения (сценарий «Цель-2») удалось получить его значительный рост с коэффициентом 0,75 (по сравнению с 0,12 в сценарии «Активный-1»). Если же сравнить итог с исходным коэффициентом (– 0,5), то общий рост уровня жизни — очень сильный, на 1,25. Все другие наблюдаемые факторы, кроме одного, также выросли больше, чем в сценарии «Активный-1». Наблюдаем очень сильный рост валового внутреннего продукта — до 1,62 (впечатляюще по сравнению с начальным значением 0,1 и соответствующим сценарию «Активный-1» значением 0,71). Падение занятости (коэффициент – 0,3) сменилось ее умеренным ростом (коэффициент 0,43). Инфляция упала вдвое (с 0,3 до 0,16). Единственный показатель, по которому результаты сценариев «Цель-1» и «Цель-2» уступают результатам сценария «Активный-1» — это рост доверия населения к государственной власти (значения коэффициента 0,21 и 0,58 соответственно при начальном значении (– 0,5)). (Уместно по этому поводу вспомнить утверждение о том, что наилучшей властью является та, действия которой незаметны для населения, все совершается как бы само собой.) Вполне естественно, что резко выросла и изначально отрицательная суммарная оценка экономической ситуации — с (– 0,14) до 0,67 (при 0,48 в сценарии «Активный-12). Примечательно, что оптимальные решения для сценариев «Цель-1» и «Цель-2» совпали. Следовательно, есть оптимальная траектория развития экономики. Поскольку при движении по ней с лихвой выполняются поставленные задания по целевым факторам, то компьютерная оптимизация дает одинаковые решения для двух сценариев.

Подведем итоги по модели НФЛ-18. Прогноз развития экономической ситуации неблагоприятен (сценарий «Пассивный-1»), поэтому необходимы управляющие воздействия. Они позволяют существенно улучшить ситуацию (сценарий «Активный-1»), особенно при оптимизации воздействий (сценарии «Цель-1» и «Цель-2»). Модель НФЛ-18, повторим, демонстрирует большие возможности улучшения экономической ситуации с помощью целенаправленных управляющих воздействий государственных органов.

Перейдем к рассмотрению модели НФЛ-19, основанной на использовании прежде всего экономических взаимодействий. Прогноз экономической ситуации здесь гораздо более неблагоприятен, чем в предыдущей модели. Согласно сценарию «Пассивный-1», налогооблагаемая база подоходного налога очень сильно сокращается (от исходного значения 0 до (– 2,56)), ВВП также очень сильно падает (от 0,1 до (– 0,91)). Еще более сильно падают уровень жизни населения (от (– 0,5) до (– 1,93)) и уровень занятости (от (– 0,3) до (– 1,6)). Вдвое растет инфляция (от 0,3 до 0,6). Естественно, что суммарная оценка экономической ситуации также резко падает (от (– 0,2) до (– 0,92)).

Очевидно, необходимы активные действия, направленные на улучшение ситуации. В сценарии «Активный-1» управляющими воздействиями являются существенное усиление борьбы государства с криминалом в экономике (на 0,5), существенное повышение таможенных сборов на импортную продукцию (на 0,6) и слабое снижение таможенных сборов на экспортную продукцию (на 0,2). В результате удалось добиться некоторого замедления ухудшения ситуации по всем показателям, кроме инфляции. Налогооблагаемая база подоходного налога по-прежнему очень сильно сокращается (от исходного значения 0 до (– 1,4), что лучше, чем при отсутствии воздействий (падение до (– 2,56))). ВВП снова очень сильно падает (от 0,1 до (– 0,72), что все-таки лучше, чем в сценарии «Пассивный-1», в котором падение достигло (– 0,91)). Чуть медленнее падают уровень жизни населения (от (– 0,5) до (– 1,81), а не до (– 1,93)) и уровень занятости (от (– 0,3) до (– 1,21), но не (– 1,6)). Однако инфляция растет втрое, а не вдвое (от 0,3 до 0,9, а не до 0,6). Суммарная оценка экономической ситуации, равная (– 0,58), показывает ее ухудшение по сравнению с исходным уровнем (– 0,2), хотя и не такое резкое, как при отсутствии воздействий (– 0,92).

Наилучшие результаты в модели НФЛ-19 получены при использовании сценария «Цель-1». Хотя целевого значения (0,7) для налогооблагаемой базы подоходного налога достичь не удалось, оказалось возможным сохранить ее практически на прежнем уровне (коэффициент (– 0,06)). Однако по сравнению с предыдущим сценарием несколько усилилось падение ВВП (до (– 0,82) по сравнению с (– 0,72)), в то время как несколько улучшилась ситуация с уровнем жизни населения (коэффициент (– 0,52) вместо (– 1,81)) и уровнем занятости (коэффициент (– 0,69), а не (– 1,21)). В то же время усилилась инфляция (коэффициент 1,1, а не 0,9). Суммарная оценка экономической ситуации, равная (– 0,27), является самой лучшей среди всех сценариев, но при этом показывает ухудшение экономической ситуации по сравнению с исходным уровнем (– 0,2).

Интересны результаты, полученные в сценарии «Цель-2». Ни по одному из целевых факторов (налогооблагаемая база подоходного налога и уровень жизни населения) не удалось достичь заданных значений (0,5 и 0,3 соответственно). Однако попытка «одним выстрелом убить двух зайцев» привела к экономическим результатам, которые являются наихудшими среди всех моделей. Налогооблагаемая база подоходного налога очень резко упала (до (– 2,43)). Остальные наблюдаемые факторы получили те же стационарные значения, что и в сценарии «Цель-1» (такова же картина по иным факторам — некоторые из них совпадают в этих двух сценариях, некоторые различаются). Суммарная оценка экономической ситуации, равная (–0,97), является самой худшей среди всех сценариев, хуже даже, чем при отсутствии каких-либо воздействий (коэффициент (– 0,92) в сценарии «Пассивный-1»).

Подведем итоги по модели НФЛ-19. Оказалось, что чисто экономическими методами невозможно добиться поставленных целей, улучшить экономическую ситуацию. Максимум, чего можно достичь — это не позволить ей слишком сильно ухудшиться, удержать почти на исходном уровне налогооблагаемую базу подоходного налога и суммарную оценку экономической ситуации (как в сценарии «Цель-1»).

Полученные с помощью моделей НФЛ-18 и НФЛ-19 результаты свидетельствуют о необходимости активного вмешательства государственных органов в экономические процессы, о невозможности улучшения ситуации чисто экономическими средствами. Этот вывод полностью соответствует концепции пяти нобелевских лауреатов по экономике (США) и отечественных академиков РАН о необходимости государственного регулирования экономики [1, 2].

Модели временных рядов НФЛ-18 и НФЛ-19, построенные на основе взвешенных ориентированных графов влияния факторов, допускают разнообразные варианты сценариев. За счет выбора тех или иных начальных значений факторов, наборов управляющих и целевых факторов. А также модификаций самих моделей путем модернизаций наборов факторов, коэффициентов их взаимовлияния, и др.

9.3. Моделирование и анализ многомерных временных рядов

Рассмотрим методы моделирования и анализа многомерных временных рядов, используемых для изучения реальных процессов взаимовлияния факторов на основе подхода ЖОК, описанного в предыдущем разделе.

Основные сведения о системе ЖОК. Компьютерная система ЖОК — это система поддержки анализа и управления в сложных ситуациях[3], описываемых многомерными временными рядами. Она предназначена для структуризации и анализа сложных, трудно формализуемых, слабо структурированных задач различной природы (экономической, управленческой, прогностической, технической, медицинской, социально-политической, экологической и пр.). Она применяется для построения моделей ситуаций на основе описания влияний факторов. Это делается с помощью ориентированных графов и использования оценок экспертов с последующим определением наиболее эффективных управленческих решений. Компьютерная система ЖОК:

- поддерживает аналитическое обоснование подходов к решению исследуемых проблем;

- позволяет спрогнозировать развитие моделируемой реальной системы; оценить результаты целенаправленного изменения тех или иных факторов;

---

[3] Использованы разработки В.Н. Жихарева, выполненные в Институте высоких статистических технологий и эконометрики.

- дает возможность выработать условия для целенаправленного поведения в исследуемой ситуации;

- обеспечивает возможность решения прямых и обратных задач управления.

Для построения модели изучаемого явления или процесса компьютерная система ЖОК предусматривает выделение основных факторов, описывающих реальную ситуацию, и установление непосредственных взаимосвязей между факторами в виде построения ориентированного взвешенного графа. Опосредованные взаимовлияния и итоговое стационарное состояние рассчитываются по описанным ниже алгоритмам. Система позволяет анализировать три основных типа сценариев:

- сценарий «Прогноз», позволяющий проследить «естественное» развитие моделируемой системы при отсутствии активных воздействий;

- сценарий типа «Активный», при котором работающий с системой специалист изменяет значения тех или иных параметров и анализирует получающуюся динамику и итоговое состояние (например, с целью ручного поиска рационального управления);

- сценарий типа «Цель», когда компьютерная система по заданной цели управления (например, значения определенных параметров должны быть не менее заданных) находит оптимальные воздействия путем решения соответствующей задачи оптимизации. В частности, проводит анализ принципиальной достижимости указанной цели из текущего состояния с использованием выбранных мероприятий (управлений).

Ядром компьютерной системы ЖОК является описанная ниже математическая модель. Преобразование задач анализа реальных явлений и процессов к математической постановке, оценка адекватности реальности и ее модели, процесс выбора управлений, процесс сравнительного анализа различных ситуаций в целом, моделирования и последующей интерпретации результатов математического моделирования относится к области «ручного труда» специалиста в соответствующей области знания и полной автоматизации, как правило, не поддается.

Компьютерная система ЖОК обеспечивает расчет равновесного (стационарного) состояния, к которому будет стремиться система взаимовлияющих факторов, и всех промежуточных состояний на пути от начального состояния к равновесному. В систему включены три варианта расчетов:

- расчет равновесного состояния без управления (учитываются только начальные данные);

- расчет равновесного состояния с управлением импульсного типа (при t = 0). (В такой модели система интерпретирует импульсное управление, как поправку к начальным данным.);

- расчет величины управления по заданным значениям величины приращения целевых факторов.

Математические алгоритмы исследовательской системы ЖОК. Используются следующие обозначения:

n — количество вершин в ориентированном графе *G* модели, т.е. число используемых в модели факторов;

$\mathbf{D} = \lfloor d_{i,j} \rfloor_{n \times n}$ — матрица порядка *n*Чn непосредственных влияний факторов (матрица смежности графа *G*);

$\mathbf{D}^T = \mathbf{A} = \left[a_{i,j} \right]_{n \times n}$ — матрица, транспонированная к матрице **D** (называемая матрицей непосредственных контрвлияний факторов);

t — время, принимающее дискретные значения 0, 1, 2, 3, …

$\mathbf{V}(t) = (V_1(t), V_2(t), \ldots, V_n(t))^T$, t = 0, 1, 2, 3, …, — вектор изменений (приращений, дифференциалов) факторов в момент дискретного времени t;

$\mathbf{W}(t) = \Delta\mathbf{V}(t) = \mathbf{V}(t) - \mathbf{V}(t-1)$, *t* = 0, 1, 2, 3, …, — вектор дифференциалов факторов второго порядка в момент дискретного времени *t*;

$\mathbf{V}_{уст} = \mathbf{V}(\infty) = (V_1(\infty), V_2(\infty), \ldots, V_n(\infty))^T$ — вектор, который обозначает величины предельных стационарных изменений (дифференциалов) факторов при безграничном росте *t*. Очевидно, что если $\mathbf{V}(\infty)$ существует, то:

$$\mathbf{W}(\infty) = \Delta\mathbf{V}(\infty) = \lim_{t \to \infty} (\mathbf{V}(t) - \mathbf{V}(t-1)) = 0;$$

$\mathbf{g}(t) = (g_1(t), g_2(t), \ldots, g_n(t))^T$ — вектор, обозначающий внешние управляющие воздействия, подаваемые на фактор V_i.в момент *t;*

$\mathbf{s} = (s_1, s_2, \ldots, s_n)^T$ — вектор, который обозначает сравнительную важность факторов V_i, задаваемую экспертным путем;

$u = (u_1, u_2, \ldots, u_n)^T$ — вектор, обозначающий отношение составителя модели к направлению изменения величин факторов V_i (+ 1 — рост значения фактора оценивается положительно, (– 1) — отрицательно, 0 — нейтрально);

E — единичная *n* · *n* матрица (на главной диагонали стоят 1, на остальных позициях — 0);

C — прореженная единичная *n* · n матрица, в которой единицы стоят на диагонали только на тех позициях, которые соответствуют целевым факторам. Очевидно, что **C** является проектором на координатную плоскость целевых факторов, и следовательно, $\mathbf{C}^2 = \mathbf{C}$, матрица **C** является псевдообратной к матрице **C**;

B — прореженная единичная n · n матрица, в которой единицы стоят на диагонали только на тех позициях, которые соответствуют управляющим факторам. Очевидно, что **B** является проектором на координатную плоскость управляющих факторов, и, следовательно **B**^{2}=**B**, матрица **B** является псевдообратной к матрице **B**;

$\mathbf{Q} = (\mathbf{E} - k_{cm}\mathbf{A})^{-1}$ — резольвента, где k_{cm} — множитель-стабилизатор, который используется в целях обеспечения достаточно устойчивой и быстрой сходимости итерационного процесса приближенного вычисления матрицы резольвентного оператора:

$$\mathbf{Q} = (\mathbf{E} - k_{cm}\mathbf{A})^{-1} = \sum_{m=0}^{\infty}(k_{cm}\mathbf{A})^m \approx \sum_{m=0}^{p}(k_{cm}\mathbf{A})^m,$$

где p достаточно велико. Полагают $k_{cm} = 1$ в том случае, если собственные числа матрицы $\mathbf{A}$ достаточно малы (обычно принимается, что $k_{cm}\mathbf{A}$ должна иметь собственные числа не только меньше единицы, но и меньше 0,9). Поскольку стабилизатор k_{cm} имеет лишь внутриматематический смысл и не используется при построении модели и интерпретации результатов расчетов, то в дальнейшем его не будем упоминать, предполагая по умолчанию $k_{cm} = 1$.

Система уравнений в математико-статистической модели. Для описания динамики факторов в компьютерной системе ЖОК используется математико-статистическая модель в виде системы линейных конечноразностных рекуррентных уравнений на трехточечном шаблоне $\{t - 1, t, t + 1\}$ следующего вида:

$$\begin{aligned} V_i(t+1) = V_i(t) + \sum_{j=1}^{n} a_{i,j}\big(V_j(t) - V_j(t-1)\big) + g_i(t) = \\ V_i(t) + \sum_{j=1}^{n} d_{j,i}\big(V_j(t) - V_j(t-1)\big) + g_i(t) \end{aligned}, \tag{1}$$

с начальными условиями:

$$V_i(0) = V_i^0, \tag{2}$$

где $i = 1, 2, \dots, n$, $t = 0, 1, 2, \dots$

Для рекуррентного уравнения на трехточечном шаблоне необходимо задать начальные условия при t = 0 ($V_i(0)=V_i^0$) и t = 1 ($V_i(1)=V_i^1$). Следовательно, первым уравнением цепочки рекуррентных уравнений (1) будет уравнение при t = 1,

При t = 1 уравнение полагается определенным и имеет вид:

$$V_i(2)=V_i(1)+\sum_{j=1}^{n}a_{i,j}\left(V_j(1)-V_j(0)\right)+g_i(1)$$

Для t = 0 уравнение определяется посредством соотношения

$$V_i(-1)=0 \qquad (3),$$

и тогда недостающие начальные данные $V_i(1)=V_i^1$ вычисляются из уравнения

$$\begin{gathered}V_i(1)=V_i(0)+\sum_{j=1}^{n}a_{i,j}\left(V_j(0)-V_j(-1)\right)+g_i(0)= \\ V_i(0)+\sum_{j=1}^{n}a_{i,j}V_j(0)+g_i(0)\end{gathered} \qquad . \qquad (4)$$

Заметим, что доопределение начальных данных $V_i(-1)=0$ нулем — всего лишь один из способов. В частности, если положить $V_i(0)=V_i(1)=V_i^0$, то результаты вычислений будут другими.

Из уравнений (1) видно, что используемая модель предполагает, что за один шаг дискретного времени (Δt = 1) происходит распространение влияния факторов-аргументов только на непосредственно от них зависящие факторы-функции. Времени можно придать содержательный смысл, если за шаг принять реальный интервал времени, необходимый для осуществления непосредственного влияния одного фактора на другой. Этот интервал может быть оценен экспертно, в ряде случаев его можно принять равным кварталу.

Уравнение (1)–(2) в векторной форме имеет вид:

$$\mathbf{V}(t+1)=\mathbf{V}(t)+\mathbf{A}\cdot\left(\mathbf{V}(t)-\mathbf{V}(t-1)\right)+\mathbf{g}(t), \qquad (5)$$

$$\mathbf{V}(-1)=0,\mathbf{V}(0)=\mathbf{V}_0, \qquad (6)$$

где t = 0, 1, 2, ... Решение задачи (5)–(6) определяются формулой:

$$\mathbf{V}(t)=\mathbf{V}(0)+\left(\sum_{k=0}^{t}\mathbf{A}^k\right)\cdot\mathbf{V}(0)+\sum_{k=0}^{t-1}\sum_{m=0}^{t-1-k}\mathbf{A}^m\cdot\mathbf{B}\cdot\mathbf{g}(k). \qquad (7)$$

Стационарное состояние и начальные условия. Стационарное состояние $\mathbf{V}(\infty)$ вычисляется приближенно при $t\to\infty$. Для практических расчетов достаточно принять, что $t\le\min(n,25)$.

Векторное уравнение (5) может быть представлено в виде уравнения для дифференциалов второго порядка:

$$\mathbf{W}(t+1)=\mathbf{A}\cdot\mathbf{W}(t)+\mathbf{g}(t), \qquad (8)$$

$$\mathbf{W}(0)=\Delta\mathbf{V}(0)=\mathbf{V}(0)-\mathbf{V}(-1)=\mathbf{V}(0), \qquad (9)$$

где t = 0, 1, 2, ... Решение уравнения (8)–(9) имеет вид:

$$\mathbf{W}(t)=\left(\sum_{k=0}^{t}\mathbf{A}^k\right)\cdot\mathbf{W}(0)+\sum_{k=0}^{t-1}\sum_{m=0}^{t-1-k}\mathbf{A}^m\cdot\mathbf{B}\cdot\mathbf{g}(k). \qquad (10)$$

Если просуммировать уравнения (8) при t = 0, 1, 2, … , то получим (при условии сходимости):

$$\mathbf{V}(\infty)-\mathbf{V}(0)=\mathbf{A}\cdot(\mathbf{V}(\infty)-\mathbf{V}(-1))+(\mathbf{g}(0)+\mathbf{g}(1)+\mathbf{g}(2)+\ldots), \qquad (11)$$

откуда следует?

$$\mathbf{V}(\infty)=(\mathbf{E}-\mathbf{A})^{-1}\circ(\mathbf{V}(0)+\mathbf{g}(0)+\mathbf{g}(1)+\mathbf{g}(2)+\ldots). \qquad (12)$$

Если же просуммировать уравнения (8) при t = 1, 2, … , то получим (при условии сходимости):

$$\mathbf{V}(\infty)-\mathbf{V}(1)=\mathbf{A}\circ(\mathbf{V}(\infty)-\mathbf{V}(0))+(\mathbf{g}(1)+\mathbf{g}(2)+\mathbf{g}(3)+\ldots), \qquad (13)$$

и соответственно:

$$\mathbf{V}(\infty) = \mathbf{V}(0) + (\mathbf{E} - \mathbf{A})^{-1} \cdot ((\mathbf{V}(1) - \mathbf{V}\,(0)) + (\mathbf{g}(1) + \mathbf{g}(2) + \mathbf{g}(3) + \ldots), \quad (14)$$

откуда видно, что при выборе начальных условий вида $V_i(0) = V_i(1) = V_i^0$ результат (14) отличается от (12).

В частности, при выборе режима прогноза развития ситуации без управления $\mathbf{g}(1) = \mathbf{g}(2) = \mathbf{g}(3) = \ldots = 0$ и выборе начальных условий $V_i(0) = V_i(1) = V_i^0$, которые выражают равенство нулю вторых производных от величин факторов при t = 0, из формулы (14) получим $\mathbf{V}(\infty) = \mathbf{V}(0)$. Это означает, что никакого развития ситуации не происходит. Она продолжает двигаться «равномерно и прямолинейно», поскольку вторые дифференциалы факторов равны нулю и первые дифференциалы факторов не изменяются во времени.

С другой стороны, формула (12) предполагает, что начальные данные оказывают такое же ударное воздействие в момент t = 0, как и внешнее импульсное при t = 0 управление, играющее роль (и имеющее «размерность») «механической силы».

Если предполагается использование только импульсных управляющих воздействий $\mathbf{g}(0) \neq 0$ при t = 0 и в дальнейшем $\mathbf{g}(1) = \mathbf{g}(2) = \mathbf{g}(3) = \ldots = 0$, то задача развития ситуации без управления и с управлением не отличаются друг от друга, поскольку управление играет роль поправки к начальным данным, и обратно, начальные данные выполняют роль поправки к управлению.

Режим поиска управления по целевым значениям факторов. Проекция стационарного решения (12) уравнения (8)–(9) на координатную плоскость целевых факторов может быть представлено в виде:

$$\mathbf{Y}_{уст} = \mathbf{Y}(\mathbf{V}(0)) + \mathbf{Y}(\mathbf{g}(0)),$$

где:

$$\mathbf{Y}(\mathbf{V}(0)) = \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{V}(0),\ \mathbf{Y}(\mathbf{g}(0)) = \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B} \cdot \mathbf{g}(0),$$

или иначе:

$$\mathbf{Y}_{\text{уст}} = \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{V}(0) + \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B} \cdot \mathbf{g}(0) = \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{V}(0) + \mathbf{B} \cdot \mathbf{g}(0). \quad (15)$$

Пусть $\mathbf{Y}^*{}_{уст}$ — вектор значений дифференциалов целевых факторов, тогда импульсное управление $\mathbf{g}(0)$ определяется по формуле:

$$\mathbf{g}(0) = (\mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B})^+ \cdot \left(\mathbf{Y}^*{}_{уст} - \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{V}(0)\right), \qquad (16),$$

где «+» обозначает операцию псевдоинверсии, и матрица $(\mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B})^+$ является псевдообратной к матрице $\mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B}$;

$\mathbf{g}^*(0) = L_1(\mathbf{g}(0))$ является результатом применения к вектору $\mathbf{g}(0)$ операции L_1 — ограничения числовых значений компонент вектора $\mathbf{g}(0)$ величинами $+1$ и -1, если эти значения выходят за пределы отрезка [–1; +1];

$\mathbf{g}^{**}(0) = Extr_1\left(\mathbf{g}^*(0)\right)$ получается из $\mathbf{g}^*(0)$ применением операции $Extr_1$ — замены числовых значений $\mathbf{g}^*(0)$ ближайшими к ним экстремальными на отрезке [– 1; + 1] величинами + 1 или – 1 соответственно.

Тогда стационарные решения, получаемые с использованием этих управлений, вычисляются по формулам:

$$\mathbf{Y}^{**}{}_{уст} = \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{V}(0) + \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B} \cdot \mathbf{g}^*(0),$$
$$\mathbf{Y}^{***}{}_{уст} = \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{V}(0) + \mathbf{C} \cdot \mathbf{Q} \cdot \mathbf{B} \cdot \mathbf{g}^{**}(0).$$

Степени матрицы смежности графа G и опосредованные взаимовлияния факторов. Пусть вершина $x1$ влияет на вершину $x2$ с силой 0,5, вершина $x2$ влияет на $x4$ с силой 0,6, вершина $x1$ влияет на $x3$ с силой 0,8, вершина $x3$ влияет на $x4$ с силой 0,4. Тогда опосредованное суммарное влияние $x1$ на $x4$ имеет силу:

$$0{,}5 \cdot 0{,}6 + 0{,}8 \cdot 0{,}4 = 0{,}62,$$

что равно сумме весов двух путей $x1 \to x2 \to x4$ и $x1 \to x3 \to$ x4 из $x1$ в $x4$, веса которых равны соответственно 0,5 · 0,6 = 0,3 и 0,8 · 0,4 = 0,32. Суммарная сила влияния одного фактора на другой равна сумме весов всех маршрутов в ориентированном графе G, ведущих из одного фактора в другой. Вес пути (маршрута) определяется как произведение весов дуг, составляющих этот путь (маршрут).

Если рассмотреть степени матрицы $\mathbf{D} = \lfloor d_{i,j} \rfloor_{n \times n}$, то их элементам можно придать вполне определенный смысл. Так, например, элемент матрицы $\mathbf{D}^2$ с координатами (1,2) равен сумме весов всех маршрутов из $x1$ в $x2$, содержащих ровно две дуги, а в $\mathbf{D}^3$ сумме весов всех маршрутов из $x1$ в $x2$, содержащих ровно три дуги и т.д. Таким образом, матрица $\sum_{m=0}^{\infty}\mathbf{D}^m$ выражает суммарные опосредованные влияния факторов друг на друга с учетом рефлексивного (при m = 0) непосредственного влияния фактора на самое себя с силой + 1, а матрица $\sum_{m=1}^{\infty}\mathbf{D}^m$ не учитывает рефлексивного непосредственного влияния.

Матрица $\mathbf{Q} = (\mathbf{E} - \mathbf{A})^{-1} = \sum_{m=0}^{\infty}\mathbf{A}^m$ является матрицей контрвлияний факторов с учетом рефлексивности, а матрица

$$\mathbf{Q} - \mathbf{E} = -\mathbf{E} + \sum_{m=0}^{\infty}\mathbf{A}^m = \sum_{m=1}^{\infty}\mathbf{A}^m = \mathbf{A} \cdot \sum_{m=0}^{\infty}\mathbf{A}^m = \mathbf{A} \cdot \mathbf{Q} = \mathbf{A} \cdot (\mathbf{E} - \mathbf{A})^{-1}$$

— матрицей контрвлияний факторов без учета рефлексивности.

Отдельный интерес представляет собой матрица $\mathrm{sign}\left(\sum_{m=0}^{\infty}\mathbf{D}^m\right)$ знаков элементов матрицы $\sum_{m=0}^{\infty}\mathbf{D}^m$, т.е. матрица направленности интегральных влияний фактора на фактор (или контрвлияний, если рассмотреть матрицу $\mathrm{sign}\left(\sum_{m=0}^{\infty}\mathbf{A}^m\right)$).

9.4. Балансовые соотношения в системе ЖОК

В настоящем разделе анализируется соотношения между временными рядами основных макроэкономических характеристик на основе балансовых соотношений. За основу взята известная схема Макконнелла и Брю [3] с данными по США за 1988 г. Анализ проведен по методологии ЖОК. Он позволил вскрыть два балансовых нарушения в схеме Макконнелла и Брю для блоков «государство» и «бизнес» и предложить способ восстановления корректности схемы путем введения двух дополнительных блоков. Установлены отличия схемы Макконнелла и Брю от используемых в методологии ЖОК, в частности, наличие двух противоположно направленных финансовых потоков, соединяющих два блока. После устранения (суммирования) подобных потоков на основе количественных данных схемы Макконнелла и Брю построена модель «Расчет ВНП» в соответствии с методологией ЖОК.

Далее кратко рассмотрены возможные перспективные направления математико-компьютерного развития системы ЖОК. В частности, применение аппарата нечетких множеств, осуществление синтеза с количественными эконометрическими моделями, организация автоматизированного изучения устойчивости выводов по отношению к малым отклонениям начальных данных и матрицы взаимовлияний.

Модели балансового типа в системе ЖОК. Методология компьютерного моделирования взаимовлияния факторов ЖОК ориентирована на использование экспертных оценок непосредственных влияний факторов. Однако она может быть успешно применена и для построения и анализа моделей балансового типа, в которых непосредственные влияния факторов описываются количественным образом.

В качестве примера возьмем классическую систему макроэкономических характеристик (в соответствии с классической монографией К.Р. Макконнелла и С.Л. Брю [3, с. 136–145]). Основная схема на с. 144 этой книги не может быть непосредственно использована, поскольку в ней два блока могут быть связаны финансовыми потоками, идущими в противоположных направлениях (например, государство изымает из личных доходов индивидуальные налоги, но при этом направляет в личные доходы трансфертные платежи). Кроме того, часть информации на указанной схеме приведена не при описании макроэкономических характеристик, а при рассмотрении финансовых потоков.

Поэтому классическая схема была переделана в соответствии с методологией ЖОК. Результат представлен на схеме 3 ниже. При этом выявилась необходимость дополнить систему характеристик Макконнелла и Брю двумя новыми, обеспечивающими выполнение балансовых соотношений для блоков «государство» и «бизнес». Все характеристики и блоки схемы 3 снабжены в качестве примеров численными значениями, соответствующими хозяйственной деятельности США за 1988 г. [3, с. 136–145].

Однако в схеме 3 имеется ряд блоков, имеющих один вход и один выход. Упростим схему, «убрав» такие блоки. В результате получена схема 4. Она уже не может удовлетворять балансовым соотношениям (сумма входов равна сумме выходов для каждого блока), поскольку при ее построении было проведено объединение всех финансовых потоков, непосредственно соединяющих те или иные два блока. Схема 4 дополнена относительными величинами, показывающими доли финансовых потоков, направленных из одного блока в другой. Их можно рассматривать как коэффициенты непосредственного влияния в исходной методологии ЖОК. Однако поскольку они получены из количественных значений, приводим два знака после запятой, а не один, как в исходных моделях типа ЖОК.

Основные макроэкономические характеристики. В табл. 6 указаны макроэкономические характеристики, используемые в [3] и схемах 3 и 4. Содержание основных макроэкономических величин раскрывается ниже через соотношения между ними.

Таблица 6

Основные макроэкономические характеристики

№ п/п	Характеристика
1.	Государство (как хозяйствующий субъект)
2.	Валовой национальный продукт
3.	Чистый национальный продукт
4.	Национальный доход
5.	Чистый доход домохозяйств (после уплаты налогов)
6.	Чистый экспорт (экспорт минус импорт)
7.	Бизнес (как объединение хозяйствующих субъектов)
8.	Другие источники финансирования государства (дополнительная характеристика по сравнению со схемой Макконнелла и Брю)
9.	Накопления бизнеса (дополнительная характеристика по сравнению со схемой Макконнелла и Брю)
10.	Государственные закупки товаров и услуг
11.	Трансфертные платежи
12.	Инвестиции бизнеса
13.	Амортизационные отчисления
14.	Нераспределенные доходы корпораций
15.	Косвенные налоги на бизнес
16.	Налоги на прибыли корпораций
17.	Взносы на социальное страхование
18.	Индивидуальные налоги
19.	Личный доход: заработная плата, рента, процент, дивиденды
20.	Личные сбережения
21.	Личные потребительские расходы

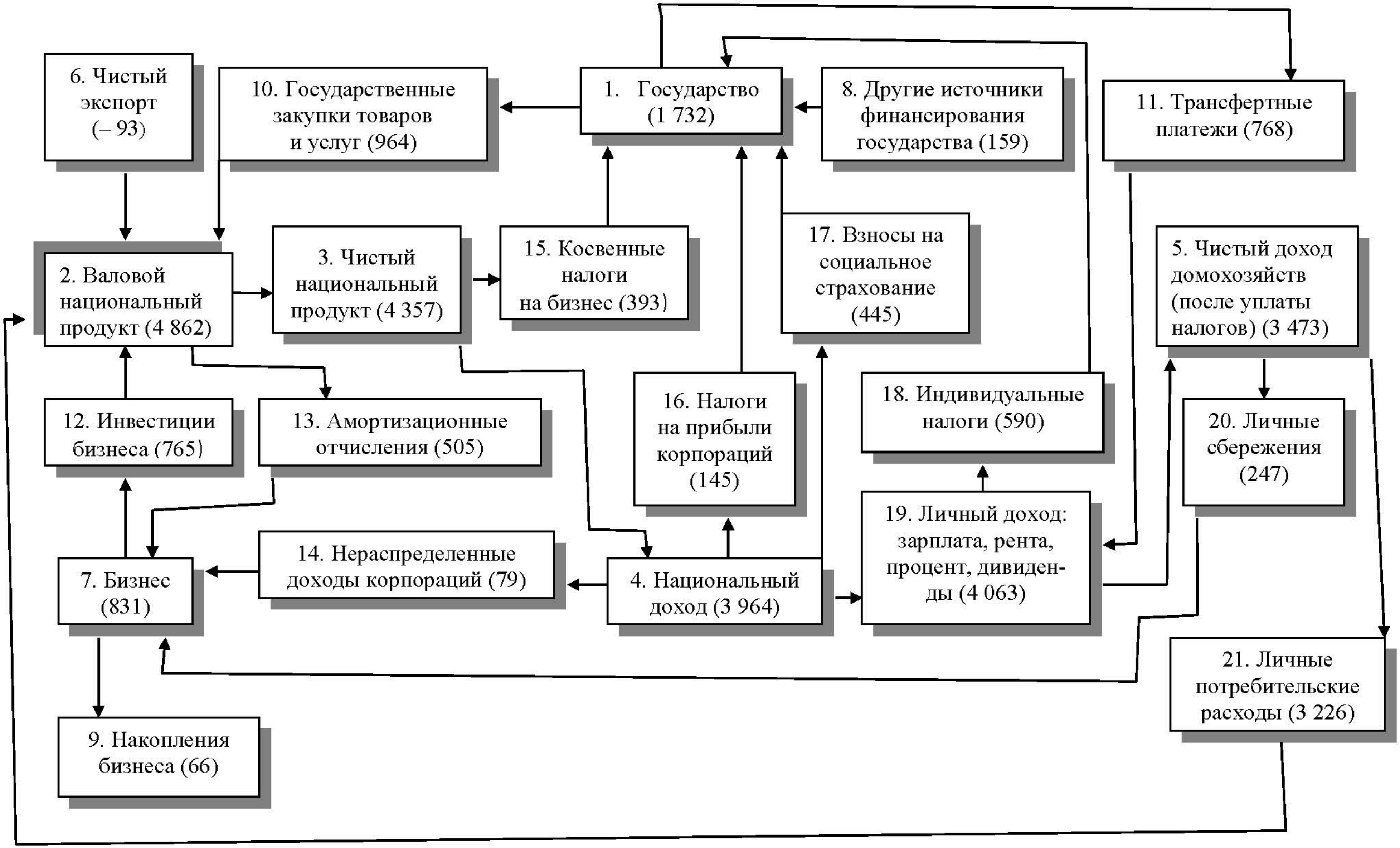


Схема 3. Соотношения основных макроэкономических величин (США, 1988 г., млрд долл.)

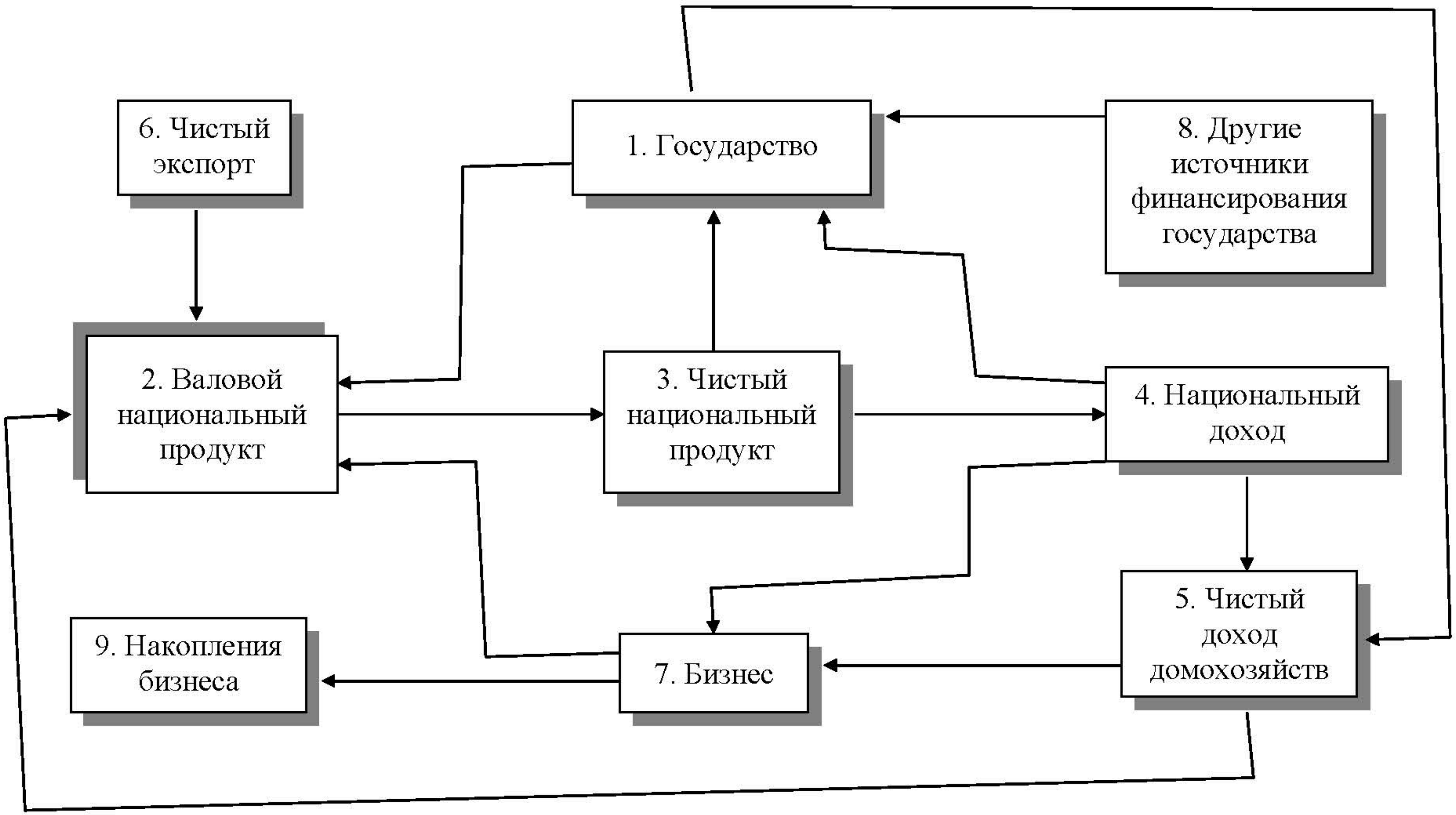


Схема 4. Взаимовлияния основных макроэкономических характеристик

Соотношения основных макроэкономических величин. Раскроем содержание и приведем количественные данные для основных макроэкономических характеристик.

В 1988 г. валовой национальный продукт (ВНП) США составлял 4 862 млрд долл. В дальнейшем указания на единицу измерения (миллиард долларов) будем опускать. Основные хозяйствующие субъекты — это государство, бизнес и домохозяйства (семьи, конечные потребители).

Будем выписывать соотношения между макроэкономическими характеристиками, а строкой ниже — соответствующие балансовые соотношения (в миллиардах долларов).

Как известно,

ВНП = (чистый национальный продукт) + (амортизационные отчисления),

$$4\,862 = 4\,357 + 505.$$

Известно, что ВНП отражает не только результат работы страны за год – примерно на 10 % он состоит из результатов труда предыдущих лет, перенесенных на продукты и услуги, выпущенные в данном году. Прошлый труд учитывается с помощью амортизационных отчислений.

По определению,

(чистый национальный продукт) = (национальный доход) + (косвенные налоги на бизнес),

в количественном отношении:

$$4\,357 = 3\,964 + 393.$$

Косвенные налоги на бизнес — это прежде всего акцизы на алкогольную и табачную продукцию, бензин, драгоценности, дорогие автомашины и т.п. Государство собирает косвенные налоги на бизнес, ничего не предоставляя взамен (бремя налогов делят между собой потребители и производители).

Поскольку составляющей (косвенные налоги на бизнес) не соответствует производство реальных товаров и услуг, вполне естественно ее исключить. *Наиболее естественной характеристикой результатов работы страны за год является национальный доход, составляющий лишь 81,5 % от ВНП.*

По определению,

(национальный доход) = (личный доход (без трансфертов): зарплата, рента, процент, дивиденды) +
+ (взносы на социальное страхование) + (налоги на прибыли корпораций) +
+ (нераспределенные доходы корпораций),

$$3\ 964 = 3\ 295 + 445 + 145 + 79.$$

К личному доходу добавляются трансфертные платежи (пенсии, пособия и др.) со стороны государства в размере 768 млрд долл. В рассматриваемом месте схемы наблюдаем круговое замыкание финансовых потоков — часть национального дохода идет на налоги, поступающие государству, а часть — в личный доход граждан, куда также поступают трансферты от государства:

(личный доход: зарплата, рента, процент, дивиденды) =
= (личный доход (без трансфертов): зарплата, рента, процент, дивиденды) +
+ (трансфертные платежи),

$$4\ 063 = 3\ 295 + 768.$$

С точки зрения сбалансированности бюджета государства настораживает тот факт, что национальный доход (3 064) оказывается меньше личного дохода (4 063), поскольку часть национального дохода — это нераспределенные доходы корпораций (79), которые никак нельзя отнести к личному доходу.

Примечание. К трансфертным платежам относятся

(1) выплаты по страхованию по старости и от несчастных случаев, пособия по безработице, основанные на социальных программах;

(2) выплаты по вспомоществованию;

(3) разнообразные выплаты ветеранам, например, субсидии на образование и пособия по нетрудоспособности;

(4) выплаты частных пенсий и пособий по безработице и вспомоществованию;

(5) процентные платежи, выплачиваемые правительством и потребителями [3, с. 143].

Ясно, что:

(личный доход: зарплата, рента, процент, дивиденды) =
= (индивидуальные налоги) + (чистый доход домохозяйств),

$$4063 = 590 + 3473.$$

Следовательно, средняя ставка индивидуальных налогов составляет 14,5 %.

Очевидно,

(чистый доход домохозяйств) = (личные сбережения) + (личные потребительские расходы),

$$3\ 473 = 247 + 3\ 226.$$

В кейнсианской макроэкономической теории большое значение имеет склонность к сбережению. Для США 1988 г. этот коэффициент равен 7,1 %.

Основное соотношение для ВНП имеет вид:

ВНП = (личные потребительские расходы) + (государственные закупки товаров и услуг) + + (инвестиции бизнеса) + (чистый экспорт),

$$4\,862 = 3\,226 + 964 + 765 - 93.$$

Заслуживает анализа хозяйственная деятельность государства (второе слагаемое в последней формуле для ВНП) и бизнеса (третье слагаемое в той же формуле).

Хозяйственная деятельность государства и бизнеса. Как видно из схемы 3, доходы государства складываются из следующих источников:

(доходы государства) = (косвенные налоги на бизнес) + (налоги на прибыли корпораций) + + (взносы на социальное страхование) + (индивидуальные налоги),

$$1\,573 = 393 + 145 + 445 + 590,$$

Согласно той же схеме, расходы государства таковы:

(расходы государства) = (государственные закупки товаров и услуг) + (трансфертные платежи),

$$1\,732 = 964 + 768.$$

Следовательно, расходы государства превышают его доходы на 159 млрд долл. Этот факт никак не разъясняется в [3].

Одно из возможных разъяснений может состоять в том, что отдельные разделы монографии Макконнелла и Брю не соответствуют друг другу, Так, проведенные выше рассуждения исходили из схемы на с. 144 этой монографии, на которой финансовый поток (трансфертные платежи) ведет от блока (государство) к блоку (личный доход). Однако на с. 143 той же книги подробно разъясняется понятие «трансфертные платежи» (это разъяснение процитировано выше), и из сказанного там ясно, что трансфертные платежи осуществляют не только государство.

Естественно выбрать другой путь снятия противоречия — ввести специальный блок (другие источники финансирования государства) с наполнением в 159 млрд долл. К «другим источникам» могут относиться доходы от государственных предприятий, от внешнеэкономической деятельности государства, от эмиссии наличности и ценных бумаг и т.д.

Перейдем к рассмотрению блока (бизнес). Согласно [3]:

(доходы бизнеса) = (амортизационные отчисления) +
+ (нераспределенные доходы корпораций) + (личные сбережения),

$$831 = 505 + 79 + 247,$$

в то время как

(расходы бизнеса) = (инвестиции бизнеса, включаемые в ВНП) = 765.

Итак, доходы бизнеса превышают его расходы на 66 млрд долл. Этот факт никак не разъясняется в [3]. Одно из объяснений может состоять в некоторой условности отнесения личных сбережений к доходам бизнеса, с помощью которых он осуществляет инвестиции. Обычно считается, что личные сбережения хранятся в банках, а банки из этих средств выдают кредиты бизнесу. Однако очевидно, что некоторая часть поступивших в банк средств должна использоваться для обеспечения бесперебойной работы банка (с помощью государственной резервной системы, т.е. храниться в качестве резерва в центральном банке, а не выдаваться бизнесу в кредит). Кроме того, часть кредитов может быть не связана с инвестициями (хотя согласно [3] строительство частных домов учитывается в блоке (инвестиции бизнеса), но кредиты частным лицам не сводятся к кредитованию строительства).

Пойдем другим путем снятия противоречия — введем специальный блок (накопления бизнеса) с наполнением в 66 млрд долл. Можно считать, что эти накопления выражены в денежной форме и в форме ценных бумаг, но не в виде материальных инвестиций, как в блоке (инвестиции бизнеса).

Выводы по результатам анализа данных, приведенных в книге Макконнелла и Брю [3], состоят в следующем.

1. Методология подхода, кратко именуемого ЖОК, позволяет адекватно анализировать схемы типа приведенной на с. 144 известной монографии Макконнелла и Брю [3].

2. Методология ЖОК позволяет выявлять «дыры» (нарушения балансовых соотношений) в подобных схемах.

3. Дальнейшие усовершенствования схем типа рассмотренной приводят к достаточно сложным построениям т.н. «национального счетоводства» [6]. Не

со всеми построениями «национального счетоводства» можно согласиться. Так, вряд ли можно научно обосновать утверждение о том, что банки создают такую большую долю ВНП, как 13 %.

4. Схемы типа приведенной на с. 144 известной монографии Макконнелла и Брю [3], основанные на количественных соотношениях, внешне представляются более обоснованными, чем схемы метода ЖОК. Однако при реальном использовании они не являются более полезными для решений стоящих в тематике ЖОК задач, чем схемы, основанные на экспертной оценке непосредственных влияний факторов. Связано это с достаточно большим произволом при формулировках определений численных значений.

5. На основе данных схемы, приведенной на с. 144 известной монографии Макконнелла и Брю [3], может быть построена модель «Расчет ВНП» типа ЖОК (см. ниже).

ВВП и ВНП. ВНП отличается от валового внутреннего (отечественного) продукта (ВВП) на величину доходов от экономической деятельности, полученных из-за границы, за вычетом аналогичных доходов, переданных другим странам [7, с. 55]. Другими словами, ВВП рассчитывается по всем предприятиям, действующим внутри рассматриваемой страны, независимо от доли иностранной собственности в тех или иных предприятиях. В отличие от ВВП, ВНП разделяет предприятия по стране принадлежности. Конкретно, для расчета ВНП надо к ВВП добавить доходы отечественных предприятий, полученные за рубежом, и вычесть доходы зарубежных структур, полученные на территории нашей страны. Понятны сложности, связанные с деятельностью транснациональных корпораций, с акционерными обществами (совместными предприятиями), объединяющими капитал из различных стран. Очевидно, расчет ВВП более прост и обоснован. Однако в учебниках по экономической теории укрепился рассказ о ВНП.

В странах, в которых подавляющая часть экономики принадлежит отечественным хозяйствующим субъектам, различие между ВНП и ВВП незначительно. Полагаем, что Россия относится к таким странам.

В российской экономической практике используется ВВП, а не ВНП. Например, согласно действующему законодательству, финансирование Вооруженных Сил, образования, науки и др. установлено в процентах от ВВП.

Модель ВНП. Модель ВНП построена на основе методологии ЖОК с помощью приведенной выше схемы 3. Она содержит 9 факторов (табл. 7), указанных в схеме 3 под номерами 1–9.

Таблица 7

Факторы модели «Расчет ВНП»

№ п/п	Фактор
1.	Государство (как хозяйствующий субъект)
2.	Валовой национальный продукт
3.	Чистый национальный продукт
4.	Национальный доход
5.	Чистый доход домохозяйств (после уплаты налогов)
6.	Чистый экспорт (экспорт минус импорт)
7.	Бизнес (как объединение хозяйствующих субъектов)
8.	Другие источники финансирования государства
9.	Накопления бизнеса

Взаимовлияния основных макроэкономических характеристик приведены на схеме 7.

Степень влияния одного фактора на другой оценивалась по количественным данным, приведенным выше. Например, если из ВНП (4 862 млрд долл.) в чистый национальный продукт переходит 4 357 млрд долл., то влияние оценивается величиной 4 357/4 862 = 0,90, Другим примером является взаимоотношение блоков (государство) и (чистый доход домохозяйств). С одной стороны, государство берет 590 млрд долл. индивидуальных налогов, с другой — выплачивает 768 млрд долл. трансфертных платежей. Эти финансовые потоки объединяются, и итог — 178 млрд долл. в пользу домохозяйств.

Модель ВНП содержит 14 связей. Они приведены в табл. 8. В каждой паре указана сила воздействия первого элемента на второй.

Таблица 8

Связи между факторами модели «Расчет ВНП»

№ п/п	Фактор
3-1	Величина чистого национального продукта влияет на государство (как хозяйствующий субъект) с коэффициентом влияния 0,09
4-1	Национальный доход влияет на государство (как хозяйствующий субъект) с коэффициентом влияния 0,15
8-1	Другие источники финансирования государства влияют на государство (как хозяйствующий субъект) с коэффициентом влияния 1,0
1-2	Государство (как хозяйствующий субъект) влияет на валовой национальный продукт с коэффициентом влияния 0,56

№ п/п	Фактор
5-2	Чистый доход домохозяйств (после уплаты налогов) влияет на валовой национальный продукт с коэффициентом влияния 0,93
6-2	Чистый экспорт (экспорт минус импорт) влияет на валовой национальный продукт с коэффициентом влияния (– 1.0)
7-2	Бизнес (как объединение хозяйствующих субъектов) влияет на валовой национальный продукт с коэффициентом влияния 0,31
2-3	Валовой национальный продукт влияет на чистый национальный продукт с коэффициентом влияния 0,90
3-4	Чистый национальный продукт влияет на национальный доход с коэффициентом влияния 0,91
1-5	Государство (как хозяйствующий субъект) влияет на чистый доход домохозяйств (после уплаты налогов) с коэффициентом влияния 0,10
4-5	Национальный доход влияет на чистый доход домохозяйств (после уплаты налогов) с коэффициентом влияния 0,68
4-7	Национальный доход влияет на бизнес (как объединение хозяйствующих субъектов) с коэффициентом влияния 0,02
5-7	Чистый доход домохозяйств (после уплаты налогов) влияет на бизнес (как объединение хозяйствующих субъектов) с коэффициентом влияния 0,07
7-9	Бизнес (как объединение хозяйствующих субъектов) влияет на накопления бизнеса с коэффициентом влияния 0,08

Перспективные направления математико-компьютерного развития системы ЖОК. Технология ЖОК и ее программное обеспечение могут быть развиты в различных направлениях с целью расширения их познавательных и прикладных возможностей.

Применение аппарата нечетких множеств. В процессе распространения влияний факторов переход от одного фактора к другому может происходить многими путями. Возникает два вопроса:

- как рассчитать итоговое влияние при движении по фиксированному пути, если заданы коэффициенты непосредственного влияния между соседними факторами на этом пути?

- как свести вместе результаты влияний по различным путям?

В описанном выше варианте системы ЖОК в ответ на первый вопрос заданные коэффициенты непосредственного влияния между соседними факторами перемножают, а результаты влияний по различным путям складывают.

Представляется целесообразным использовать альтернативный вариант на основе концепций теории нечетких множеств. Например, итоговое влияние при движении по фиксированному пути рассчитывается как минимум заданных коэффициентов непосредственного влияния между соседними факторами на этом пути, а для получения итогового результата влияния находится максимум из результатов влияний по различным путям. Имеется ряд аргументов против используемой ныне процедуры и за предлагаемую процедуру.

Синтез с количественными эконометрическими моделями. Соединение основанной на не вполне количественных экспертных оценках технологии ЖОК с достаточно развитой технологией количественных эконометрических уравнений, в частности, использование лагов, т.е. не непосредственных влияний, а влияний с задержкой, представляет несомненный теоретический и практический интерес. Можно сравнить методологию ЖОК с эконометрической моделью, в которой исходные данные и правила перехода задаются экспертами в шкале порядка. При движении в обсуждаемом направлении необходимо дистанцироваться от попыток использовать нормальные распределения, поскольку подобных распределений в реальной экономике не может быть в принципе. Другие параметрические модели, в частности, логарифмически нормальные распределения, могут иметь ограниченную область полезности, но наиболее обоснованным является непараметрический подход.

Автоматизированное изучение устойчивости выводов по отношению к малым отклонениям начальных данных и матрицы взаимовлияний. Поскольку точность всех исходных экспертных оценок очевидным образом мала, необходимо изучить выводы на устойчивость. Это может быть сделано «малой кровью», путем добавления в компьютерную систему ЖОК блока, который бы моделировал описанные выше малые отклонения и выдавал исследователи распределение (разброс) получаемых при этом выводов. С научной точки зрения этот подход — один из вариантов метода размножения выборок, кратко — бутстрепа (см. главу 16).

Система ЖОК получила название по первым буквам основных разработчиков (В.Н. Жихарев, А.И. Орлов, В.Г. Кольцов). Опыт практического применения этой системы описан в работах [8, 9]. Система ЖОК развивает идеи когнитивного подхода при решении слабоструктурированных задач, разработанного в Институте проблем управления РАН [10, 11], но на основе иного математического обеспечения.

Метод ЖОК может найти широкое применение для анализа экономического состояния и перспектив развития промышленных предприятий, банков, различных государственных и коммерческих структур. Представленные в настоящем подразделе материалы и результаты свидетельствуют о целесообразности дальнейшего развития рассматриваемой тематики с целью получения новых теоретических и практических результатов.

Подведем итоги главы. Рассмотрены методы анализа и моделирования временных рядов. Они используются прежде всего для прогнозирования технических, социально-экономических, медицинских и иных процессов. Надо отметить, что как самим временным рядам, так и вопросам их прогнозирования посвящена огромная литература. Наряду с вероятностно-статистическими методами при прогнозировании активно применяют экспертные методы (см. главу 11).

В настоящей главе рассмотрены лишь основы и отдельные вопросы статистики временных рядов. Эта область — одна из наиболее обширных и сложных (с математической точки зрения) в прикладной статистике. Читатель, желающий глубже познакомиться с такой специфической ветвью прикладной статистики, как статистика временных рядов, должен обратиться к литературе, в частности, указанной в конце главы.

Литература

1. Стратегия эффективного перехода и шоковые методы реформирования российской экономики / М. Интрилигейтор, Р. Макинтайр, Л. Тейлор, А. Эмсден // Шансы российской экономики / под редакцией Ю.М. Осипова, Е.С. Зотовой. — Москва : ТЕИС, 1997. — С. 168–195.

2. *Орлов, А.А.* Нобелевские лауреаты — за государственное регулирование экономики / А.А. Орлов, А.И. Орлов // Обозреватель — Observer. — 1998. — № 1. — С. 44–46. Перепечатано в книге: Современная политическая история России (1985–1998). Том 1. Хроника и аналитика. — Москва : Духовное наследие : РАУ-Корпорация, 1999. — С. 909–911.

3. *Макконнелл, К.Р.* Экономикс: Принципы, проблемы и политика : в 2 томах / К.Р. Макконнелл, С.Л. Брю ; перевод с английского 11-го издания. — Т. 1. — Москва : Республика, 1995. — 400 с.

4. Налоги / под редакцией Д.Г. Черника. — Москва : Финансы и статистика, 1998. — 688 с.

5. Социальная статистика / под редакцией И.И. Елисеевой. — Москва : Финансы и статистика, 1997. — 416 с.

6. Национальное счетоводство / под редакцией Г.Д. Кулагиной. — Москва : Финансы и статистика, 1997. — 448 с.

7. Статистический словарь. — Москва : Финансы и статистика, 1989. — 623 с.

8. *Жихарев, В.Н.* Эконометрический метод оценки результатов влияния / В.Н. Жихарев, В.Г. Кольцов, А.И. Орлов // Тезисы конференции «Организация производства на предприятиях в современных условиях». — Москва : Изд-во МГТУ им. Н.Э. Баумана, 1999. — С. 113–114.

9. *Орлов, А.И.* Новый эконометрический метод «ЖОК» оценки результатов взаимовлияний факторов в инженерном менеджменте / А.И. Орлов, В.Н. Жихарев, В.Г. Кольцов // Проблемы технологии, управления и экономики / под общей редакцией В.А. Панкова. — Ч. 1. — Краматорск : Донбасская государственная машиностроительная академия, 1999. — С. 87–89.

10. *Максимов, В.И.* Аналитические основы применения когнитивного подхода при решении слабоструктурированных задач / В.И. Максимов, Е.К. Корноушенко // Труды Института проблем управления РАН. — 1998. — № 2. — С. 95–109.

11. *Корноушенко, Е.К.* Управление процессами в слабоформализированных средах при стабилизации графовых моделей среды / Е.К. Корноушенко, В.И. Максимов // Труды Института проблем управления РАН. — 1998. — № 2. — С. 82–94.

12. *Микрюков, А.А.* Когнитивные технологии в процессах поддержки принятия решений в цифровой экономике // Инновации и инвестиции. — 2018. — № 6. — С. 127–131.

13. *Лойко, В.И.* Высокие статистические технологии и системно-когнитивное моделирование в экологии : монография / В.И. Лойко, Е.В. Луценко, А.И. Орлов. — Краснодар : КубГАУ, 2019. — 258 с.

Контрольные вопросы

1. Сколько факторов используют при построении моделей многомерных временных рядов с помощью системы ЖОК?

2. Сколько взаимосвязей между факторами используют при построении моделей с помощью системы ЖОК?

3. В чем состоит один шаг учета взаимовлияний факторов в системе ЖОК?

4. Какая эконометрическая модель лежит в основе системы ЖОК?

5. На чем основана сходимость к соответствующим пределам оценок факторов в системе ЖОК?

6. Чем сценарии типа «Активный» отличаются от сценариев типа «Цель»?

Темы докладов, рефератов, исследовательских работ

1. Различные виды экспертных оценок при построении моделей многомерных временных рядов с помощью системы ЖОК.

2. Роль оцифровки качественных оценок при использовании системы ЖОК.

3. Существование равновесных состояний при различных режимах использования системы ЖОК.

4. Система ЖОК как способ получения нового экономического знания (на примере двух типов моделей динамики налогооблагаемой базы подоходного налога НФЛ-18 и НФЛ-19).

5. Сравнение вариантов практического применения системы ЖОК в случае использования качественных и количественных признаков.

ГЛАВА 10. СТАТИСТИЧЕСКИЕ МОДЕЛИ УПРАВЛЕНИЯ КАЧЕСТВОМ

Одна из наиболее важных областей статистических методов, приносящая к тому же наибольший доход — это обеспечение качества, основанное на применении статистического моделирования. Статистическим методам управления качеством и сертификации и посвящена настоящая глава.

10.1. Основы статистического контроля качества

Сначала дадим общие сведения о месте статистических методов в принятии решений при управлении качеством и сертификации продукции. Затем рассмотрим статистический контроль качества и продемонстрируем его высокую экономическую эффективность.

Качество продукции в современных условиях. Слова «сертификация», «международные стандарты ИСО (т.е. разработанные *International Standardization Organization* — Международной организацией по стандартизации, сокращенно *ISO*, по-русски — ИСО) серии 9000 по системам качества» уже навязли в зубах. Менее осознано в нашей стране, что управление качеством — прежде всего применение современных методов статистического моделирования. На Западе (США) и на Востоке (Япония) это — аксиома. Вот типичное высказывание японского менеджера и инженера: «Методы статистики — именно то средство, которое необходимо изучить, чтобы внедрить управление качеством. Они — наиболее важная составная часть комплексной системы всеобщего управления качеством на фирме. В японских корпорациях все, начиная от председателя Совета Директоров и до рядового рабочего в цехе, обязаны знать хотя бы основы статистических методов». Так считает Каору Исикава, президент промышленного института Мусаси, заслуженный профессор Токийского университета [1, с. 15].

Раз все японские работники знают про статистические методы — значит, их научили в школе. Во всем мире — в США, Японии и Ботсване — школьники учат статистические методы как один из обязательных школьных предметов, вместе с физикой, химией, математикой и историей. ЮНЕСКО регулярно проводит конференции по преподаванию статистики в средней школе. И вот всем виден результат — качество компьютеров IBM и японских телевизоров. Справедливости ради надо отметить, что популярные ныне международные стандарты ИСО серии 9000 ничем принципиально не отличаются от давних документов КС УКП[4], а в некоторых отношениях КС УКП были более прогрессивными, чем нынешние стандарты ИСО 9000.

Очевидно, овладение основами статистического контроля качества продукции — неотъемлемая часть образования менеджера и инженера.

О сертификации. Сертификация — это официальная гарантия поставки производителем продукции, удовлетворяющей установленным требованиям. Поставщики и продавцы должны иметь сертификаты качества на предлагаемые ими товары и услуги. Маркетинг включает в себя работы по сертификации.

За новыми терминами зачастую скрываются хорошо известные понятия, несколько модернизированные в соответствии с современной обстановкой. Так, целесообразно связать комплексную систему управления качеством продукции с маркетингом: «маркетинг в широком смысле — это усовершенствованная, ориентированная на рыночную экономику КС УКП» [2, с. 61].

Есть несколько уровней сертификации. Говоря о сертификации продукции, могут иметь в виду качество конкретной ее партии. В ряде случаев это оправдано — рядового потребителя интересует качество лишь той единицы продукции, которую он сам приобрел. Однако установление долговременных хозяйственных связей целесообразно лишь в случае, когда поставщик гарантирует высокое качество не одной, а всех партий своей продукции. Очевидно, для этого должны быть проведены оценка и сертификация технологических процессов и производств, обеспечивающих выпуск этой продукции.

Еще больше повышается доверие к поставщику, если не только отдельные технологические процессы, но и все предприятие в целом гарантированно выпускает продукцию высокого качества. Это обеспечивается действующей на предприятии системой качества, удовлетворяющей требованиям Международной организации по стандартизации, выраженным в системе стандартов ИСО 9000.

В условиях рыночной экономики основная характеристика товара — его конкурентоспособность. Очевидно, производителю необходимо уметь оценивать конкурентоспособность перед запуском продукции в производство или началом работы по продвижению на зарубежный рынок. Одним из основных компонентов конкурентоспособности является технический уровень продукции. Фирма, обладающая патентом или новой научно-технической разработкой, имеет более высокий излишек производителя по сравнению с другими фирмами. При принятии решений о выборе направления инвестиционных вложений

[4] В 1970-е и 1980-е гг. в СССР активно разрабатывались КС УКП — Комплексные Системы Управления Качеством Продукции. Имелись областные варианты — горьковская, львовская, днепропетровская и иные системы качества.

одна из основных учитываемых характеристик — технический уровень продукции.

Из сказанного вытекает, что сертификация продукции — это современная форма управления качеством продукции. На Западе общепринято, что основная составляющая в управлении качеством продукции — это статистические методы (см., например, отчет Комитета ИСО по изучению принципов стандартизации [3]). В нашей стране внедрение комплексных систем управления качеством (КС УКП), надо признать, сводилось во многом к подготовке документации организационного характера. Статистические методы использовались в промышленности недостаточно, а государственные стандарты по этой тематике зачастую содержали грубейшие ошибки (см. ниже).

Подготовка предприятий к сертификации продукции, технологических процессов и производств, систем качества требует приложения труда квалифицированных специалистов, причем в достаточно большом объеме. Подобную работу обычно проводят специализированные организации.

О развитии статистических методов сертификации в России. Более 150 лет статистические методы применяются в России для проверки соответствия продукции установленным требованиям, т.е. для сертификации. Так, еще в 1846 г. действительный член Петербургской академии наук М.В. Остроградский рассматривал задачу статистического контроля партий мешков муки или штук сукна армейскими поставщиками. С тех пор в России в статистическом контроле качества было сделано многое, особенно в области теории. Так, монографии проф. Ю.К. Беляева и проф. Я.П. Лумельского можно смело назвать классическими. Было выпущено и большое количество практических руководств, в основном переводных.

С начала 1970-х гг. стали разрабатываться государственные стандарты по статистическим методам. В связи с обнаружением в них грубых ошибок 24 из 31 государственного стандарта по статистическим методам были отменены в 1986–87 гг. (перечень стандартов и описание ошибок приведены в работе [4]). К сожалению, потеряв правовую силу как нормативные документы, ошибочные стандарты продолжают использоваться как научно-технические издания.

В 1989 г. был организован Центр статистических методов и информатики (ЦСМИ) для работ по развитию и внедрению современных статистических методов. Уже к середине 1990 г. ЦСМИ были разработаны 7 диалоговых систем по современным статистическим методам управления качеством, а именно, СПК, АТСТАТ-ПРП, СТАТКОН, АВРОРА-РС, ЭКСПЛАН, ПАСЭК, НАДИС (описания этих систем даны в работе [5]).

Параллельно ЦСМИ вел работу по объединению статистиков. В апреле 1990 г. в Большом Актовом Зале Московского Энергетического института прошла Учредительная конференция Всесоюзной организации по статистическим методам и их применениям. На Учредительном съезде Всесоюзной статистической ассоциации (ВСА) в октябре 1990 г. в Московском экономико-статистическом институте эта организация вошла в состав ВСА в качестве секции статистических методов. В 1992 г. после развала СССР и фактического прекращения работы ВСА на основе секции статистических методов ВСА организована Российская ассоциация по статистическим методам (РАСМ), а затем и Российская академия статистических методов, существующие и в настоящее время. Коллективными усилиями разработан единый подход к проблемам принятия решений в сертификации и управлении качеством на основе применения современных статистических методов.

Статистический контроль — это выборочный контроль на научной основе. Контроль качества продукции всем знаком хотя бы по названию — им обычно занимается отдел технического контроля (ОТК) предприятия. Есть различные виды контроля — входной контроль, приемочный контроль (готовой продукции), и контроль при передаче полуфабрикатов и комплектующих из цеха в цех. Кроме сплошного контроля всех изделий подряд применяют выборочный, когда о качестве партии продукции судят по результатам контроля некоторой части — выборки.

Зачем нужен выборочный контроль? Чтобы проверить качество спички — надо чиркнуть ею. Загорится — должное качество, не загорится — брак. Но повторно однажды зажженную спичку использовать уже нельзя. Поэтому партию спичек можно контролировать только выборочно. Партии консервов, лампочек, патронов — тоже. При разрушающем контроле необходимо пользоваться выборочными методами и судить о качестве партии продукции по результатам контроля ее части — выборки.

Выборочные методы контроля могут применяться и из экономических соображений, когда стоимость контроля высока по сравнению со стоимостью изделия. Например, вряд ли целесообразно визуально проверять качество каждой скрепки в каждой коробке.

Для проведения выборочного контроля необходимо сформировать выборку, выбрать план контроля. А если план имеется — полезно знать его свойства. Анализ и синтез планов проводят с помощью математического моделирования на основе теории вероятностей и математической статистики, применяя компьютерные диалоговые системы (пакеты программ).

Компьютерные диалоговые системы позволяют, прежде всего, проводить анализ и синтез планов контроля. Пусть перед Вами — ГОСТ на продукцию, в нем есть раздел «Правила приемки» с планами контроля. Хороша эта система планов или плоха? С помощью диалоговых систем Вы найдете характеристики конкретного плана, приемочный и браковочный уровни дефектности (см. ниже) и т.д. Можно провести и синтез планов, т.е. компьютер поможет принять решение в новых условиях — подберет план, удовлетворяющий Вашим условиям.

Российской ассоциацией статистических методов были проанализированы сотни стандартов на конкретную продукцию (разделы «Правила приемки») и ГОСТы по статистическим методам. Обнаружено, что более половины и тех и других стандартов содержат грубые ошибки, пользоваться ими нельзя. Причины этого печального положения проанализированы в статье [4].

По оценкам, полученным в работе [6], применение современных статистических методов позволяет в среднем вдвое сократить трудозатраты на контрольные операции (как известно, они составляют расходуют примерно 10 % от себестоимости машиностроительной продукции). Следовательно, от повышения эффективности решений менеджеров на основе внедрения современных статистических методов обеспечения качества продукции Россия может получить более 5 миллиардов долларов США дополнительного дохода в год.

Приведем еще два сообщения о высокой экономической эффективности статистического контроля. «Мы документально зафиксировали экономию от применения методов статистического контроля и методов принятия решений, которым обучили наших сотрудников. Мы приближаемся к степени окупаемости около 30 долларов на 1 вложенный доллар. Вот почему мы получили такую серьезную поддержку от высшего руководства», — сообщает Билл Виггенхорн, ответственный за подготовку специалистов фирмы «Моторола» (цитируем по статье [7]).

По подсчетам профессора Массачусетского технологического института Фримена (см. монографию [8]), только статистический приемочный контроль давал промышленности США 4 млрд долл. в 1958 г. (32,8 млрд долл. в ценах 2021 г.), т.е. 0,8 % ВВП — валового внутреннего продукта.

Основы теории статистического контроля. Выборочный контроль, построенный на научной основе, т.е. исходящий из теории вероятностей и математической статистики, называют статистическим контролем. Организатора производства и менеджера выборочный контроль может интересовать не только в связи с качеством продукции, но и в связи, например, с контролем экологической обстановки, поскольку зафиксированные государственными органами экологические нарушения влекут штрафы и иные «неприятные» последствия. К выборочному контролю вынуждены прибегать при аудиторской проверке организации, поскольку сплошной анализ всего массива документов бухгалтерского или управленческого учета весьма трудоемок, а потому практически невозможен. Обсудим основные идеи статистического контроля.

При статистическом контроле решение о генеральной совокупности — об экологической обстановке в данном регионе или о партии продукции — принимается по выборке, состоящей из некоторого количества единиц (отдельных документов бухгалтерского или управленческого учета, единиц экологического контроля или единиц продукции). Следовательно, выборка должна представлять партию, т.е. быть репрезентативной (представительной). Как эти слова понимать, как проверить репрезентативность? Ответ может быть дан лишь в терминах вероятностных моделей выборки (см. главу 1).

Как известно, наиболее часто применяют биномиальную и гипергеометрическую модели. В первой из них предполагается, что результаты контроля n рассматриваемых единиц можно рассматривать как совокупность n независимых одинаково распределенных случайных величин $X_1, X_2, ..., X_n$, где $X_i = 1$, если i-е измерение показывает, что имеется нарушение, т.е. превышено ПДК (предельная норма концентрации) или i-е изделие дефектно, и $X_i = 0$, если это не так. Тогда число X превышений ПДК (при другой интерпретации — дефектных единиц продукции в партии) равно:

$$X = X_1 + X_2 + ... + X_n . \tag{1}$$

Из формулы (1) и Центральной Предельной Теоремы теории вероятностей (см. главу 14) вытекает, что при увеличении объема выборки n распределение X сближается с нормальным распределением. Известно, что распределение X имеет вид:

$$P(X = k) = C_n^k p^k (1 - p)^{n-k} , \tag{2}$$

где C_n^k — число сочетаний из n элементов по k, а p —уровень дефектности (в другой предметной области — в экологии — доля превышений ПДК в генеральной совокупности), т.е. $p = P(X_i = 1)$. Как известно, формула (2) задает биномиальное распределение.

Вторая модель — гипергеометрическая — соответствует случайному отбору единиц в выборку. Пусть среди N единиц, составляющих генеральную совокупность, имеется D дефектных. Случайность отбора означает, что каждая единица имеет одинаковые шансы попасть в выборку. Более того, ни одна пара единиц не имеет преимущества перед любой другой парой при отборе в выборку. То же самое — для троек, четверок и т.д. Итак, каждое из C_N^n сочетаний по n единиц из N имеет одинаковую вероятность быть отобранным в качестве выборки, равную, очевидно, $1/C_N^n$.

Отбор случайной выборки, согласно описанным правилам, организуют при проведении различных лотерей. Пусть Y — число дефектных единиц в такой выборке. Известно, что тогда $P(Y = k)$ — гипергеометрическое распределение, т.е.

$$P(Y=k)=\frac{C_n^k C_{N-n}^{D-k}}{C_N^n}. \qquad (3)$$

Известно, что биномиальная и гипергеометрическая модели *весьма близки*, когда объем генеральной совокупности (партии) по крайней мере в 10 раз превышает объем выборки, т.е.

$$P(X = k) = P(Y = k), \qquad (4)$$

если объем выборки мал по сравнению с объемом партии. При этом в качестве p в формуле (4) берут D/N.

Близость результатов, получаемых с помощью биномиальной и гипергеометрической моделей, весьма важна с методологической точки зрения. Дело в том, что эти модели исходят из принципиально различных модельных предпосылок. В биномиальной модели случайность *присуща каждой единице* — она с какой-то вероятностью дефектна, а с какой-то — годна. В то же время в гипергеометрической модели качество определенной единицы детерминировано, задано, а случайность проявляется лишь *в отборе*, вносится инженером, экологом или экономистом при составлении выборки. .

Соотношение (4) показывают, что во многих случаях нет необходимости выбирать одну из моделей, поскольку обе дают близкие численные результаты. Отличия проявляются при обсуждении вопроса о том, какую выборку считать *представительной*. Является ли таковой выборка, составленная из 20 изделий, лежащих сверху в первом вскрытом ящике? В биномиальной модели вполне допустим ответ «да», в гипергеометрической — только «нет».

Биномиальная модель легче для теоретического изучения, поэтому будем ее рассматривать в дальнейшем. Однако при реальном контроле лучше формировать выборку, исходя из гипергеометрической модели. Это делают, выбирая номера изделий (для включения в выборку) с помощью датчиков псевдослучайных чисел на ЭВМ или с помощью таблиц псевдослучайных чисел. Алгоритмы формирования выборки встраивают в современные программные продукты по статистическому контролю.

Планы статистического контроля и правила принятия решений. Под планом статистического контроля понимают алгоритм, т.е. правила действий при контроле. На «входе» при этом — генеральная совокупность (партия продукции), а на «выходе» — одно из двух решений: «принять партию» либо «забраковать партию». Рассмотрим несколько примеров.

Одноступенчатые планы контроля (n,c): отобрать выборку объема n; если число дефектных единиц в выборке X не превосходит c, то партию принять, в противном случае забраковать. Число c называется «приемочным числом».

Частные случаи: план $(n, 0)$ —партию принять тогда и только тогда, когда все единицы в выборке являются годными; план $(n, 1)$ — партия принимается, если в выборке все единицы являются годными или ровно одно — дефектное, во всех остальных случаях партия бракуется.

Двухступенчатый план контроля $(n, a, b) + (m, c)$: отобрать первую выборку объема n; если число дефектных единиц в первой выборке X не превосходит a, то партию принять; если число дефектных единиц в первой выборке X больше или равно b, то партию забраковать; во всех остальных случаях, т.е. когда X больше a, но меньше b, следует взять вторую выборку объема m; если число дефектных единиц во второй выборке Y не превосходит c, то партию принять, в противном случае забраковать.

Рассмотрим в качестве примера план (20, 0, 2) + (40, 0). Сначала берется первая выборка объема 20. Если все единицы в ней — годные, то партия принимается. Если две или больше — дефектные, партия бракуется. А если только одно — дефектное? В реальной ситуации в таких случаях начинаются споры между представителями предприятия и экологического контроля, или поставщика и потребителя. Говорят, например, что дефектная единица случайно попала в партию, что ее подсунули конкуренты или что при контроле случайно сделан неправильный вывод. Поэтому, чтобы споры пресечь, берут вторую выборку объема 40 (вдвое большего, чем в первый раз). Если все единицы во второй выборке — годные, то партию принимают, в противном случае — бракуют.

В реальной нормативно-технической документации — договорах на поставку, технических регламентах, стандартах, технических условиях, инструкциях по экологическому контролю и т.д. — не всегда четко сформулированы планы статистического контроля и правила принятия решений. Например, при описании двухступенчатого плана контроля вместо задания приемочного числа c может стоять загадочная фраза «результат контроля второй выборки считается окончательным». Остается гадать, как принимать решение по второй выборке. Менеджер, администратор (государственный служащий), инженер, эколог или экономист, занимающийся вопросами экологического контроля или контроля качества, должен первым делам добиваться кристальной ясности в формулировках правил принятия решений, иначе ошибочные и необоснованные решения, а потому и убытки неизбежны.

Оперативная характеристика плана статистического контроля. Каковы свойства плана статистического контроля? Они, как правило, определяются с помощью функции $f(p)$, связывающей вероятность p дефектности единицы контроля с вероятностью $f(p)$ приемки партии, положительной оценки экологической обстановки или правильности ведения бухгалтерской документации по результатам контроля. При этом вероятность p того, что конкретная единица дефектна, называется входным уровнем дефектности, а указанная функция называется оперативной характеристикой плана контроля. Если дефектные единицы отсутствуют, $p = 0$, то партия всегда принимается, т.е. $f(0) = 1$. Если все единицы дефектные, $p = 1$, то партия наверняка бракуется, $f(1) = 0$. Между этими крайними значениями p функция $f(p)$ монотонно убывает. При изучении свойств плана входной уровень дефектности p — свободный параметр, он может принимать любые значения между 0 и 1.

Вычислим оперативную характеристику плана $(n,0)$. Поскольку партия принимается тогда и только тогда, когда все единицы являются годными, а вероятность того, что конкретная единица — годная, равна $(1-p)$, то оперативная характеристика имеет вид

$$f(p) = P(X = 0) = (1 - p)^n. \qquad (5)$$

Для плана $(n,1)$ оперативная характеристика, как легко видеть, такова:

$$f(p) = P(X = 0) + P(X = 1) = (1 - p)^n + np\,(1 - p)^{n-1}. \qquad (6)$$

Оперативные характеристики для конкретных планов статистического контроля не всегда имеют такой простой вид, как в случае формул (5) и (6). Рассмотрим в качестве примера план (20, 0, 2) + (40, 0). Сначала найдем вероятность того, что партия будет принята по результатам контроля первой партии. Согласно формуле (5) имеем:

$$f_1(p) = P(X=0) = (1-p)^{20}.$$

Вероятность того, что понадобится контроль второй выборки, равна:

$$P(X = 1) = 20p(1-p)^{19}.$$

При этом вероятность того, что по результатам ее контроля партия будет принята, равна

$$f_2(p) = P(X = 0) = (1-p)^{40}.$$

Следовательно, вероятность того, что партия будет принята со второй попытки, т.е. что при контроле первой выборки обнаружится ровно одна дефектная единица, а затем при контроле второй — ни одной, равна

$$f_3(p) = P(X=1)f_2(p) = 20p(1-p)^{19}(1-p)^{40} = 20p(1-p)^{59}.$$

Следовательно, вероятность принятия партии с первой или со второй попытки равна

$$f(p) = f_1(p) + f_3(p) = (1-p)^{20} + 20p(1-p)^{59}.$$

При практическом применении методов статистического приемочного контроля для нахождения оперативных характеристик планов контроля вместо формул, имеющих обозримый вид лишь для отдельных видов планов, применяют численные компьютерные алгоритмы или заранее составленные таблицы.

Риск поставщика и риск потребителя, приемочный и браковочный уровни дефектности. С оперативной характеристикой связаны важные понятия *приемочного и браковочного уровней дефектности*, а также понятия *«риск поставщика»* и *«риск потребителя»*. Чтобы ввести эти понятия, на оперативной характеристике выделяют две характерные точки, делящие входные уровни дефектности на три зоны (области) — *А*, *Б* и *В*. В зоне *А* почти всегда все хорошо, а именно — почти всегда экологическая обстановка признается благополучной, почти все партии принимаются. В зоне *В*, наоборот, почти всегда все плохо, а именно — почти всегда экологический контроль констатирует экологические нарушения, почти все партии бракуются. Зона. *Б* — буферная, переходная, промежуточная, в ней как вероятность приемки, так и вероятность браковки заметно отличаются от 0 и 1. Для задания границ между зонами выбирают два малых числа—риск поставщика (производителя, предприятия) α и риск потребителя (заказчика, системы экологического контроля) β, при этом границы между зонами задают два уровня дефектности — приемочный $p_{пр}$ и браковочный $p_{бр}$, определяемые из уравнений

$$f(p_{пр}) = 1 - \alpha, f(p_{бр}) = \beta. \qquad (7)$$

Таким образом, если входной уровень дефектности не превосходит $p_{пр}$, то вероятность забракования партии мала, т.е. не превосходит α. Приемочный уровень дефектности выделяет зону *А* значений входного уровня дефектности, в которой нарушения экологической безопасности почти всегда не отмечаются, партии почти всегда принимаются, т.е. соблюдаются интересы проверяемого предприятия (в экологии), поставщика (при контроле качества). Это — зона комфортности для поставщика. Если он обеспечивает работу (входной уровень дефектности) в этой зоне, то его практически никогда никто не потревожит.

Если же входной уровень дефектности больше браковочного уровня дефектности $p_{бр}$, то нарушения почти наверняка фиксируются, партия почти всегда бракуется, т.е. экологи узнают о нарушениях, потребитель оказывается защищен от попадания к нему партий со столь высоким уровнем брака. Поэтому можно сказать, что в зоне *В* соблюдаются интересы потребителей — брак к ним не попадает.

При выборе плана контроля часто начинают с выбора приемочного и браковочного уровней дефектности. При этом выбор конкретного значения приемочного уровня дефектности отражает интересы поставщика, а выбор конкретного значения браковочного уровня дефектности — интересы потребителя. Можно доказать (см. следующий раздел), что для любых положительных чисел α и β и любых входных уровней дефектности $p_{пр}$ и $p_{бр}$, причем $p_{пр}$ меньше $p_{бр}$, найдется план контроля (n,c) такой, что его оперативная характеристика $f(p)$ удовлетворяет неравенствам

$$f(p_{пр}) \geq 1 - \alpha, f(p_{бр}) \leq \beta.$$

При практических расчетах обычно принимают α = 0,05 (т.е. 5 %) и β = 0,1 (т.е. 10 %).

В качестве примера вычислим приемочный и браковочный уровни дефектности для плана (n,0). Из формул (5) и (7) вытекает, что

$$(1 - p_{\text{пр}})^n = 1 - \alpha,\ p_{\text{пр}} = 1 - (1 - \alpha)^{1/n}.$$

Поскольку риск поставщика α мал, то из известного соотношения математического анализа:

$$\sqrt[n]{1-\alpha} = 1 - \frac{\alpha}{n} + O\left(\frac{\alpha^2}{n^2}\right)$$

вытекает приближенная формула

$$p_{\text{пр}} \approx \frac{\alpha}{n}.$$

Для браковочного уровня дефектности имеем

$$p_{\text{бр}} = 1 - \beta^{1/n}.$$

При практическом применении методов статистического приемочного контроля аналитическими формулами, имеющих обозримый вид лишь для отдельных видов планов, не пользуются. Для нахождения приемочных и браковочных уровней дефектности планов контроля вместо них применяют численные компьютерные алгоритмы или заранее составленные таблицы. Такие таблицы имеются в нормативно-технической документации или научно-технических публикациях.

Предел среднего выходного уровня дефектности. Обсудим судьбу забракованной партии продукции. В зависимости от ситуации эта судьба может быть разной. Партия может быть утилизирована. Например, забракованная партия гвоздей может быть направлена на переплавку. У партии может быть понижена сортность, и она может быть продана по более низкой цене (при этом результаты выборочного контроля будут использованы не только для констатации того, что не выдержан заданный уровень качества, но и для оценки реального уровня качества). Наконец, партия продукции может быть подвергнута сплошному контролю (для этого обычно привлекают инженеров из всех заводских служб). При сплошном контроле все дефектные изделия обнаруживаются и либо исправляются на месте, либо извлекаются из партии. В результате в партии остаются только годные изделия. Такая процедура называется «контроль с разбраковкой».

При среднем входном уровне дефектности p и применении контроля с разбраковкой с вероятностью $f(p)$ партия принимается (и уровень дефектности в ней по-прежнему равен p) и с вероятностью $(1 - f(p))$ бракуется и подвергается сплошному контролю, в результате чего к потребителю поступают только годные изделия. Следовательно, по формуле полной вероятности средний выходной уровень дефектности равен

$$f_1(p) = pf(p) + 0(1 - f(p)) = pf(p).$$

Средний выходной уровень дефектности $f_1(p)$ равен 0 при p=0 и p=1, положителен на интервале (0;1), а потому достигает на нем максимума, который в теории статистического контроля называется пределом среднего выходного уровня дефектности (сокращенно ПСВУД):

$$\text{ПСВУД} = \max_{0 \le p \le 1} f_1(p).$$

Пример. Рассмотрим план $(n,0)$. Для него $f(p) = (1 - p)^n$ и $f_1(p) = p(1 - p)^n$. Чтобы найти ПСВУД, надо приравнять 0 производную среднего выходного уровня дефектности по среднему входному уровню дефектности:

$$\frac{df_1(p)}{dp} = \left(p(1-p)^n\right)' = (1-p)^n + pn(1-p)^{n-1} =$$
$$= (1-p)^{n-1}(1-p-pn) = (1-p)^{n-1}(1-(n+1)p) = 0.$$

В полученном уравнении корень $p = 1$ соответствует минимуму, а не максимуму. Поскольку непрерывная функция на замкнутом отрезке достигает максимума, то максимум достигается при

$$p_n = \frac{1}{n+1}.$$

Следовательно,

$$\text{ПСВУД} = p_n(1-p_n)^n = \frac{1}{n+1}\left(1-\frac{1}{n+1}\right)^n. \qquad (8)$$

По выражению (8) могут быть проведены конкретные расчеты. Однако оно довольно громоздко. Его можно упростить, используя один замечательный предел из курса математического анализа, а именно:

$$\lim_{n\to\infty}\left(1-\frac{1}{n+1}\right)^{n+1} = e^{-1} = \frac{1}{2{,}718281828...} \approx 0{,}368. \qquad (9)$$

Сравнивая соотношения (8) и (9), видим, что

$$\text{ПСВУД} = \left(\frac{\frac{1}{n+1}}{1-\frac{1}{n+1}}\right)\left(1-\frac{1}{n+1}\right)^{n+1}.$$

Первая скобка равна $1/n$, а вторая согласно соотношению (9) приближается к 0,368 при росте объема выборки. Поэтому получаем простую асимптотическую формулу

$$\text{ПСВУД} \approx \frac{0{,}368}{n}.$$

Для более сложных планов ПСВУД рассчитывают с помощью более или менее сложных компьютерных программ.

При рассмотрении основ статистического контроля в настоящем пункте расчетные формулы удалось получить лишь для простейших планов, в основном для планов вида $(n,0)$. Если ослабить требования и рассчитывать не на точные формулы, а на асимптотические, при $n \to \infty$, то можно справиться и с одноступенчатыми планами вида (n, c).

10.2. Асимптотическая теория одноступенчатых планов

Пусть X — число дефектных единиц продукции в выборке объема n. Как уже отмечалось, распределение X является биномиальным и имеет вид

$$P(X = k) = C_n^k p^k (1-p)^{n-k},$$

где C_n^k — число сочетаний из n элементов по k, а p — входной уровень дефектности.

Пусть используется одноступенчатый план контроля (n, c). Тогда оперативная характеристика этого плана имеет вид:

$$f(p) = \sum_{1 \le k \le c} P(X = k) = \sum_{1 \le k \le c} C_n^k p^k (1-p)^{n-k}.$$

Пусть $n \to \infty$. Тогда по Закону Больших Чисел теории вероятностей (по теореме Бернулли)

$$\frac{X}{n} \to p$$

(сходимость по вероятности). Значит, если c/n окажется заметно больше входного уровня дефектности p, то партии будут почти всегда приниматься, а если c/n окажется заметно меньше входного уровня дефектности p, то партии будут почти всегда отклоняться. Ситуация будет нетривиальной только там, где величины c/n и p близки друг к другу.

Хотя оперативная характеристика рассчитывается с помощью сумм биномиальных вероятностей, анализировать эти суммы затруднительно. Поэтому целесообразно найти для нее приближение с помощью теоремы Муавра-Лапласа. Имеем цепочку тождественных преобразований:

$$f(p) = P(X \le c) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \le \frac{c - np}{\sqrt{np(1-p)}} \right).$$

Справа строит именно то выражение, которое участвует в теореме Муавра-Лапласа. Воспользовавшись равномерной сходимостью в этой теореме, можно записать, что

$$\lim_{n \to \infty} f(p) = \Phi\left(\frac{c - np}{\sqrt{np(1-p)}} \right),$$

где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

Последняя формула позволяет указать асимптотические выражения для приемочного и браковочного уровней дефектности. Действительно, согласно определениям этих понятий

$$\Phi\left(\frac{c - np_{пр}}{\sqrt{np_{пр}(1 - p_{пр}))}} \right) = 1 - \alpha, \quad \Phi\left(\frac{c - np_{бр}}{\sqrt{np_{бр}(1 - p_{бр}))}} \right) = \beta, \qquad (10)$$

откуда с помощью элементарных преобразований получаем, что

$$p_{пр} = \frac{c}{n} - \frac{1}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}\Phi^{-1}(1-\alpha), \quad p_{бр} = \frac{c}{n} - \frac{1}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}\Phi^{-1}(\beta). \qquad (11)$$

Так как величины c/n и p близки друг к другу, то при переходе от формулы (10) к формуле (11) в подкоренных выражениях приемочный и браковочный уровни дефектности заменены на c/n (с точностью до бесконечно малых более высокого порядка).

Поскольку при практическом применении статистического приемочного контроля, как уже отмечалось, принимают α = 0,05, β =0,10, то в предыдущие формулы следует подставить Φ^{-1} (0,95) =1,64 и Φ^{-1} (0,10) = – 1,28. Итак, итоговые формулы для приемочного и браковочного уровней дефектности имеют вид

$$p_{пр} = \frac{c}{n} - \frac{1{,}64}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}, \quad p_{бр} = \frac{c}{n} + \frac{1{,}28}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}.$$

Перейдем к задаче синтеза. Пусть заданы приемочный и браковочный уровни дефектности. Требуется построить одноступенчатый план, имеющий эти характеристики. Из формул (10) следует, в частности, что

$$c - np_{пр} = \Phi^{-1}(1-\alpha)\sqrt{np_{пр}(1-p_{пр})}, \quad c - np_{бр} = \Phi^{-1}(\beta)\sqrt{np_{бр}(1-p_{бр})}. \qquad (12)$$

Вычитая из первого уравнения второе, получаем, что

$$np_{бр} - np_{пр} = \sqrt{n}\{\Phi^{-1}(1-\alpha)\sqrt{p_{пр}(1-p_{пр})} - \Phi^{-1}(\beta)\sqrt{p_{бр}(1-p_{бр})}\}.$$

Следовательно, оценка n^* необходимого объема выборки имеет вид

$$n^* = \left(\frac{\Phi^{-1}(1-\alpha)\sqrt{p_{пр}(1-p_{пр})} - \Phi^{-1}(\beta)\sqrt{p_{бр}(1-p_{бр})}}{p_{бр} - p_{пр}}\right)^2.$$

Для стандартных значений рисков α = 0,05, β = 0,10 имеем:

$$n^* = \left(\frac{1{,}64\sqrt{p_{пр}(1-p_{пр})} + 1{,}28\sqrt{p_{бр}(1-p_{бр})}}{p_{бр} - p_{пр}}\right)^2 \qquad (13)$$

С помощью уравнений (12) нетрудно найти оценку c^* приемочного числа, заменив неизвестный объем выборки на его оценку n^*. Будем использовать оценку

$$c^* = n^* p_{бр} + \Phi^{-1}(\beta)\sqrt{n^* p_{бр}(1 - p_{бр})}.$$

Для стандартного значения β=0,10 имеем

$$c^* = n^* p_{бр} - 1{,}28\sqrt{n^* p_{бр}(1 - p_{бр})}. \qquad (14)$$

Итак, по формуле (13) можно рассчитать оценку объема выборки, затем по формуле (14) найти оценку приемочного числа. Необходимо отметить, что результаты расчетов по рассматриваемым асимптотическим формулам отнюдь не всегда дают целые числа, поэтому необходима корректировка полученных результатов.

Полученные формулы позволяют решить сформулированную выше задачу — по заданным приемочному и браковочному уровням дефектности подобрать такой одноступенчатый план контроля, что его оперативная характеристика f(p) удовлетворяет неравенствам

$$f(p_{пр}) > 1 - \alpha,\ f(p_{бр}) < \beta.$$

Поэтому при практической работе корректировка асимптотических результатов должна быть направлена на выполнение указанных неравенств.

Пример. Пусть $p_{пр} = 0{,}02$, $p_{бр} = 0{,}09$. Тогда по формуле (13) оценка объема выборки равна

$$n^* = \left(\frac{1{,}64\sqrt{0{,}02(1-0{,}02)} + 1{,}28\sqrt{0{,}09(1-0{,}09)}}{0{,}09 - 0{,}02}\right)^2 = \left(\frac{1{,}64 \cdot 0{,}14 + 1{,}28 \cdot 0{,}286}{0{,}07}\right)^2 =$$
$$= \left(\frac{0{,}2296 + 0{,}3661}{0{,}07}\right)^2 = \left(\frac{0{,}5957}{0{,}07}\right)^2 = 8{,}51^2 = 72{,}42.$$

Полученное число не является натуральным, поэтому вполне естественно откорректировать объем выборки до ближайшего целого, т.е. до $n^* = 72$.

Оценку приемочного числа находим по формуле (14):

$$c^* = 72 \cdot 0{,}09 - 1{,}28\sqrt{72 \cdot 0{,}09 \cdot 0{,}91} = 6{,}48 - 1{,}28 \cdot 2{,}428 = 3{,}37.$$

Полученное число не является целым, поэтому в качестве приемочного числа надо взять ближайшее целое, т.е. до 3.

Если объем выборки округлить до 73, то аналогично получим

$$c **= 73 \cdot 0{,}09 - 1{,}28\sqrt{73 \cdot 0{,}09 \cdot 0{,}91} = 6{,}57 - 1{,}28 \cdot 2{,}445 = 3{,}44.$$

При округлении снова получаем 3.

С помощью первого из уравнений (12) можно построить оценку c^* на основе приемочного уровня дефектности:

$$c^* = n^* p_{пр} + \Phi^{-1}(1-\alpha)\sqrt{n^* p_{пр}(1-p_{пр})} = n^* p_{пр} + 1{,}64\sqrt{n^* p_{пр}(1-p_{пр})}\,.$$

Подставив конкретные значения, получим практически ту же оценку, что и раньше:

$$c *= n * p_{пр} + 1{,}64\sqrt{n * p_{пр}(1-p_{пр})} = 72 \cdot 0{,}02 + 1{,}64\sqrt{72 \cdot 0{,}02 \cdot 0{,}98} = 3{,}39\,.$$

Итак, в результате асимптотических расчетов найден одноступенчатый план (72, 3).

10.3. Практическое применение статистического контроля

Познакомившись с некоторыми основными понятиями, подходами и идеями теории статистического контроля качества, обсудим подробнее практические стороны этой технико-экономической области.

Анализ и синтез планов контроля. На основе теории статистического контроля можно проанализировать планы контроля качества, имеющиеся в нормативно-технической документации (стандартах, технических условиях) и в договорах на поставку продукции и оказание услуг. Достаточно часто оказывается, что формулировки соответствующих разделов (разделов «Правила приемки», «Методы контроля» и др.) имеют различные недостатки и неточности, что может послужить в дальнейшем причиной к возникновению арбитражных ситуаций (т.е. решаемых через арбитражные или иные суды).

Если обсуждаемая система контроля качества выдерживает чисто логическую проверку, то наступает вторая стадия — анализ с точки зрения теории статистического контроля. На этой стадии рассчитывают характеристики применяемых планов контроля. О некоторых из них уже шла речь — приемочный и браковочный уровни дефектности, предел среднего выходного уровня дефектности. Есть и иные показатели, например, средний используемый объем выборки, средняя стоимость контроля, и т.п. Особенно важна прогнозируемая доля арбитражных ситуаций (споров между предприятиями) при используемой системе контроля.

На стадии анализа возможны неожиданные «открытия». Например, может оказаться, что существующая система контроля качества, хотя и является формально безупречной, но защищает лишь от приемки столь плохих партий продукции, в которых более половины единиц продукции дефектно (т.е. для применяемых планов контроля браковочный уровень дефектности больше 0,5). Или что система контроля защищает интересы поставщиков, у которых каждое пятое изделие является бракованным (приемочный уровень дефектности равен 0,2).

Замечание. До сих пор постоянно говорилось о контроле единиц и партий продукции. Однако нет никакого принципиального отличия с контролем услуг (медицинских, туристических, транспортных, образовательных, банковских и иных) или документации. Поэтому теория и практика статистического контроля качества продукции дает полезные рекомендации для банковского дела и бухгалтерского аудита. Надо только аккуратно заменить слова, описывающие предметную область применения теории статистического контроля.

После анализа ситуации с системой контроля естественно перейти к улучшению этой системы, к обоснованному выбору планов, к этапу синтеза. В зависимости от конкретных условий используются разнообразные подходы к выбору планов. Например, задают приемочный и браковочный уровни дефектности. В случае контроля с разбраковкой естественно использовать ограничения на предел среднего выходного уровня дефектности.

Какая оптимизация планов статистического контроля возможна? Обсудим подробнее оптимизационные постановки в статистическом приемочном контроле. Очевидно, имеется три вида затрат и потерь:

- затраты непосредственно на проведение контроля единиц продукции, включенных в выборку,

- потери в случае неверного решения о забраковании партии продукции (в которой на самом деле доля дефектной продукции *соответствует* требованиям договора между поставщиком и потребителем или иной нормативно-технической документации);

- потери в случае неверного решения о принятии партии продукции (в которой на самом деле доля дефектной продукции *не соответствует* требованиям договора между поставщиком и потребителем или иной нормативно-технической документации).

При этом первые два вида затрат непосредственно связаны с деятельностью предприятия, на котором производится продукция, третий же вид затрат (потерь) формируется там, где она потребляется. С этим связана принципиальная сложность подсчета затрат третьего вида. Особенно эта сложность проявляется тогда, когда попадание к потребителю дефектных изделий может привести к авариям с человеческими жертвами. Тогда возникает вопрос: сколько стоит человеческая жизнь? Только оценив потери здоровья и жизни в денежных единицах, можно сформировать функционал качества плана статистического контроля и затем оптимизировать его. Поскольку невозможно (прежде всего, из этических и религиозных соображений) выразить стоимость человеческой жизни в денежных единицах, то невозможно сформировать функционал качества плана статистического контроля и тем более оптимизировать его.

К счастью, для большинства видов продукции вопрос о денежной оценке человеческой жизни не возникает. Проблема обычно «всего лишь» в том, что выпущенная продукция используется разнообразными конечными потребителями, а потому оценить отрицательный эффект повышения доли ее дефектности затруднительно.

Поэтому наряду с функционалом качества, включающим все три вида затрат, рассматривают «условный» функционал на основе затрат первых двух типов, а на вероятность принятия партии продукции, в которой доля дефектной продукции не соответствует требованиям нормативно-технической документации, накладывают ограничение, т.е., грубо говоря, третий вид затрат учитывают в качестве ограничения.

Должны ли совпадать планы контроля у поставщика и потребителя? Естественно также по-разному проводить контроль у поставщика (производителя) и потребителя (заказчика). Пусть для определенности поставщик используют план $(n_1,0)$, а потребитель — $(n_2,0)$. Тогда естественно зафиксировать в договоре о поставке, что $n_1 >> n_2$. Такая договоренность обеспечит тщательный контроль со стороны изготовителя и почти автоматическое подтверждение приемки со стороны потребителя (т.е. отсутствие спора).

Одна из распространенных догм состоит в том, что изготовитель и потребитель должны проводить контроль по одним и тем же планам контроля. Если план контроля и входной уровень контроля таков, что ситуация контроля относится к буферной зоне *Б*, т.е. вероятность приемки партии заметно отличается от 0 и 1, то указанная догма приводит к высокой вероятности спорных ситуаций. Пусть, например, оперативная характеристика равна 0,5. Пусть изготовитель принял партию (с вероятностью 0,5). После этого при независимом контроле у потребителя с той же вероятностью 0,5 она может быть отклонена и с вероятностью 0,5 принята. Значит, общий итог таков: 50 % за то, что партия будет забракована у поставщика, 25 % — за спорную ситуацию (поставщик принял, потребитель забраковал), 25 % — за принятие и поставщиком и потребителем. Конечно, рассмотрен крайний случай — наиболее частое появление спорных ситуаций. Но реальное появление 10–15 % арбитражных споров — это типовая ситуация в 1980-е гг.

Один из вариантов выбора планов контроля поставщиком и потребителем выглядит так. Стороны договариваются о некотором «приемлемом» входном уровне дефектности p^*. Затем поставщик выбирает план контроля, используя p^* как браковочный уровень дефектности, а потребитель — рассматривая p^* как приемочный уровень дефектности. Подробнее об анализе, синтезе и оптимизации планов статистического контроля рассказано в специальной литературе, в частности, в работах [6, 8].

Усеченные планы. Пусть единицы продукции контролируются одна за другой (т.е. последовательно). Рассмотрим план статистического контроля (60, 3). Пусть при проверке единицы продукции появляются в таком порядке: дефектная, дефектная, дефектная, дефектная, … Четыре дефектные единицы подряд! Надо ли дальше проверять выборку? Исходя из здравого смысла — нет. Ведь совершенно неважно, каковы будут результаты по остальным 59-ти единицам продукции, окажутся они годными или дефектными — 4 дефектные единицы уже есть, и партию следует забраковать. Контроль мог бы быть прекращен и тогда, когда при проверке 57 единиц все 57 окажутся годными — независимо от качества остальных 3 партию надо принимать.

Усеченные планы — это планы статистического контроля, в которых контроль разрешается прекращать, если итог (принятие или забракование партии) становится ясен ранее, чем проведен контроль всех включенных в выборку единиц продукции. Усеченные планы применяются, когда единицы продукции поступают на контроль последовательно, одна за другой (или группа за группой). Это не всегда так. Если, например, план (60, 3) применяется для контроля качества электролампочек, и все 60 лампочек ввернуты в гнезда на испытательном стенде и одновременно включены, то подход на основе усеченных планов применить нельзя.

Возможность применения усеченных планов должна быть явным образом указана в нормативно-технической документации и в договорах на поставку. Опишем юридический казус, связанный с усеченными планами. В ГОСТе на штангенциркули был предусмотрен план контроля (20, 0). Органы Госстандарта проверяли завод «Точнометр» (название изменено). Проверили первый штангенциркуль — дефектен, второй — дефектен, …, десятый — дефектен. На этом комиссия остановилась, вполне резонно (с точки зрения здравого смысла) решив, что партия штангенциркулей должна быть забракована. Органы Госстандарта наложили на завод «Точнометр» штраф за выпуск некачественной продукции (в соответствии с действующим в то время правопорядком). Однако завод опротестовал это решение в суде. И суд удовлетворил протест, ссылаясь на то, что порядок проведения контроля качества штангенциркулей был нарушен! Бракоделы не смогли бы уйти от наказания, если бы в соответствующих документах была бы прописана возможность использования усеченных планов.

Выделение единиц бесформенной (жидкой, газообразной) продукции. Во всем предыдущем изложении постоянно встречается термин «единица продукции». Он вполне ясен, если речь идет об отдельных изделиях — дискетах, коробках спичек, патронах, бутылках минеральной воды, электробритвах, или отдельных деталях — болтах, гвоздях, пластмассовых дисках… Однако многие виды продукции имеют иной вид — газообразный, жидкий или, как говорят, бесформенный (порошкообразный, желеобразный, …). Как быть с ними? В работе [9] предложен подход, позволяющий применить к бесформенной продукции методы статистического контроля качества.

Основное — это выделить единицу продукции. Она не должна быть очень малой, поскольку ясно, что в бесформенной продукции свойства вещества в близких точках близки. Основная идея состоит в том, чтобы взять некоторое количество пар точек, отстоящих друг от друга на определенное расстояние, и выяснить, есть связь (т.е. значим ли ранговый коэффициент корреляции Спирмена — см. главу 6.1) между значениями изучаемого свойства в этих парах точек или нет. Если связь есть, значит, точки разнесены на недостаточное расстояние, другими словами, точки относятся к одной и той же единице продукции. Поэтому расстояние между точками надо увеличить. Если связь уже не обнаруживается, то это значит, что они относятся к разным единицам продукции. В процессе увеличения расстояния тем самым была оценена величина ребра куба, в виде которого условно представляем себе единицу бесформенной продукции. Разбив бесформенную продукцию на единицы, можно применять описанные выше подходы для контроля ее качества (подробнее см. [9]).

Отбор случайной выборки при статистическом контроле качества продукции. Как и при любом выборочном обследовании, при статистическом контроле качества продукции остро строит проблема отбора репрезентативной (представительной) выборки (см. главу 1 выше). Эта проблема усугубляется экономической заинтересованностью участников процесса. Наиболее научно-обоснованным является использование датчиков псевдослучайных чисел (а не таблиц или физических датчиков случайных чисел).

Исходя из экономической и технической целесообразности, популярна схема многоступенчатой выборки. Например, из 15 вагонов отобрать вагон № 5, из него — контейнер №3 около двери (из 12 контейнеров), из контейнера №3 — ящики № 7, 15 и 23, а из этих ящиков — каждое пятое изделие. При этом описании составления выборки совершенно ясно, что реально классическая случайная выборка может быть организована лишь при контроле контейнера №3, и остается только надеяться, что он является типичным для всей партии.

Всегда ли нужен контроль качества продукции? Чем выше достигнутый уровень качества, тем больше необходимый объем контроля — таков парадокс классической теории статистического контроля. Возможный выход состоит в переходе к принципиально новому подходу, обеспечивающему существенное расширение возможностей менеджера при выборе технической политики на основе учета экономических рисков. При этом оказывается, что даже «перекладывание» контроля на потребителя может быть экономически выгодно, если производитель организовал защиту от риска методом пополнения партий или путем развития технического обслуживания.

В государственных стандартах, технических условиях, другой нормативно-технической документации, относящейся к потребительским товарам и услугам, различным изделиям, веществам, материалам, иным видам продукции, а также в договорах между поставщиками и потребителями обычно присутствуют разделы «Правила приемки и методы контроля». Поэтому, в частности, методы статистического контроля качества продукции являются важной составной частью статистических методов сертификации, которым посвящена работа [4]. Как уже говорилось, имеется соответствующая вероятностно-статистическая теория, посвященная анализу и синтезу (выбору) планов контроля. Однако эта теория вообще не предусматривает отказа от контроля, поскольку игнорирует возможность перехода на иную стратегию организации взаимоотношений поставщика и потребителя, например, на стратегию технического обслуживания, при которой выходной контроль не проводится, а обнаруженные потребителями дефектные изделия заменяются годными или ремонтируются. Основная обсуждаемая в настоящем пункте идея — обоснование необходимости включения теории статистического приемочного контроля в более широкую технико-экономическую теорию взаимоотношений поставщиков и потребителей и целесообразности перехода при повышении качества продукции от контроля качества к иным способам защиты потребителя, например, к развитому техническому обслуживанию или к поставке запасных единиц продукции.

Использование экономических показателей при выборе планов статистического (выборочного) контроля пропагандировалось давно, но делалось это в рамках парадигмы обязательности контроля. Здесь рассматривается более широкая система взглядов, согласно которой контроль качества продукции — лишь один из способов урегулирования взаимоотношений между поставщиками и потребителями.

В более широком плане речь идет об отказе от получения детальной информации, если она стоит слишком дорого, и переходе к использованию иных механизмов управления. Так, качественные методы химического анализа часто используют именно потому, что соответствующие количественные методы более трудоемки и дороги, но ненамного полезнее с практической точки зрения. Пример из всем знакомой области: в средней школе знания учащихся контролируются еженедельно, в высшей же — один или несколько раз в семестр. Однако разница с точки зрения эффективности управления процессом обучения невелика. Другой пример: как показано в статистике интервальных данных, из-за погрешностей измерений нецелесообразно увеличивать их число сверх некоторого «рационального объема выборки», а для увеличения точности оценивания характеристик вероятностных распределений необходимо использовать более точные средства измерения. С учетом сказанного описываемый здесь подход представляется не столь необычным.

Оценка снизу необходимого объема выборки. Как известно, в теории статистического приемочного контроля качества продукции разработано много подходов к выбору планов контроля:

- на основе приемочного и браковочного уровней дефектности;
- исходя из предела среднего выходного уровня дефектности (при контроле с разбраковкой);
- с использованием экономических показателей, относящихся к предприятию (см., например, ГОСТ 24660-81);
- с использованием экономических показателей, относящихся к народному хозяйству в целом; и т.д.

Имеется обширная литература, посвященная обоснованию и сравнению этих подходов, разработке соответствующей математической теории и программного обеспечения. Не углубляясь в эти проблемы, сосредоточим внимание на одном парадоксальном явлении: при повышении качества выпускаемой продукции теория рекомендует увеличивать объем контроля!

Действительно, при повышении качества выпускаемой продукции требования потребителя, очевидно, обеспечиваются все лучше. Следовательно, должен уменьшаться браковочный уровень дефектности, т.е. то значение входного уровня дефектности, при котором вероятность приемки партии равна риску потребителя. Из всех планов с общим объемом контроля n минимум вероятности приемки партии (т.е. оперативной характеристики) достигается на одноступенчатом плане (n, 0). (Напомним, что, согласно этому плану, партия принимается тогда и только тогда, когда из n проверенных единиц продукции все оказываются годными.) Другими словами, оперативная характеристика для плана (n, 0) является огибающей (снизу) множества всех оперативных характеристик. Следовательно, из всех планов с общим объемом контроля n минимум браковочного уровня дефектности достигается также на плане (n, 0).

В дальнейшем будем исходить из биномиальной модели выборки, согласно которой число дефектных единиц продукции в выборке объема n имеет биномиальное распределение с параметрами n и p, где p — входной уровень дефектности. Как хорошо известно, эта модель является приближением для модели простой случайной выборки из партии, согласно которой указанное число имеет гипергеометрическое распределение. Напомним, что по чисто математическим причинам гипергеометрическая модель переходит в биномиальную, если объем партии безгранично возрастает, а доля дефектных единиц продукции в партии приближается к p. Если объем выборки составляет не более 10 % объема партии, то с достаточной для практики точностью принимают, что соответствующее биномиальное распределение хорошо приближает гипергеометрическое.

Примем обычное предположение о том, что риск потребителя равен 0,10. Как известно, браковочный уровень дефектности $p_{бр}$ для плана (n,0) определяется из условия

$$(1 - p_{бр})^n = 0{,}10.$$

Это соотношение дает возможность по заданному браковочному уровню дефектности $p_{бр}$ найти необходимый объем выборки:

$$n = \ln 0{,}10 / \ln (1 - p_{бр}) = -2{,}30/\ln (1 - p_{бр}).$$

Поскольку в силу сказанного ранее представляют интерес малые значения браковочного уровня дефектности, воспользуемся тем, что при малых x, согласно правилам математического анализа,

$$\ln(1+x) = x + O(x^2).$$

Вторым слагаемым в правой части последней формулы, как обычно в асимптотических рассуждениях, можно пренебречь. Следовательно, необходимый объем выборки с достаточной точностью может быть найден по формуле

$$n = 2{,}30/p_{бр}. \qquad (15)$$

(При конкретных расчетах надо, очевидно, правую часть округлить до ближайшего целого числа.) Например, при довольно низком (с точки зрения мирового рынка) качестве выпускаемой продукции можно задать $p_{бр} = 0{,}01$, т.е. потребовать, чтобы почти все (точнее, не менее 90 %) партии, в которых дефектных единиц больше, чем 1 из 100, были забракованы и не достигли потребителя. Тогда объем контроля должен составлять не менее $n = 230$.

Основной парадокс теории статистического приемочного контроля. Как следует из сказанного выше, необходимый объем выборки, определяемый для какого-либо плана контроля по заданному браковочному уровню дефектности $p_{бр}$, будет не меньше, чем для плана $(n,0)$, т.е. не меньше, чем $2{,}30/p_{бр}$.Таким образом, если достигнут достаточно высокий уровень качества, такой, что потребителю может попасть не более 1 дефектной единицы продукции из 10 000, т.е. $p_{бр} = 0{,}0001$, то объем контроля должен быть не меньше $n = 23\,000$. Если же качество повысится в 100 раз, т.е. потребителю сможет попасть не более 1 дефектной единицы продукции из 1 000 000, то объем контроля и затраты на него возрастут также в 100 раз, и минимально необходимый объем контроля составит 2,3 миллиона единиц продукции. Поскольку объем партий большинства (практически всех!) видов продукции существенно меньше этого числа, то проведенные выше расчеты говорят о необходимости перехода на сплошной контроль.

Итак, выводы парадоксальны: если качество выпускаемой продукции не очень хорошее, то целесообразно проводить статистический (выборочный) контроль, если же качество возрастает, то объем контроля и затраты на него увеличиваются, вплоть до перехода на сплошной контроль. Если это возможно, т.е. контроль не является разрушающим. А если невозможно, то попадаем в тупиковую ситуацию — высокое качество не может быть подтверждено.

В реальных ситуациях объемы контролируемых выборок — единицы или десятки, но обычно отнюдь не сотни и тысячи. Если контролируются 100 изделий, то согласно формуле (15) браковочный уровень дефектности равен 2,3 %. И это — предел для реально используемых объемов контроля. Следовательно, статистический приемочный контроль (в том числе выходной или входной) может быть применен для контроля лишь такой продукции, в которой из 50 изделий хотя бы одно дефектно. Другими словами, этот метод управления качеством предназначен лишь для продукции сравнительно низкого качества (входной уровень дефектности не менее 1–2 %) или при обслуживании потребителя, согласного на довольно высокий браковочный уровень дефектности (не менее 2,3 %).

Следовательно, для повышения качества необходимо использовать контрольные карты и другие методы статистического регулирования технологических процессов на предприятии. О них подробно рассказано в следующем разделе, а также, например, в монографиях [1,10]. В частности, упомянем методы «всеобщего» (в другом переводе — тотального) контроля качества и др. Недаром этим методам уделяется больше внимания в зарубежных методических изданиях, чем собственно статистическому приемочному контролю.

От контроля к пополнению партии. Рассмотрим простую идею: отказываемся от контроля качества вообще, но зато по первому требованию потребителя заменяем дефектную единицу продукции на новую. При этом экономим на контроле, но вместо этого тратим средства на замену продукции. Выгодно это или не выгодно?

Замена продукции может проводиться различными способами. Для многих видов товаров народного потребления это делается с помощью системы гарантийного обслуживания, гарантийных сроков и мастерских, через сеть розничной торговли и т.д.

Другой вариант — к партии поставляемой продукции добавляется некоторое количество единиц продукции для замены имеющихся, возможно, в ней дефектных единиц. Сначала обсудим подробнее именно этот вариант идеи замены продукции.

Пусть поставщик выпускает продукцию с известным ему уровнем дефектности p. Тогда число X дефектных единиц в партии объема N имеет биномиальное распределение с параметрами N и p. По теореме Муавра-Лапласа X не превосходит (при достаточно большом N) величины

$$D_0(t) = Np + t(Np(1-p))^{1/2}$$

с вероятностью Φ(t). где Φ(.) — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Поскольку Φ(4) = = 0,999968329, то для практических целей достаточно положить t = 4, при этом более чем $D_0(4)$ дефектных единиц продукции попадет в партию лишь в 3 случаях из 100 000.

Пусть C_0 — цена одной единицы продукции, C_1 — стоимость неразрушающего контроля одной единицы продукции (с исправлением дефектов при их обнаружении). Сравним сначала две стратегии технико-экономических отношений поставщика с потребителями:

1) сплошной контроль (затраты C_1N);

2) пополнение партии дополнительными изделиями в числе $D_0(4)$ (затраты $C_0D_0(4)$). Вторая стратегия лучше (экономически выгоднее), если

$$C_1N > C_0D_0(4) = C_0(Np + 4\sqrt{Np(1-p)}). \qquad (16)$$

Поделим на C_0N, получим равносильное неравенство

$$\frac{C_1}{C_0} > p + 4\frac{\sqrt{p(1-p)}}{\sqrt{N}}.$$

Поскольку $p(1\text{-}p)$ не превосходит 1/4 при всех p, то из неравенства

$$C_1/C_0 > p + 2 / N^{1/2} \qquad (17)$$

вытекает неравенство (16). Ясно, что в случае, если

$$C_1/C_0 > p,$$

неравенство (17) (а потому и неравенство (16)) выполняется при достаточно больших объемах партии, а именно, при

$$N > \{2\,C_0 / (C_1 - C_0\,p)\}^2.$$

Например, если стоимость контроля составляет 10 % от стоимости продукции (типовая ситуация в машиностроении), т.е. C_1/C_0 = 0,1, а уровень дефектности p = 0,01, то последнее неравенство дает N>493. В то же время нетрудно проверить, что неравенство (16) выполняется при

$$0{,}1 > 0{,}01 + 4\,(0{,}01 \cdot 0{,}99)^{1/2} / N^{1/2},$$

т.е. при $N > 19$. Расхождение более чем на порядок (в 26 раз) объясняется заменой при переходе от формулы (16) к формуле (17) величины $p(1 - p)$ на 1/4, т.е. на гораздо большую величину — при малом входном уровне дефектности p.

Выгодно ли введение статистического контроля? Пусть рассматривается описанная выше стратегия пополнения партий. Мы сравнивали ее со стратегией сплошного контроля, которая во многих случаях оказалась хуже. Может быть, поставщику имеет смысл использовать статистический контроль? Понятно, что речь может идти лишь о (неразрушающем) контроле с разбраковкой, поскольку только в этом случае меняется доля дефектности в потоке партий, направляемых потребителям.

Пусть используется план (n,0) с приемочным уровнем дефектности, равным реально достигнутому предприятием уровню дефектности p. Как известно, тогда объем выборки определяется из условия

$$(1 - p)^n = 0{,}95,$$

т.е.

$$n = \ln 0{,}95 / \ln (1 - p) = -\,0{,}0513 / \ln (1 - p).$$

При малом p уже не раз применявшееся соотношение из математического анализа дает с достаточной для практики точностью

$$n = 0{,}05 / p.$$

С вероятностью $(1 - p)^n = 0{,}95$ партия принимается, с вероятностью 0,05 подвергается разбраковке. В первом случае партия поступает к потребителю с тем же уровнем дефектности, что и до контроля, но при этом добавляются затраты на контроль, равные $C_1 n$. Партию необходимо пополнить $D_0(4)$ изделиями (затраты $C_0 D_0(4)$), общие затраты (в среднем на одну выпущенную партию) равны

$$0{,}95\,(C_1 n + C_0\, D_0\,(4)).$$

Во втором случае фактически проводится сплошной контроль с исправлением дефектов и затратами C_1N. Суммарные затраты при использовании выборочного контроля равны

$$0{,}95\,(C_1 n + C_0\,D_0\,(4)) + 0{,}05\,C_1 N.$$

Он более выгоден, чем отсутствие контроля (с добавлением «запасных» изделий), в случае справедливости неравенства

$$0{,}95\,(C_1 n + C_0\,D_0\,(4)) + 0{,}05\,C_1 N < C_0\,D_0(4),$$

что эквивалентно неравенству

$$19\,C_1 n + C_1 N < C_0\,D_0(4).$$

Сравнение с формулой (16) показывает, что если контроль не является разрушающим, то выборочный контроль менее выгоден, чем сплошной. По сравнению с формулой (16) добавляется первое слагаемое в левой части последней формулы. И тем более выборочный контроль в экономической эффективности весьма проигрывает по сравнению с отсутствием контроля в сочетании с пополнением партии.

Итак, введение статистического контроля в схеме пополнения партии не выгодно.

От системы контроля к системе технического обслуживания. Вернемся к первому из указанных ранее вариантов замены продукции. Что выгоднее — сплошной контроль на предприятии или замена дефектных изделий, обнаруженных потребителями? Реальное перекладывание контроля на потребителей влечет потери, связанные с удовлетворением их претензий, но при малой доле дефектных изделий эти потери малы по сравнению с затратами на контроль.

Действительно, пусть W — средние потери поставщика, связанные с пропуском потребителю дефектной единицы продукции. Сюда входят, в частности, такие виды потерь:

- стоимость новой единицы продукции (при замене изделия или возврате его стоимости);

- расходы системы распределения продукции и гарантийного ремонта, включая издержки на устранение дефектов;

- потери из-за нежелательного изменения предпочтений потребителя, из-за снижения имиджа фирмы;

- затраты на возмещение ущерба, понесенного потребителем, страховые сборы, судебные издержки, и т.д.

Потери W в несколько раз (по экспертной оценке — обычно в 5–10 раз) превышают расходы C_0 на изготовление единицы продукции. Кроме того, для быстрого решения проблем потребителей, связанных с обнаружением дефектов, необходима развитая система технического обслуживания.

Пусть изготовлена партия продукции объема N. Тогда расходы на сплошной (неразрушающий) контроль составляют C_1N (при этом дефектные единицы продукции извлекаются и утилизируются, расходами на утилизацию или доходами от нее в настоящем изложении пренебрегаем). Пусть p — доля дефектных единиц продукции в партии. Тогда Np — математическое ожидание числа дефектных единиц продукции в партии, а WNp — математическое ожидание потерь. Если

$$WNp < C_1 N, p < C_1 / W, \quad (18)$$

то выгоднее отказаться от сплошного контроля. При повышении качества, т.е. снижении доли дефектности, целесообразно переходить к поиску и устранению дефектов не непосредственно на предприятии, а в пунктах системы технического обслуживания.

В формуле (18) участвует математическое ожидание WNp. Реальные потери могут быть больше, но ненамного. Как и выше, с помощью теоремы Муавра-Лапласа можно утверждать, что практически наверняка они не превышают $WD_0(4)$, а потому преимущество решения об отказе от контроля неоспоримо при

$$WD_0(4) < C_1N, p + 4(p(1-p))^{1/2}/N^{1/2} < C_1/W. \quad (19)$$

Аналогично выводу неравенства (17) заключаем, что неравенство (19) наверняка будет выполнено, если

$$p + 2/N^{1/2} < C_1/W. \quad (20)$$

Пусть $C_1/W = 0{,}1$, выпускается партия объема $N = 1\ 600$. Тогда согласно неравенству (20) отказ от контроля выгоден уже при $p < 0{,}05$, т.е. граничное значение соответствует довольно низкому уровню качества — 1 единица продукции из 20.

Выгодно ли в рассматриваемой ситуации вводить выборочный контроль? Пусть объем контроля равен n, приемочное число $c = 0$, с вероятностью y партия принимается, а с вероятностью $1 - y$ бракуется (и затем подвергается разбраковке). В первом случае расходы на контроль равны C_1n, а остальная часть партии содержит в среднем $(N - n)p$ дефектных единиц продукции, и средние издержки равны $y\{C_1n + W(N - n)p\}$. Во втором случае суммарные затраты равны $(1 - y)C_1N$. Следовательно, введение контроля выгодно, если

$$y\{C_1n + W(N - n)p\} + (1 - y)C_1 N < WNp.$$

Преобразуем это неравенство к виду

$$yn\{C_1 - Wp\}(1 - y)^{-1} + C_1N < WNp. \qquad (21)$$

Если выполнено неравенство $p<C_1/W$, то второе слагаемое в левой части неравенства (21) больше правой части этого неравенства, в то время как первое слагаемое в левой части (21) положительно. Следовательно, неравенство (21) неверно, и введение выборочного контроля нецелесообразно — как и в разобранном ранее случае метода пополнения партий.

Выше приведен базовый (простейший, исходный) метод сравнения различных систем взаимоотношений поставщиков и потребителей. При разработке практически пригодных систем принятия решений целесообразно дальнейшее его развитие.

Отметим в заключение, что реально статистический контроль качества продукции, осуществляемый поставщиком (выходной контроль), решает две основные задачи: обеспечение интересов потребителя и обнаружение разладок собственных технологических процессов (по результатам контроля последовательности партий). Как показано выше, для решения первой из этих задач он не всегда оптимален. Вторую из названных задач также часто эффективнее решать с помощью иных методов, например, обнаруживать разладку технологических процессов с помощью тех или иных контрольных карт. Таким образом, область применения методов статистического приемочного контроля является довольно ограниченной. Очевидно, однако, что нельзя исключать эти методы из арсенала менеджеров по качеству, в частности, при использовании концепции «всеобщего управления качеством (*TQM — Total Quality Management*)». Хотя бы потому, что они незаменимы при использовании разрушающих методов контроля.

Наиболее перспективным представляется использование полученных результатов в рамках концепции контроллинга (см., например, [11–13]). Итак, выше сформулирован основной парадокс теории статистического приемочного контроля — повышение качества выпускаемой продукции приводит к увеличению объема контроля. Описан способ разрешения этого парадокса — на основе перехода от чисто технической политики выбора плана контроля к технико-экономической. Она исходит из сравнения по экономическим показателям схем контроля и схем технического обслуживания и пополнения партий. Проанализирован базовый метод такого сравнения, позволяющий выделить область экономического преимущества схемы пополнения партий и схемы технического обслуживания по сравнению со схемой контроля.

10.4. Статистические методы управления качеством

Как уже отмечалось, в России намечается все расширяющаяся тенденция к сертификации продукции, т.е. к официальной гарантии поставки производителем продукции, удовлетворяющей установленным требованиям. Поставщики и продавцы должны иметь сертификаты качества на предлагаемые ими товары и услуги. Маркетинг, т.е. производственная и коммерческая политика, нацеленная на получение максимальной прибыли на основе изучения рынка, создания конкурентоспособной продукции и ее полной реализации, включает в себя работы по сертификации [2].

Не будем останавливаться на быстро меняющейся организационной стороне процесса сертификации и соответствующих отечественных и зарубежных нормативных документах, а также на различных системах сертификации. Как общие проблемы сертификации, так и выбор схемы сертификации для конкретной продукции активно обсуждаются специалистами. Приведем лишь несколько замечаний, необходимых для дальнейшего изложения.

Напомним, что, говоря о сертификации продукции, могут иметь в виду качество конкретной ее партии. В ряде случаев это оправдано — рядового потребителя интересует качество лишь той единицы продукции, которую он приобрел. Однако установление долговременных хозяйственных связей целесообразно лишь в случае, когда поставщик гарантирует высокое качество не одной, а всех партий своей продукции. Другими словами, должны быть проведены оценка и сертификация технологических процессов и производств.

Еще больше повышается доверие к поставщику, если не только отдельные технологические процессы, но и все предприятие в целом гарантированно выпускает продукцию высокого качества. В современных условиях это обеспечивается действующей на предприятии системой качества, удовлетворяющей требованиям Международной организации по стандартизации ИСО.

Одна из основных характеристик товара — его конкурентоспособность. Очевидно, производителю необходимо уметь оценивать конкурентоспособность перед запуском продукции в производство или началом работы по продвижению на зарубежный рынок (подробнее см. рекомендации [2]). Следует отметить, что в литературе имеются различные мнения по поводу понятия «конкурентоспособность». В частности, нельзя согласиться с крайне упрощенным подходом в монографии [14], согласно которому конкурентоспособность сводится к соотношению цен на внутреннем и внешнем рынках. Достаточно напомнить о таких приемах конкурентной борьбы, как демпинг и (добросовестная или недобросовестная) реклама, таможенные пошлины и квоты.

Одним из основных компонентов конкурентоспособности продукции является ее технический уровень. В западных публикациях справедливо отмечается, что фирма, обладающая патентом или новой научно-технической разработкой, имеет более высокий «излишек производителя» по сравнению с другими фирмами (см., например, [15–17]). В частности, согласно одному из наиболее популярных западных учебников [18] технический уровень продукции — одна из основных учитываемых характеристик при выборе направления инвестиционных вложений.

Из сказанного вытекает, что сертификация — это современная форма управления качеством. Среди зарубежных специалистов общепринято, что основная составляющая в управлении качеством продукции — это статистические методы (см., например, отчет Комитета ИСО по изучению принципов стандартизации [3]). В нашей стране внедрение комплексных систем управления качеством (КС УКП), к сожалению, сводилось во многом всего лишь к подготовке документации организационного характера. Статистические методы использовались в промышленности недостаточно, прежде всего из-за недостаточной подготовки кадров, а государственные стандарты по этой тематике зачастую содержали грубейшие ошибки (см. ниже). Ситуация в области применения статистических методов и причины нашего отставания достаточно подробно разобраны в публикациях [4, 19].

Как отмечалось в начале главы, более 150 лет статистические методы применяются в России для проверки соответствия продукции установленным требованиям, т.е. для сертификации. С начала 1970-х гг. стали разрабатываться государственные стандарты по статистическим методам. В связи с обнаружением в них грубых ошибок в 1985 г. была организована «Рабочая группа по упорядочению системы стандартов по прикладной статистике и другим статистическим методам». В ее работе приняли участие 66 специалистов, в том числе 15 докторов и 36 кандидатов наук. Выводы Рабочей группы[5] кратко отражены в статьях [4, 19]. В соответствии с рекомендациями Рабочей группы 24 из 31 государственного стандарта по статистическим методам были отменены в 1986–1987 гг.

К сожалению, потеряв правовую силу как нормативные документы, ошибочные стандарты до настоящего времени продолжают использоваться отдельными «специалистами» как научно-технические издания. Полученные Рабочей группой результаты и выводы не были широко и подробно опубликованы, ошибки в государственных стандартах не были публично вскрыты, и авторы дальнейших публикаций продолжают ссылаться на издания с грубейшими ошибками. Так, в ряде работ пропагандируются ошибочные стандарты, посвященные применению контрольных карт при статистическом регулировании технологических процессов. Продолжает широко использоваться грубо ошибочный ГОСТ 11.006-74 (СТ СЭВ 1190-78) «Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим» (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация), хотя разбору ошибок в этом стандарте посвящена статья [20] двадцатилетней давности (см. также главу 2). Перечисленные факты делают целесообразным популяризацию результатов и выводов Рабочей группы и в настоящее время.

С точки зрения теории и практики принятия решений рассматриваемая ситуация весьма поучительна. Она показывает, что недостаточно продуманная система разработки управленческих документов (в рассматриваемом случае — стандартов по статистическим методам управления качеством продукции) позволяет отдельным лицам бесконтрольно определять содержание нормативно-технических документов, в том числе на десятилетия закреплять в них ошибочные положения. Не так уж важны мотивы поведения подобных лиц — невежество в сочетании с самонадеянностью, корысть или стремление нанести вред действующей научно-технической системе — важен результат.

В 1988–1989 гг. наиболее активная часть Рабочей группы (10 докторов и 15 кандидатов наук) составили «Аванпроект комплекса методических документов и пакетов программ по статистическим методам стандартизации и управления качеством». Это обширное сочинение (около 1 600 стр.) и на настоящий момент является наиболее полным руководством по рассматриваемой тематике. Информация о нем приложена к переводу книги японских авторов по аналогичной тематике [1].

---

[5] Автор учебника был ученым секретарем Рабочей группы.

К сожалению, под влиянием авторов ошибок в стандартах Госстандарт принял решение не финансировать реализацию заказанного им «Аванпроекта». Тогда сообщество специалистов решило действовать самостоятельно. В 1989 г. был организован Центр статистических методов и информатики (ЦСМИ; в настоящее время — Институт высоких статистических технологий и эконометрики). К середине 1990 г. в ЦСМИ были разработаны 7 диалоговых систем по современным статистическим методам управления качеством, а именно, СПК, АТСТАТ-ПРП, СТАТКОН, АВРОРА-РС, ЭКСПЛАН, ПАСЭК, НАДИС (описания этих систем приведены в работе [5]). В работе участвовали 128 специалистов. В дальнейшем к ЦСМИ присоединялись новые группы научно-технических работников. К концу 1991 г. число сотрудников ЦСМИ превысило 300. Информация о программных продуктах и другой деятельности ЦСМИ постоянно помещалась в журналах «Заводская лаборатория» и «Надежность и контроль качества». Программные продукты, разработанные Центром статистических методов и информатики, были поставлены и использовались более чем в 100 организациях и предприятиях. Среди них — производственные объединения «Уралмаш», «АвтоВАЗ», «Пластик», ЦНИИ черной металлургии им. Бардина, НИИ стали, ВНИИ эластомерных материалов и изделий, НИИ прикладной химии, ЦНИИ химии и механики, НПО «Орион», НИЦентр по безопасности атомной энергетики, ВНИИ экономических проблем развития науки и техники, ВНИИ нефтепереработки, МИИТ, Казахский политехнический институт, Ульяновский политехнический институт, Донецкий государственный университет и др.

Как уже отмечалось, параллельно ЦСМИ вел работу по объединению статистиков. В апреле 1990 г. в Большом Актовом Зале Московского Энергетического института прошла Учредительная конференция Всесоюзной организации по статистическим методам и их применениям. На Учредительном съезде Всесоюзной статистической ассоциации (ВСА) в октябре 1990 г. в Московском экономико-статистическом институте эта организация вошла в состав ВСА в качестве секции статистических методов (подробнее о создании и задачах ВСА рассказано, например, в статьях [21, 22]). В 1992 г. после развала СССР и фактического прекращения работы ВСА на основе секции статистических методов ВСА организована Российская ассоциация по статистическим методам (РАСМ), а затем и Российская академия статистических методов. В мероприятиях секции статистических методов ВСА и РАСМ активно участвовали несколько сот специалистов по статистическим методам. В ЦСМИ и РАСМ, объединивших большинство ведущих российских специалистов, коллективными усилиями разработан единый подход к проблемам применения статистических методов в сертификации и управлении качеством.

Классификация статистических методов сертификации. Рассмотрим два основания для классификации. Первый — по виду статистических методов. Второй — по этапам жизненного цикла продукции, на которых соответствующий метод применяется. Первое основание привычно для специалистов по разработке статистических методов и соответствующего программного обеспечения, второе — для тех, кто эти методы применяет на конкретных предприятиях.

В ЦСМИ сложилось пятичленное деление по первому основанию (в скобках указаны наименования диалоговых систем ЦСМИ, рассмотренных, в частности, в работе [5]):

а) прикладная статистика — иногда с дальнейшим выделением статистики случайных величин, многомерного статистического анализа, статистики случайных процессов и временных рядов, статистики объектов нечисловой природы (Система Регрессионного Статистического Моделирования СРСМ, или СТАТМАСТЕР; АДДА, ГРАНТ, КЛАМС, ЭКОНОМЕТРИК, РЕГРЕССИЯ, ЛИСАТИС, ЭКОСТАТ, РЕСТ);

б) статистический приемочный контроль (СПК, АТСТАТ-ПРП, КОМПЛАН);

в) статистическое регулирование технологических процессов, в частности, методом контрольных карт (СТАТКОН, АВРОРА-РС);

г) планирование эксперимента (ПЛАН, ЭКСПЛАН, ПАСЭК, ПЛАНЭКС);

д) надежность и испытания (НАДИС, ОРИОН, СЕНС).

Быстрое развитие компьютерной техники имеет свою оборотную сторону. Вполне добротные программные продукты устаревают и выходят из обращения просто потому, что они сделаны на отработавшем свой срок операционном (системном) программном обеспечении. «Выжить» может только то программное обеспечение, которое поддерживается соответствующей фирмой и постоянно совершенствуется с чисто программистской точки зрения. Важна система технической поддержки, обучение и, конечно, реклама. При этом чисто научная сторона дела отходит на задний план. Эти простые соображения объясняют, почему за 30 лет (с 1991 г. по 2021 г.) отечественный рынок программных продуктов по эконометрике и статистическим методам стал гораздо более бедным по числу продуктов, научный уровень явно понизился, зато дизайн явно стал более привлекательным.

Перейдем ко второму основанию классификации методов сертификации. Согласно п. 5.1 «Петля качества» стандарта ИСО 9004-87 2Общее руководство качеством и элементы системы качества. Руководящие указания» система качества функционирует «... одновременно со всеми остальными видами деятельности, влияющими на качество продукции или услуг, и взаимодействует с ними.

Ее воздействие распространяется на все этапы от первоначального определения и до конечного удовлетворения требований и потребностей потребителя. Эти этапы и виды деятельности включают:

1) маркетинг, поиски и изучение рынка;

2) проектирование и/или разработку технических требований, разработку продукции (опытного образца);

3) поиски поставщиков и оптовых покупателей, организацию материально-технического снабжения (решение задач логистики);

4) подготовку и разработку производственных (технологических) процессов;

5) непосредственно производство продукции;

6) контроль качества продукции, проведение испытаний и обследований;

7) упаковку и хранение продукции;

8) реализацию (сбыт) и распределение (доставку) продукции;

9) монтаж и эксплуатацию продукции у потребителей;

10) технические помощь и обслуживание;

11) утилизацию после использования.

Подробное рассмотрение применения основных типов статистических методов на перечисленных этапах жизненного пути продукции не входит в задачу настоящей книги. Сводка, приведенная в табл. 1, показывает, что статистические методы широко применяются на всех этапах жизненного пути продукции.

Таблица 1

Применение статистических методов на различных этапах жизненного цикла продукции по ИСО 9004-87

Номер этапа	а	б	в	г	д	Специальные модели
1	+	–	–	+	–	+
2	+	–	–	+	+	+
3	+	–	–	–	–	+
4	+	+	+	+	+	+
5	+	+	+	+	–	+
6	+	+	+	+	+	+
7	+	+	+	+	+	+
8	+	+	–	–	–	+
9	+	+	+	+	+	+
10	+	–	–	–	–	+
11	+	+	+	+	–	+

Помимо компьютерных диалоговых систем широкого назначения, на каждом конкретном предприятии и на любом конкретном этапе жизненного пути продукции могут быть использованы специальные модели, например, на этапе 3 «материально-техническое снабжение» — модели управления запасами (см. о них, например, главу 5 монографии [23]).

Среди диалоговых систем по статистическому анализу выделим пакеты, ориентированные:

- на восстановление зависимостей (СТАТМАСТЕР, он же СРСМ — Система Регрессионного Статистического Моделирования, и его развитие ЭКОНОМЕТРИК, а также РЕГРЕССИЯ);
- анализ нечисловых данных на основе методов статистики объектов нечисловой природы (АДДА, КЛАМС, а также ориентированный на экспертное оценивание ГРАНТ, на анализ интервальных данных РЕСТ);
- прогнозирование (ЛИСАТИС и его развитие ЭКОСТАТ, а также относящиеся к временным рядам разделы пакета АВРОРА-РС — Анализ Временных Рядов и Обнаружение РАзладки).

Для регулярного и обоснованного принятия решений на основе решения обширных комплексов задач сертификации и управления качеством на конкретном предприятии в ряде случаев целесообразно создать диалоговую систему, предназначенную для использования именно на этом предприятии. В частности, для решения задач этапа 4 используют созданные для конкретного предприятия программные системы, соединяющие в себе банки данных и пакеты статистических методов анализа этих данных. Примерами являются «Автоматизированное рабочее место материаловеда (АРМ материаловеда)» и «Автоматизированное рабочее место математика (АРМ математика)», разработанные Центром статистических методов и информатики для ВНИИ эластомерных материалов и изделий.

Для объединения типовых пакетов в индивидуальную систему полезно программное средство ИНТЕГРАТОР — универсальный инструмент, предназначенный для создания интегрированных программных систем и обеспечивающий возможность совместного использования различных пакетов прикладных программ на персональных компьютерах IBM PC. Так, с помощью ИНТЕГРАТОРА был разработан АРМ математика на основе пакетов СРСМ, ПЛАН, АТСТАТ-ПРП, соответствующей базы данных и ряда программ, ориентированных на специфику ВНИИ эластомерных материалов и изделий.

На всех этапах жизненного цикла продукции, особенно на этапах 3, 8, 10, часто используют специализированные вероятностно-статистические модели, в том числе модели управления запасами (см., например, монографию [23, гл. 5]), массового обслуживания и др. Такие модели и их программное обеспечение, как правило, разрабатываются для конкретного предприятия, и потому они хорошо приспособлены к особенностям этого предприятия.

Как избежать ошибок в нормативно-технической документации и инструктивно-методической документации? Как уже отмечалось, многие ошибочные государственные стандарты по статистическим методам управления качеством были отменены (хотя результаты анализа, проведенного Рабочей группой, так и не были вовремя и полностью опубликованы). Однако эти стандарты продолжают и до сих пор использоваться как авторитетные научно-технические публикации. Почему так происходит? Как вообще могли появиться ошибки в нормативно-технических документах и почему в течение ряда лет эти документы использовались, несмотря на очевидные для специалистов ошибки?

Дело в том, что инженеру, экономисту, менеджеру, работнику прикладной науки (короче — инженеру) несвойственно менять свою специальность, становиться математиком и самостоятельно повторять выкладки и рассуждения, положенные в основу ГОСТа. Поэтому инженер обычно не может самостоятельно обнаружить математические ошибки в ГОСТе, даже грубейшие. Главное — он не хочет этим заниматься. С другой стороны, математику несвойственно анализировать нормативно-техническую документацию. Он также обычно не хочет этим заниматься. «Рабочая группа по упорядочению системы стандартов по прикладной статистике и другим статистическим методам» была уникальным примером совместной работы математиков и инженеров, именно поэтому ей удалось сопоставить тексты стандартов с результатами современной науки.

Вполне естественно, что виновные в ошибках сделали все возможное в рамках действовавшей системы принятия решений, чтобы помешать признанию и исправлению допущенных ими ошибок в государственных стандартах. До сих пор продолжаются попытки навязать промышленности (в качестве нормативных документов!) тексты, грубая ошибочность которых давно установлена. Кроме того, государственные стандарты, отмененные как нормативно-технические документы, продолжают физически существовать как издания (брошюры) и использоваться при проведении инженерных расчетов, проектировании систем контроля и т.д. Все это делает необходимым пропаганду выводов Рабочей группы относительно ГОСТов по статистическим методам. Они приведены в работе [4].

«Шесть сигм» — система внедрения контроллинга и его статистических инструментов. Как улучшить организацию производства? Как повысить эффективность управления? Волна за волной накатывают на руководителей и специалистов все новые сочетания слов и стоящие за ними концепции. И в каждой волне есть что-то новое и что-то давно известное. Основное — новый угол взгляда на старые проблемы и методы.

И вот появилось новое модное поветрие — «Шесть сигм». Что стоит за этими словами, наводящими на мысли о статистических методах (греческой буквой «сигма» традиционно обозначают показатель разброса статистических данных)? Как сказано в [26], «Шесть сигм» — это более разумный способ управлять всей компанией или отдельным подразделением. Фактически речь идет о развитии системы контроллинга на предприятии, в организации, фирме, компании. Концепция «Шесть сигм» ставит на первое место потребителя и помогает находить самые лучшие решения, опираясь на факты и данные. Она нацелена на три основные задачи:

- повысить удовлетворенность клиентов;
- сократить время цикла (производственного, операционного);
- уменьшить число дефектов.

Внедрение «Шести сигм» дает значительный экономический эффект. Исполнительный директор *General Electric* Джек Уэлч писал в ежегодном докладе, что всего за три года «Шесть сигм» сэкономили компании более 2 миллиардов долларов.

Совершенно справедливо «Шесть сигм» рассматривают как «революционный метод управления качеством». Согласно «Шести сигмам» следует стремиться к достижению самого малого разброса контролируемого параметра по сравнению с полем допуска. Желательно добиться, чтобы ширина поля допуска была в 6 раз больше типового разброса, традиционно описываемого как «плюс-минус сигма». Отсюда и название концепции — «Шесть сигм». Соотношение поля допуска с полем разброса (в «сигмах») связывают с числом дефектов (на миллион возможностей) и с выходом годной продукции (в %). Так, 6 «сигма» соответствуют 3,4 дефектов на 1 000 000 возможностей, или выходу годной продукции 99,99966 %. А пока такой высокий уровень не достигнут, можно оценивать ситуацию в «сигма». И промежуточная задача может формулироваться так: с уровня 2,5 «сигма» подняться до уровня 4 «сигма».

С помощью каких инструментов достигается успех в системе «Шести сигм»? Это инструменты генерации идей и структурирования информации — экспертные оценки (голосования, мозговой штурм), диаграммы (сродства, древовидные, «рыбий скелет» — схема Исикава), блок-схемы. Это инструменты сбора данных — выборочный метод, методики измерений, методы определения «голоса потребителя», контрольные листки и электронные таблицы. Третья группа — инструменты анализа процесса и данных — анализ течения процесса, добавленной ценности, различные графики и диаграммы. В том числе диаграмма Парето, график временного ряда (тренда), диаграмма разброса (поле корреляции). Затем — инструменты статистического анализа (проверка статистических гипотез, методы корреляции и регрессии, планирования экспериментов и др.). Наконец, четвертая группа инструменты реализации решений и управления процессом. Среди них — методы управления проектами (планирование, бюджетирование, составление графиков, коммуникации, управление коллективом, диаграммы Ганта и др.), анализ потенциальных проблем и анализ видов и последствий отказов, анализ заинтересованных сторон, диаграмма поля сил, документирование процесса, сбалансированная система показателей и «приборная» панель процесса.

Инструментарий системы «Шести сигм» весьма широк. Эти инструменты помогают принимать правильные решения, решать проблемы и управлять переменами. Среди них, как следует из проведенного выше перечисления, основное место занимают различные статистические инструменты. Однако нельзя считать, что система «Шести сигм» и инструменты «Шести сигм» — это одно и то же.

Как справедливо подчеркнуто в цитированной книге о системе «Шести сигм», возможно, вы говорите себе: «Мы уже делаем кое-что из этого». И уж, безусловно, читали почти обо всем из названных выше инструментов, в том числе на страницах настоящего учебника. Совершенно бесспорно, что многое в концепции «Шести сигм» не ново. Что действительно ново — так это *соединение всех этих элементов системы и ее инструментов в согласованный процесс управления.*

Действительно, различные виды инструментов повышения эффективности управления известны давно. Чтобы их успешно использовать, НУЖНА СИСТЕМА ВНЕДРЕНИЯ. Нужна тщательно разработанная методика создания и функционирования творческих коллективов, занимающихся анализом ситуации, подбором и внедрением современных инструментов управления. Все это создано в системе «Шести сигм». В этом и состоит суть нового шага в науке и практике управления предприятием.

Выделяют шесть элементов, составляющих квинтэссенцию системы «Шесть сигм»:

- ориентация на потребителя;

- управление на основе данных и фактов;
- процессный подход (где действия, там и процессы);
- проактивный менеджмент (т.е. основанный на прогнозировании);
- безграничное сотрудничество;
- стремиться к совершенству, но не бояться поражений.

Конечно, каждый из этих элементов сам по себе хорошо известен. Дело в системе «Шести сигм», в которую они объединены. В этой системе подробно расписаны роли различных участников команды — «черные пояса», «зеленые пояса», «мастера черных поясов», «чемпионы». Подчеркивается основополагающая роль членов руководства компании («спонсоров»), лично занимающихся развитием системы «Шесть сигм».

Анализ системы «Шесть сигм» показывает, что, несмотря не некоторое различие терминов, связанное с корнями этой системы (лежащими в проблемах управления качеством), фактически «Шесть сигм» — это глубоко проработанная система внедрения современного контроллинга. Отметим большое место, которое занимают статистические методы среди ее инструментов. Система «Шесть сигм» трудоемка, на внедрение нужны годы. Но и эффект велик [27]. Можно взглянуть на систему «Шесть сигм» и как на инструмент инновационного менеджмента. Тогда естественно рассматривать «Шесть сигм» как новую систему внедрения математических методов исследования и управления на промышленном предприятии.

10.5. Обнаружение разладки с помощью контрольных карт[6]

Как уже отмечалось, контрольные карты Шухарта и кумулятивных сумм первоначально использовались для статистического контроля технологических процессов. В последние годы область их применения значительно расширилась, охватив временные ряды различной природы, от экологического мониторинга до анализа динамики экономических показателей. В разделе даются общие сведения о методе контрольных карт и подробно разбирается пример, относящийся к деятельности предприятия ОАО «Северсталь-авто».

Области применения контрольных карт. Контрольные карты предназначены для организации и проведения статистического контроля стохастических процессов, представленных в форме временных рядов, на предмет наискорейшего обнаружения момента спонтанного изменения (разладки) вероятностных характеристик контролируемых процессов. Во многих практических зада-

[6] Раздел 10.5 написан совместно с И.Н. Митрохиным, ОАО «Северсталь-авто», Москва, Россия.

чах только на основании анализа временных рядов возможно выявить скрытые изменения, происходящие в наблюдаемом объекте или явлении.

Методы обнаружения разладки имеют многообразные области применения: контроль качества технологических процессов, обнаружение аварийных ситуаций, обработка информации в автоматизированных системах научных исследований (АСНИ), в автоматизированных системах управления технологическими процессами (АСУ ТП), в автоматизированных испытательных стендах. Текущий контроль систем подвижных объектов, обнаружение сейсмических сигналов на фоне шума, оперативное выявление отказов агрегатов без остановки технологического процесса, статистическое регулирование технологических процессов с помощью контрольных карт, мониторинг экологической обстановки, интенсивное наблюдение за тяжелобольными в клиниках — вот лишь краткий перечень задач, которые решают с помощью контрольных карт [1, 10, 28, 29].

Основные идеи метода контрольных карт. Задача статистического регулирования технологического процесса состоит в том, чтобы на основании результатов периодического контроля выборок малого объема принимать решение «процесс налажен» или «процесс разлажен». Соответствующая вероятностно-статистическая модель в простейшем варианте следующая. Выдвигаются две гипотезы: нулевая гипотеза H_0—технологический процесс налажен, если параметр (характеристика) θ распределения контролируемого показателя качества X равен θ_0, и альтернативная гипотеза H_1—технологический процесс разлажен, если параметр θ равен θ_1. На основании результатов контроля X_1, X_2, ..., X_n , включенных в выборку единиц продукции, необходимо с помощью определенных статистических критериев принять одну из этих двух гипотез.

Обычно [24] выборку значений контролируемого показателя качества моделируют совокупностью независимых одинаково распределенных случайных величин с математическим ожиданием μ и дисперсией σ^2. При статистическом регулировании технологических процессов проверяют гипотезы:

$$H_0 : \mu = \mu_0 \text{ (технологический процесс налажен),}$$
$$H_1 : \mu = \mu_1 \text{ (технологический процесс разлажен),}$$

если разладка связана с изменением математического ожидания μ. Если же разладка связана с увеличением дисперсии σ^2, то в этом случае проверяют гипотезы:

$$H_0 : \sigma = \sigma_0 \text{ (технологический процесс налажен),}$$
$$H_1 : \sigma = \sigma_1 \text{ (технологический процесс разлажен).}$$

На стадии предварительного анализа состояния технологического процесса необходимо оценить параметры μ_0 и σ_0. Для этого следует отобрать на контроль определенное количество единиц продукции при нормальном ходе производства, т.е. при надлежащем качестве сырья и при отлаженном оборудовании. При этих условиях мы по правилам прикладной статистики [30] получим оценки параметров μ и σ при налаженном состоянии технологического процесса, т. е. μ_0 и σ_0.

Значения параметров, соответствующие разлаженному процессу, либо задают с помощью экспертов, либо определяют при анализе забракованной продукции.

На контрольной карте отмечают границы регулирования, ограничивающие область допустимых значений статистики (при справедливости нулевой гипотезы). Контрольная карта является наглядным графическим средством, отражающим состояние технологического процесса. Выход точки за границу регулирования служит сигналом о разладке технологического процесса.

Впервые контрольные карты были разработаны в 1924 г. американцем У.А. Шухартом, работавшим в *Bell Telephone Laboratories* [1]. Методология построения контрольных карт Шухарта имеет ту особенность, что проводимый контроль не учитывает предыдущего поведения контролируемого параметра. В середине XX века была разработана методика, учитывающая информацию о прошлых данных для анализа текущего состояния. В основе этой методики лежит исследование не индивидуальных значений признака, а учет их кумулятивных сумм. Такой подход получил название метода кумсум (или кумсум-метода).

Кумсума образуется следующим образом:

$$S_r = \sum_{i=1}^{r}(x_i - k) = S_{r-1} + (x_r - k), \tag{22}$$

где k — константа, представляющая собой некоторое эталонное значении; x_i — значение используемой статистики в i-й момент времени. Нанесенные на график в порядке появления кумулятивные суммы (S_r) образуют кумсум-карту.

В модели, когда разладка связана с изменением математического ожидания, величину k обычно приравнивают математическому ожиданию в налаженном состоянии. Если среднее значение параметра процесса возрастет, то будет иметь место и общий рост уровня кумсуммы, поскольку все большее число значений ($x_i - k$) станут положительными. Аналогичным образом, если среднее значение будет падать, то и график будет стремиться вниз. Таким образом изменение среднего значения исходных данных приводит к изменению угла наклона графика кумсумм.

Пример практического применения. На примере контроля премиального фонда (ПФ) аппарата управления производственного объединения покажем, какой контроль можно осуществлять, применив карты Шухарта и карты кумулятивных сумм.

Исходные данные — ПФ за 45 месяцев (табл. 2). Причем ПФ за первые 3 года считается стабильным. Проверим, не разлажен ли процесс выплаты ПФ, при условии, что норматив ПФ в течение всего контролируемого периода не изменяется. На рис. 1 показана динамика ПФ.

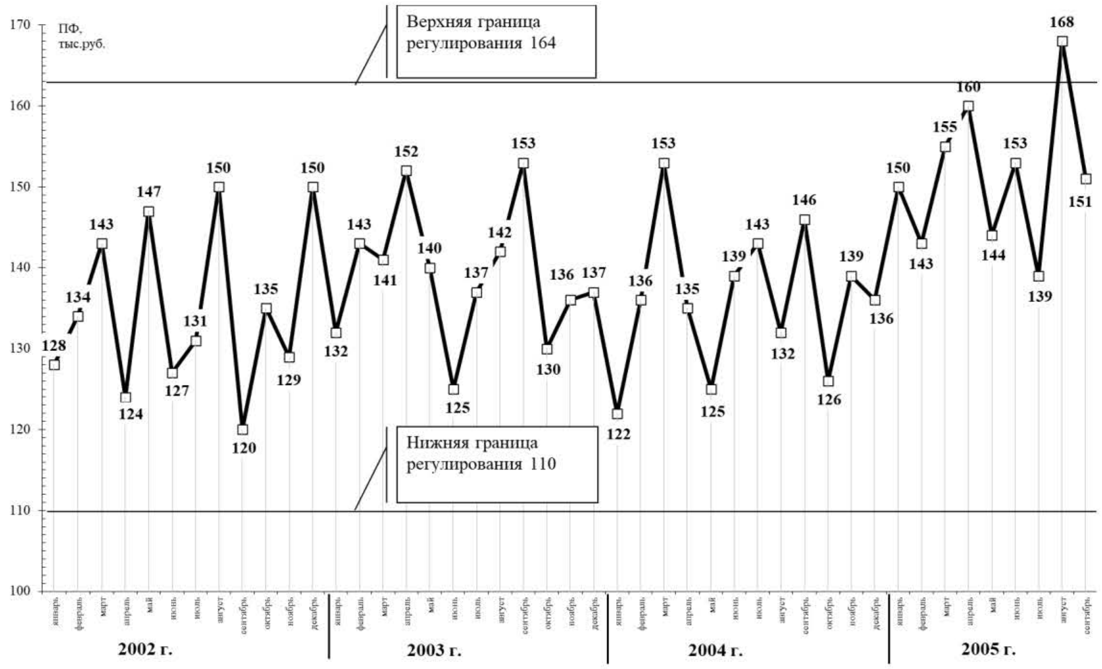


Рис. 1. Контрольная карта Шухарта для ПФ

Таблица 2

Премиальный фонд по месяцам (млн руб.)

Год	янв	фев	мар	апр	май	июн	июл	авг	сен	окт	ноя	дек
X	128	134	143	124	147	127	131	150	120	135	129	150
X+1	132	143	141	152	140	125	137	142	153	130	136	137
X+2	122	136	153	135	125	139	143	132	146	126	139	136
X+3	150	143	155	160	144	153	139	168	151	–	–	–

На основании табличных данных слабо прослеживается изменение среднего уровня ПФ в X+3 году. Построим карту Шухарта. Делаем предположение, что с X по X+2 год норматив ПФ за месяц не менялся. Для расчета выборочного среднего арифметического значения и выборочной дисперсии воспользуемся данными X–X+2 гг., поскольку в эти периоды ПФ считается стабильным, управляемым.

Среднее значение выборки рассчитываем по формуле:

$$\overline{ПФ} = \frac{ПФ_1 + ПФ_2 + ... + ПФ_{36}}{36} = \frac{128 + 134 + ... + 136}{36} = 137 \text{ млн руб.}$$

Выборочную дисперсию рассчитываем по формуле:

$$\sigma^2 = \frac{1}{36}\sum_{i=1}^{36}\left(\overline{ПФ} - ПФ_i\right) = \frac{1}{36}\left((137 - 128) + (137 - 134) + ...(137 - 136)\right) = 80{,}8$$

Выборочное среднеквадратическое отклонение таково:

$$\sigma = \sqrt{\sigma^2} = 9{,}0 \text{ млн руб.}$$

Будем использовать контрольную карту Шухарта, в которой границы регулирования отстоят от значения, соответствующего налаженному процессу, на 3 среднеквадратических отклонения. Тогда допустимые значения премиального фонда будут лежать в диапазоне $\left[\overline{МД} - 3\sigma; \overline{МД} + 3\sigma\right]$, нижняя граница регулирования ПФ составит: 137 – 9 · 3 = 110 (млн руб.), а верхняя — 137 + 9 · 3 = = 164 (млн руб.).

На основании расчетов построим контрольную карту премиального фонда (рис. 1). Анализ карты показывает, что в августе X+3 г. произошла разладка процесса, премиальный фонд превысил границу верхнюю границу регулирования. Необходимо выяснить причину разладки и принять решение о дальнейших действиях.

Вариация значений ПФ затрудняет визуальный анализ контрольной карты и не позволяет понять тенденцию изменения среднего значения ПФ и вовремя заметить разладку. Решить эту проблему помогают карты кумулятивных сумм. Найдем кумулятивные суммы. Эталонное значение *k* выберем, как выборочное среднее арифметическое: $k = \overline{ПФ} = 137$ млн руб.

Далее рассчитываем кумулятивные суммы по формуле (22) (табл. 3). Карта кумсумм представлена на рис. 2.

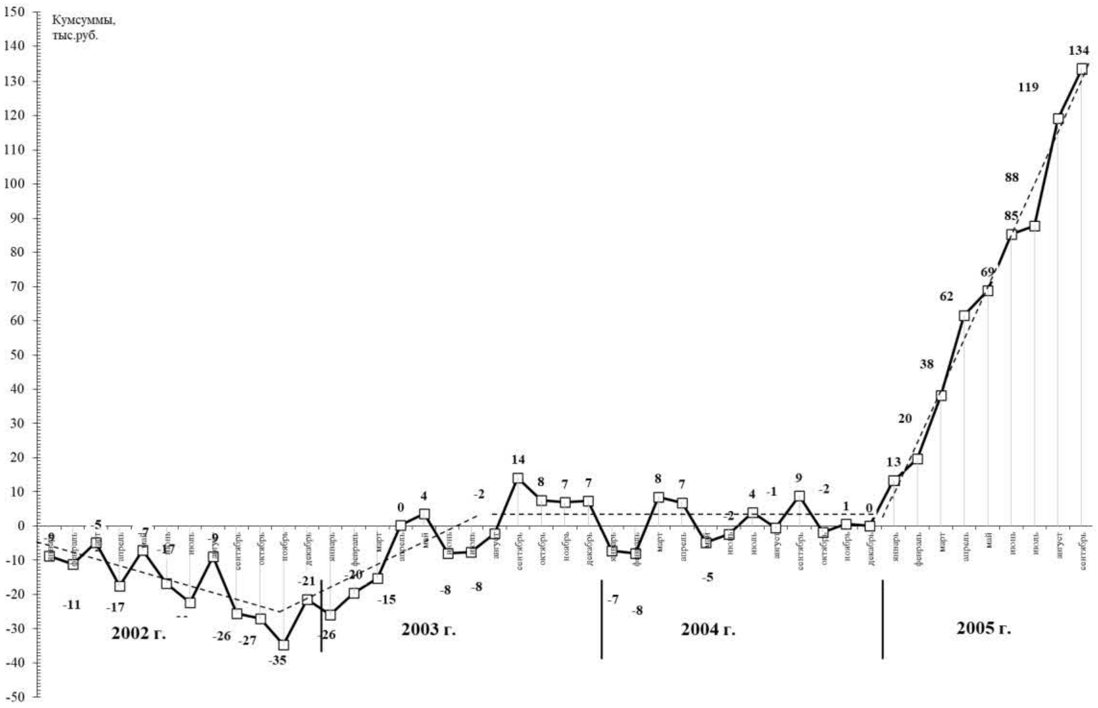


Рис. 2. Карта куммулятивных сумм

Таблица 3

Значения кумулятивных сумм

Год	янв	фев	мар	апр	май	июн	июл	авг	сен	окт	ноя	дек
X	– 9	– 11	– 5	– 17	– 7	– 17	– 22	– 9	– 26	– 27	– 35	– 21
X+1	– 26	– 20	– 15	0	4	– 8	– 8	– 2	14	8	7	7
X+2	– 7	– 8	8	7	– 5	– 2	4	– 1	9	–2	1	0
X+3	13	20	38	62	69	85	88	119	134	–	–	–

Именно наклон графика служит мерой измерения случайной величины. Так, в рассматриваемом случае, нулевой наклон означает, что система работает в своем среднем режиме 137 млн руб. в мес. и т.д. согласно шкале кумулятивных сумм.

Анализ карты свидетельствует о том, что уровень ПФ меняется в течение времени контроля:

- ПФ для первых 11 месяцев X г. = 134 млн руб.;

- ПФ с декабря X г. по август X+1 г. = 139 млн руб.;

- ПФ с сентября Х+1 г. по декабрь Х+2 г. = 137 млн руб.;
- ПФ с января Х+3 г. = 147 млн руб.

Пример демонстрирует достоинства контрольных карт кумулятивных сумм как средства визуального контроля. График кумулятивных сумм более чувствителен к выявлению изменений в уровне признака (ПФ), чем карта Шухарта, и, таким образом, карты кумулятивных сумм — хорошая графическая форма представления данных для визуального контроля. Карты кумулятивных сумм позволяют обнаружить постепенное изменение среднего значения параметра. Однако при этом методе целесообразно проверять на практике значимость любого обнаруженного изменения кумулятивных сумм, особенно в случаях, когда изменение признака нельзя объяснить какими-либо конкретными фактами.

Многообразие контрольных карт. В разобранном примере рассматривался временной ряд значений контролируемого параметра. Часто используют временной ряд значений той или иной статистики, рассчитанной по выборке значений контролируемого параметра в рассматриваемый момент времени. Этой статистикой может быть выборочное среднее арифметическое значение, медиана, выборочные дисперсия, среднее квадратическое или размах, а также число дефектов или число дефектных единиц продукции, и т.д. Во всех этих случаях могут применяться контрольные карты Шухарта или карты кумулятивных сумм.

Контрольные карты Шухарта предназначены для обнаружения резких изменений контролируемого параметра. Примером является поломка резца. В то же время постепенную разладку эти карты «чувствуют» не сразу. Это объясняется тем, что статистики, отражающие состояние процесса, рассматриваются независимо друг от друга, т.е. каждый последующий результат выборочного контроля никак не учитывает предыдущую информацию.

Контрольные карты кумулятивных сумм наиболее чувствительны к постепенной разладке процесса. Примером является постепенное стачивание резца. В то же время на резкое отклонение контролируемого параметра они реагируют медленнее, чем карты Шухарта. Это объясняется тем, что для оценки состояния процесса здесь используются накопленные суммы выборочных статистик, например, кумулятивные суммы выборочных средних или кумулятивные суммы выборочных дисперсий. Таким образом, здесь учитывается не только результат контроля текущей выборки, но также используются результаты контроля предыдущих выборок.

Средняя длина серий. Чувствительность контрольной карты к разладке определяется средней длиной серии (СДС) выборок проходящего процесса. СДС определяется как среднее число выборок, предшествующих наладке технологического процесса при неизменном распределении вероятностей контролируемого параметра.

При налаженном процессе сигнал о разладке является ложным, поэтому предпочтительным является максимально возможное значение СДС выборок L_0. Чем больше значение L_0, тем реже появляется сигнал о разладке при налаженном процессе. При разлаженном процессе, наоборот, предпочтительным является возможно меньшее значение СДС выборок L_1. Чем меньше значение L_1 при разлаженном процессе, тем быстрее будет обнаружена разладка процесса.

СДС определяет эффективность плана контроля с помощью контрольной карты. Наиболее эффективным планом контроля будет тот, который обеспечит, при равных исходных условиях, наибольшее значение СДС выборок налаженного процесса L_0 и наименьшее значение СДС выборок разлаженного процесса L_1. Однако оптимизация одновременно по двум критериям невозможна. Поэтому один из критериев переводят в ограничение, обычно связанный с L_0. Величина L_0 определяет положение границ регулирования.

Метод контрольных карт в России. Судьба метода контрольных карт в России драматична. В 1970-е гг. они стали внедряться на отечественных предприятиях, давая заметный экономический эффект. Появились государственные стандарты по статистическому регулированию технологических процессов, основанные на использовании контрольных карт. Однако Рабочая группа по упорядочению системы стандартов по прикладной статистике и другим статистическим методам (1985–1987) установила, что все государственные стандарты СССР по статистическому регулированию технологических процессов содержат грубейшие ошибки и их невозможно использовать ни как справочные, ни тем более как нормативные материалы [4]. Эти стандарты были отменены. К сожалению, ряд литературных источников [29], как и отмененные ГОСТы, содержит грубые ошибки.

Диалоговая программная система СТАТКОН разработана в качестве замены ГОСТам с ошибками и содержит в себе все основные результаты теории контрольных карт. (Руководитель разработки системы СТАТКОН профессор Г.Ф. Филаретов (Московский энергетический институт) возглавлял подкомиссию по статистическому регулированию технологических процессов Рабочей группы.) В частности, в диалоговую систему СТАТКОН входят следующие разновидности контролирующих алгоритмов, относящиеся к группе алгоритмов кумулятивных сумм и предназначенные для обнаружения:

- скачкообразного изменения математического ожидания гауссовского дискретного случайного процесса с независимыми наблюдениями;

- линейного возрастающего изменения математического ожидания случайного процесса указанного типа;

- скачкообразного изменения вектора математических ожиданий векторного гауссовского случайного процесса с независимыми наблюдениями;

- скачкообразного изменения дисперсии для гауссовского дискретного процесса с независимыми наблюдениями;

- скачкообразного изменения вероятности появления некоторого события в последовательности независимых случайных событий (нерандомизированный и рандомизированный варианты).

Как известно, в подавляющем большинстве технико-экономических ситуаций распределения элементов выборки отличны от нормальных [31] (см. также раздел 2.1). Поэтому выводы, сделанные на основе использования при моделировании нормального распределения, имеют лишь теоретическое значение. Однако они дают первое представление о ситуации, первоначальные ориентиры, которые могут быть уточнены при более тщательном исследовании.

На несколько иных принципах построена диалоговая система АВРОРА-РС «Анализ временных рядов и обнаружение разладки», разработанная под руководством кандидата технических наук И.В. Никифорова (академический Институт проблем управления) и доктора физико-матемматических наук А.А. Новикова (Математический институт им. В.А. Стеклова). К ее основным функциям относятся последовательное обнаружение изменения свойств независимых случайных последовательностей и зависимых случайных последовательностей типа авторегрессии с помощью ряда жестко настроенных и адаптивных алгоритмов скорейшего обнаружения, обнаружение изменения свойств многомерных сигналов, наблюдаемых с избыточностью, расчет статистических характеристик алгоритмов обнаружения, статистическое моделирование алгоритмов на специально генерируемых или вводимых пользователем данных, оптимальная настройка алгоритмов и подбор их характеристик, параметрический анализ временных рядов на основе авторегрессии.

В нашей стране разработано большое число программных продуктов, посвященных статистическим методам управления качеством продукции [5].

Итоги. Подведем итоги настоящей главы. В России активно разрабатываются теоретические, программные и практические вопросы статистических методов сертификации и управления качеством продукции. Некоторые из них кратко рассмотрены выше. Ранее разработанные нормативно-техническая и методическая документация, диалоговые компьютерные системы по статистическим методам продолжают использоваться, несмотря на социально-политические преобразования 1990-х гг. В частности, стандарты СССР и СЭВ продолжают оставаться широко известными методическими документами, хотя

СССР и СЭВ уже нет. Большое значение имеет работа по устранению ошибок в нормативно-технических и инструктивно-методических документах с целью уменьшения числа ошибок в практической работе. Важно создать такую систему управления в научно-технической сфере, чтобы никто не мог навязать стране свои ошибки в качестве стандартов, проигнорировав протесты ведущих специалистов. При этом условии внедрение современных статистических методов сертификации и управления качеством продукции могут дать нашей стране экономический эффект, измеряемый миллиардами долларов США в год.

Литература

1. Статистические методы повышения качества. / под редакцией Х. Кумэ ; перевод с японского. — Москва : Финансы и статистика, 1990. — 301 с.

2. Организационно-методические материалы по маркетингу на предприятии. — Москва : Всесоюзный центр статистических методов и информатики, 1991. — 91 с.

3. Цели и принципы стандартизации. / под редакцией Т. Сандерса. — Москва : Изд-во стандартов, 1974. — 132 с.

4. *Орлов, А.И.* Сертификация и статистические методы / А.И. Орлов // Заводская лаборатория. — 1997. — Т. 63 — № 3. — С. 55–62.

5. *Орлов, А.И.* Внедрение современных статистических методов с помощью персональных компьютеров / А.И. Орлов // Качество и надежность изделий. — 1992. — № 5 (21). — С. 51–78.

6. *Орлов, А.И.* Об оптимизации выборочного контроля качества продукции / А.И. Орлов // Стандарты и качество. — 1989. — № 3. — С. 91–94.

7. *Broody, M.* Helping Workers Work Smarter / M. Broody // Fortune. — 1987. — June 8. — P. 86–88.

8. *Гнеденко, Б.В.* Математика и контроль качества продукции / Б.В. Гнеденко. — Москва: Знание, 1978. — 64 с.

9. *Кравченко, Г.Г.* О статистическом приемочном контроле порошкообразных материалов / Г.Г. Кравченко, А.И. Орлов // Надежность и контроль качества. — 1991. — № 2. — С. 37–39.

10. *Мердок, Дж.* Контрольные карты / Дж. Мердок ; перевод с английского. — Москва : Финансы и статистика, 1986. — 132 с.

11. *Фалько, С.Г.* Контроллинг на предприятии / С.Г. Фалько, В.М. Носов. — Москва : Знание, 1995. — 80 с.

12. *Хан, Д.* Планирование и контроль: концепция контроллинга / Д. Хан ; перевод с немецкого. — Москва : Финансы и статистика, 1997. — 800 с.

13. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях / А.М. Карминский, Н.И. Оленев, А.Г. Примак, С.Г. Фалько. — Москва : Финансы и статистика, 1998. — 256 с.

14. *Лэйард, Р.* Макроэкономика : курс лекций для российских читателей / Р. Лэйард. — Москва : Джон Уайли энд Санз, 1994. — 160 с.

15. *Пиндайк, Р.* Микроэкономика / Р. Пиндайк, Д. Рубинфельд. — Москва : Экономика : Дело, 1992. — 510 с.

16. *Varian, H.R.* Intermediate Microeconomics. A Modern Approach / H.R. Varian. — New York : W.W.Norton & Company, 1993. — 623 p.

17. *Begg, D.* Economics / D. Begg, S. Fischer, R. Dornbusch. — London : McGraw-Hill Book Company, 1991. — 667 p.

18. *Brealey, R.A.* Principles of Corporate Finance / R.A. Brealey, S.C. Myers. — New York : McGraw-Hill, Inc., 1991. — 924 p.

19. *Орлов, А.И.* О современных проблемах внедрения прикладной статистики и других статистических методов. / А.И. Орлов // Заводская лаборатория. — 1992. — Т. 58. — № 1. — С. 67–74.

20. *Орлов, А.И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат / А.И. Орлов // Заводская лаборатория. — 1985. — Т. 51. — № 1. — С. 60–62.

21. *Орлов, А.И.* Создана единая статистическая ассоциация / А.И. Орлов // Вестник Академии наук СССР. — 1991. — № 7. — С. 152–153.

22. *Орлов, А.И.* Всесоюзная статистическая ассоциация — гарантия успешного внедрения современных статистических методов. / А.И. Орлов // Надежность и контроль качества. — 1991. — № 6. — С. 54–55.

23. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

24. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — Изд. 3-е, перераб. и доп. — Москва : Экзамен, 2004. — 576 с.

25. *Орлов, А.И.* Математика случая. Вероятность и статистика — основные факты : справочник для вузов / А.И. Орлов. — Москва : Экзамен, 2006. — 136 с.

26. *Панде, П.* Что такое «Шесть сигм»? Революционный метод управления качеством / П. Панде, Л. Холп ; перевод с английского. — Москва : Альпина Бизнес Букс, 2004. — 158 с.

27. *Фалько, С.Г.* «Шесть сигм» как подход к совершенствованию бизнеса / С.Г. Фалько, А.И. Орлов // Контроллинг. — 2004. — № 4 (12). — С. 42–46.

28. *Андерсен, Б.* Бизнес-процессы. Инструменты совершенствования / Б. Андерсен. — Москва : Стандарты и качество, 2005. — 272 с.

29. *Богатырев, А.А.* Стандартизация статистических методов управления качеством / А.А. Богатырев, Ю.Д. Филиппов. — Москва : Изд-во Стандартов, 1989. — 120 с.

30. *Орлов, А.И.* Непараметрическое точечное и интервальное оценивание характеристик распределения / А.И. Орлов // Заводская лаборатория. — 2004. — Т. 70. — № 5. — С. 65–70.

31. *Орлов, А.И.* Часто ли распределение результатов наблюдений является нормальным? / А.И. Орлов // Заводская лаборатория. — 1991. — Т. 57. — № 7. — С. 64–66.

32. *Орлов, А.И.* «Шесть сигм» — новая система внедрения математических методов исследования / А.И. Орлов // Заводская лаборатория. — 2006. — Т. 72. — № 5. — С. 50–53.

33. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в цифровой экономике (на примере управления качеством) / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 169. — С. 216–242.

34. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в организации производства в эпоху цифровой экономики / А.И. Орлов // Инновации в менеджменте. — 2021. — № 2 (28). — С. 36–45.

35. *Фалько, С.Г.* «Шесть сигм» как подход к совершенствованию бизнеса / С.Г. Фалько, А.И. Орлов // Контроллинг. — 2004. — № 4 (12). — С. 42–46.

36. *Орлов, А.И.* «Шесть сигм» — новая система внедрения математических методов исследования // Заводская лаборатория. Диагностика материалов. — 2006. — Т. 72. — № 5. — С. 50–53.

37. *Орлов, А.И.* Предельные теоремы в статистическом контроле / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 116. — С. 462–483.

38. *Орлов, А.И.* Метод проверки гипотез по совокупности малых выборок и его применение в теории статистического контроля / А.И. Орлов // Научный журнал КубГАУ. —2014. — № 104. — С. 38–52.

39. *Орлов, А.И.* Всегда ли нужен контроль качества продукции у поставщика? / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 96. — С. 709–724.

40. *Орлов, А.И.* Основные проблемы контроллинга качества // Научный журнал КубГАУ. — 2015. — № 111. — С. 20–52.

41. *Орлов, А.И.* Выявление отклонений в контроллинге (на примере мониторинга уровня безопасности полетов) / А.И. Орлов, В.Д. Шаров // Научный журнал КубГАУ. — 2014. — № 95. — С. 460–469.

42. *Филаретов, Г.Ф.* Последовательный алгоритм обнаружения момента изменения дисперсии временного ряда / Г.Ф. Филаретов, А.А. Червова // Заводская лаборатория. Диагностика материалов. — 2019. — Т. 85. — № 3. — С. 75–82.

Контрольные вопросы и задачи

1. Какие решения необходимо принимать в связи с качеством продукции и сертификацией?

2. Почему необходимо использование выборочного контроля?

3. Для плана (n, 0) с n = 27 найти приемочный уровень дефектности и браковочный уровень дефектности.

4. Для плана (n, 0) предел среднего выходного уровня дефектности не превышает t = 0,02. Каково минимально возможное n?

5. Даны приемочный уровень дефектности $p_{пр}$ = 0,03 и браковочный уровень дефектности $p_{бр}$ = 0,09. Указать какой-либо допустимый план вида (n, с), т.е. план, значение оперативной характеристики которого в точке $p_{пр}$ не меньше 0,95, а в точке $p_{бр}$ не больше 0,10.

6. Как могли быть утверждены государственные стандарты по статистическим методам управления качеством продукции, содержащие грубейшие ошибки?

7. Чем контрольные карты Шухарта отличаются от карт кумулятивных сумм?

Темы докладов, рефератов, исследовательских работ

1. Комплексные системы управления качеством продукции и международные стандарты ИСО по менеджменту качества.

2. Экономическая эффективность усеченных планов статистического приемочного контроля.

3. Взаимосвязь технического уровня, качества и конкурентоспособности продукции.

4. Методы проведения статистического приемочного контроля порошкообразных материалов.

5. Различные варианты взаимодействия поставщика и потребителя в связи с системой принятием решений о качестве продукции.

6. Расчет средней длины серий для контрольных карт Шухарта и карт кумулятивных сумм при скачкообразной и постепенной разладке.

ГЛАВА 11. СТАТИСТИЧЕСКИЕ МОДЕЛИ ЭКСПЕРТНЫХ ИССЛЕДОВАНИЙ

Глава посвящена основным вопросам теории и практики экспертных оценок, в том числе связанным с типовыми стадиями экспертного опроса, методами подбора экспертов, разработкой регламентов проведения сбора и анализа экспертных мнений. Рассмотрены основные идеи современной теории измерений, метода согласования кластеризованных ранжировок и ряда других математических методов анализа экспертных оценок.

11.1. Примеры процедур экспертных оценок

Как изменится экологическая обстановка через десять лет? Будет ли обеспечена экологическая безопасность промышленных производств или же вокруг будет простираться рукотворная пустыня? Достаточно вдуматься в эту постановку вопроса, проанализировать, как десять лет назад мы представляли себе сегодняшний день, чтобы понять, что стопроцентно надежных прогнозов просто не может быть. Вместо утверждений с конкретными числами можно ожидать лишь качественных оценок. Тем не менее, мы должны принимать решения, например, об экологических и иных проектах и инвестициях, последствия которых скажутся через десять, двадцать и более лет.

Бесспорно, что для принятия обоснованных решений необходимо опираться на опыт, знания и интуицию специалистов. После второй мировой войны в рамках кибернетики, теории управления, менеджмента и исследования операций стала развиваться самостоятельная дисциплина — теория и практика экспертных оценок.

Методы экспертных оценок — это методы организации работы со специалистами-экспертами и обработки мнений экспертов. Эти мнения обычно выражены частично в количественной, частично в качественной форме. Экспертные исследования я проводят с целью подготовки информации для принятия решений лицом, принимающим решения (ЛПР). Для проведения работы по методу экспертных оценок создают Рабочую группу (сокращенно РГ), которая и организует по поручению ЛПР деятельность экспертов, объединенных (формально или по существу) в экспертную комиссию (ЭК).

Экспертные оценки бывают *индивидуальные* и *коллективные*. *Индивидуальные оценки* — это оценки одного специалиста. Например, преподаватель единолично ставит отметку студенту, а врач — диагноз больному. Но в сложных случаях заболевания или при угрозе отчисления студента за плохую учебу обращаются к *коллективному* мнению — симпозиуму врачей или комиссии преподавателей. Аналогичная ситуация — в армии. Обычно командующий принимает решение единолично. Но в сложных и ответственных ситуациях проводят военный совет. Один из наиболее известных примеров такого рода — военный совет 1812 г. в Филях, на котором под председательством М.И. Кутузова решался вопрос: «Давать или не давать французам сражение под Москвой?»

Другой простейший пример экспертных оценок — оценка номеров в КВН. Каждый из членов жюри поднимают фанерку со своей оценкой, а технический работник вычисляет среднюю арифметическую оценку, которая и объявляется как коллективное мнение жюри (ниже мы увидим, что такой подход некорректен с точки зрения теории измерений).

В фигурном катании процедура усложняется — перед усреднением *отбрасываются самая большая и самая маленькая оценки.* Это делается для того, чтобы не было соблазна завысить оценку одной спортсменке (например, соотечественнице) или занизить другой. Такие резко выделяющиеся из общего ряда оценки будут сразу отброшены.

Экспертные оценки часто используются при выборе — одного варианта технических устройств из нескольких, группы космонавтов из многих претендентов, набора проектов научно-исследовательских работ для финансирования из массы заявок, получателей экологических кредитов из многих желающих, выбор инвестиционных проектов для реализации среди представленных, и т.д.

Существует масса методов получения экспертных оценок. В одних с каждым экспертом работают отдельно, он даже не знает, кто еще является экспертом, а потому высказывает свое мнение независимо от авторитетов. В других экспертов собирают вместе для подготовки материалов для ЛПР, при этом эксперты обсуждают проблему друг с другом, учатся друг у друга, и неверные мнения отбрасываются. В одних методах число экспертов фиксировано и таково, чтобы статистические методы проверки согласованности мнений и затем их усреднения позволяли принимать обоснованные решения. В других — число экспертов растет в процессе проведения экспертизы, например, при использовании метода «снежного кома» (о нем — ниже). Не меньше существует и методов обработки ответов экспертов, в том числе весьма насыщенных математикой и компьютеризированных

Один из наиболее известных методов экспертных оценок — это *метод «Дельфи».* Название дано по ассоциации с Дельфийским храмом, куда, согласно древнему обычаю, было принято обращаться для получения поддержки при принятии решений. Он был расположен у выхода ядовитых вулканических газов. Жрицы храма, надышавшись отравы, начинали пророчествовать, произнося непонятные слова. Специальные «переводчики» — жрецы храма толковали эти слова и отмечали на вопросы пришедших со своими проблемами паломников.

В США в 1960-х гг. методом Дельфи назвали экспертную процедуру прогнозирования научно-технического развития. В первом туре эксперты называли вероятные даты тех или иных будущих свершений. Во втором туре каждый эксперт знакомился с прогнозами всех остальных. Если его прогноз сильно отличался от прогнозов основной массы, его просили пояснить свою позицию, и часто он изменял свои оценки, приближаясь к средним значениям. Эти средние значения и выдавались заказчику как групповое мнение. Надо сказать, что реальные результаты исследования оказались довольно скромными — хотя дата высадки американцев на Луну была предсказана с точностью до месяца, все остальные прогнозы провалились — холодного термоядерного синтеза и средства от рака в XX в. человечество не дождалось. Однако сама методика оказалась популярной — за последующие годы она использовалась не менее 40 тыс. раз. Средняя стоимость экспертного исследования по методу Дельфи — 5 тыс. долларов США, но в ряде случаев приходилось расходовать и более крупные суммы — до 130 тыс. долларов.

Несколько в стороне от основного русла экспертных оценок лежит *метод сценариев*, применяемый прежде всего для экспертного прогнозирования. Рассмотрим основные идеи технологии сценарных экспертных прогнозов. Экологическое или социально-экономическое прогнозирование, как и любое прогнозирование вообще, может быть успешным лишь при некоторой стабильности условий. Однако решения органов власти, отдельных лиц, иные события меняют условия, и события развиваются по-иному, чем ранее предполагалось. При разработке методологического, программного и информационного обеспечения *анализа риска* химико-технологических проектов необходимо составить детальный каталог сценариев аварий, связанных с утечками токсических химических веществ. Каждый из таких сценариев описывает аварию своего типа, со своим индивидуальным происхождением, развитием, последствиями, возможностями предупреждения. Таким образом, метод сценариев — это метод декомпозиции задачи прогнозирования, предусматривающий выделение набора отдельных вариантов развития событий (сценариев), в совокупности охватывающих все возможные варианты развития. При этом каждый отдельный сценарий должен допускать возможность достаточно точного прогнозирования, а общее число сценариев должно быть обозримо.

Возможность подобной декомпозиции не очевидна. При применении метода сценариев необходимо осуществить два этапа исследования:

- построение исчерпывающего, но обозримого набора сценариев;

- прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы.

Каждый из этих этапов лишь частично формализуем. Существенная часть рассуждений проводится на качественном уровне, как это принято в общественно-экономических и гуманитарных науках. Одна из причин заключается в том, что стремление к излишней формализации и математизации приводит к *искусственному* внесению определенности там, где ее нет по существу, либо к использованию громоздкого математического аппарата. Так, рассуждения на словесном уровне считаются доказательными в большинстве ситуаций, в то время как попытка уточнить смысл используемых слов с помощью, например, теории нечетких множеств приводит к весьма громоздким математическим моделям. Набор сценариев должен быть обозрим. Приходится исключать различные маловероятные события. Само по себе создание набора сценариев — предмет экспертного исследования. Кроме того, эксперты могут оценить вероятности реализации того или иного сценария. Прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы также осуществляется в соответствии с описанной выше методологией прогнозирования. При стабильных условиях могут быть применены статистические методы прогнозирования временных рядов. Однако этому предшествует анализ с помощью экспертов, причем зачастую прогнозирование на словесном уровне является достаточным (для получения интересующих исследователя и ЛПР выводов) и не требующим количественного уточнения.

Как известно, при принятии решений на основе *анализа ситуации*, в том числе результатов прогнозных исследований, можно исходить из различных критериев. Так, можно ориентироваться на то, что ситуация сложится наихудшим, или наилучшим, или средним (в каком-либо смысле) образом. Можно попытаться наметить мероприятия, обеспечивающие минимально допустимые полезные результаты при любом варианте развития ситуации, и т.д.

Еще один вариант экспертного оценивания — *мозговой штурм*. Организуется он как собрание экспертов, на выступления которых наложено одно, но очень существенное ограничение — нельзя критиковать предложения других. Можно их развивать, можно высказывать свои идеи, но нельзя критиковать! В ходе заседания эксперты, «заражаясь» друг от друга, высказывают все более экстравагантные соображения. Часа через два записанное на магнитофон или видеокамеру заседание заканчивается, и начинается второй этап мозгового штурма — анализ высказанных идей. Обычно из 100 идей 30 заслуживают дальнейшей проработки, из 5–6 дают возможность сформулировать прикладные проекта, а 2–3 оказываются в итоге приносящими полезный эффект — прибыль, повышение экологической безопасности и т.п. При этом интерпретация идей — творческий процесс. Например, при обсуждении возможностей защиты кораблей от торпедной атаки была высказана идея: «Выстроить матросов вдоль борта и дуть на торпеду, чтобы изменить ее курс». После проработки эта идея привела к созданию устройств, создающих волны, сбивающиеся торпеду с курса.

11.2. Основные стадии экспертного опроса

Как показывает опыт проведения экспертных исследований, с точки зрения менеджера — организатора такого исследования целесообразно выделять следующие стадии проведения экспертного опроса.

1. *Принятие решения* о необходимости проведения экспертного опроса и формулировка Лицом, Принимающим Решения (ЛПР) его цели. Таким образом, инициатива должна исходить от руководства, что в дальнейшем обеспечит успешное решение организационных и финансовых проблем.

2. *Подбор и назначение ЛПР основного состава Рабочей группы,* сокращенно РГ (обычно — научного руководителя и секретаря). При этом научный руководитель отвечает за организацию и проведение экспертного исследования в целом, а также за анализ собранных материалов и формулировку заключения экспертной комиссии. Он участвует в формировании коллектива экспертов и выдаче задания каждому (вместе с ЛПР или его представителем). Дело секретаря — ведение документации, решение организационных задач.

3. *Разработка* РГ (точнее, ее основным составом, прежде всего научным руководителем и секретарем) и утверждение у ЛПР *технического задания* на проведение экспертного опроса. На этой стадии решение о проведении экспертного опроса приобретает четкость во времени, финансовом, кадровом, материальном и организационном обеспечении. В частности, в РГ выделяются различные группы специалистов — аналитическая, эконометрическая (специалисты по методам анализа данных), компьютерная, по работе с экспертами (например, интервьюеров), организационная. Очень важно для успеха, чтобы все эти направления работ были утверждены ЛПР.

4. *Разработка* аналитической группой РГ подробного сценария (т.е. *регламента*) проведения сбора и анализа экспертных мнений (оценок). Сценарий включает в себя прежде всего конкретный вид информации, которая будет получена от экспертов (например, тексты (слова), условные градации, числа, ранжировки, разбиения или иные виды объектов нечисловой природы). Так, довольно часто экспертов просят высказаться в свободной форме, ответив при этом на некоторые количество заранее сформулированных вопросов. Кроме того, их просят заполнить формальную карту, в каждом пункте выбрав одну из нескольких градаций. Сценарий должен содержать и конкретные методы анализа собранной информации. Например, вычисление медианы Кемени, статистический анализ люсианов, применение иных методов статистики объектов нечисловой природы и других разделов современной эконометрики. Эта работа ложится на эконометрическую и компьютерную группу РГ. Традиционная ошибка — сначала собрать информацию, а потом думать, что с ней делать. В результате, как показывает опыт, информация используется лишь на 1–2 %.

5. *Подбор экспертов* в соответствии с их компетентностью. На этой стадии РГ составляет список возможных экспертов.

6. *Формирование экспертной комиссии*. На этой стадии РГ проводит переговоры с экспертами, получает их согласие на работу в экспертной комиссии (сокращенно ЭК), возможно, часть намеченных РГ экспертов отказывается по тем или иным причинам. ЛПР утверждает состав экспертной комиссии, возможно, вычеркнув или добавив часть экспертов к предложениям РГ. Проводится заключение договоров с экспертами об условиях их работы и ее оплаты.

7. Проведение *сбора* экспертной информации. Часто перед этим проводится набор и обучение интервьюеров — одной из групп, входящих в РГ.

8. Компьютерный *анализ экспертной информации* с помощью включенных в сценарий методов. Ему обычно предшествует введение информации в компьютеры.

9. При применении согласно сценарию экспертной процедуры из нескольких туров — *повторение* двух предыдущих этапов.

10. Итоговый анализ экспертных мнений, интерпретация полученных результатов аналитической группой РГ и *подготовка заключительного документа* ЭК для ЛПР.

11. Официальное *окончание* деятельности РГ, в том числе утверждение ЛПР заключительного документа ЭК, подготовка и утверждение научного и финансового отчетов РГ о проведении исследования, оплата труда экспертов и сотрудников РГ, официальное прекращение деятельности (роспуск) ЭК и РГ.

Разберем подробнее отдельные стадии экспертного исследования.

Подбор экспертов. Проблема подбора экспертов является одной из наиболее сложных в теории и практике экспертных исследований. Очевидно, в качестве экспертов необходимо использовать тех людей, чьи суждения наиболее помогут принятию адекватного решения. Но как выделить, найти, подобрать таких людей? Надо прямо сказать, что нет методов подбора экспертов, наверняка обеспечивающих успех экспертизы.

В проблеме подбора экспертов можно выделить две составляющие — *составление списка возможных экспертов и выбор из них экспертной комиссии в соответствии с компетентностью кандидатов.*

Составление списка возможных экспертов облегчается тогда, когда рассматриваемый вид экспертизы проводится многократно. В таких ситуациях обычно ведется *реестр* возможных экспертов, например, в области государственной экологической экспертизы или судейства фигурного катания, из которого можно выбирать по различным критериям или с помощью датчика псевдослучайных чисел.

Как быть, если экспертиза проводится впервые, устоявшиеся списки возможных экспертов отсутствуют? Однако и в этом случае у каждого конкретного специалиста есть некоторое представление о том, что требуется от эксперта в подобной ситуации. Для формирования списка есть полезный метод «снежного кома», при котором от каждого специалиста, привлекаемого в качестве эксперта, получают несколько фамилий тех, кто может быть экспертом по рассматриваемой тематике. Очевидно, некоторые из этих фамилий встречались ранее в деятельности РГ, а некоторые — новые. Каждого вновь появившегося опрашивают по той же схеме. Процесс расширения списка останавливается, когда новые фамилии практически перестают встречаться. В результате получается достаточно обширный список возможных экспертов. Метод «снежного кома» имеет и недостатки. Число туров до остановки процесса наращивания кома нельзя заранее предсказать. Кроме того, ясно, что если на первом этапе все эксперты были из одного «клана», придерживались в чем-то близких взглядов или занимались сходной деятельностью, то и метод «снежного кома» даст, скорее всего, прежде всего лиц из этого «клана». Мнения и аргументы других «кланов» будут упущены.

Вопрос об оценке компетентности экспертов не менее сложен. Успешность участия в предыдущих экспертизах — хороший критерий для деятельности дегустатора, врача, судьи в спортивных соревнованиях, т.е. таких экспертов, которые участвуют в длинных сериях однотипных экспертиз. Однако, увы, наиболее интересны и важны уникальные экспертизы больших проектов, не имеющих аналогов. Использование формальных показателей экспертов (должность, ученые степень и звание, стаж, число публикаций...), очевидно, в современных условиях может носить лишь вспомогательный характер, хотя подобные показатели проще всего применять.

Часто предлагают использовать методы самооценки и взаимооценки компетентности экспертов. Обсудим их, начав с метода самооценки, при котором эксперт сам дает информацию о том, в каких областях он компетентен, а в каких — нет. С одной стороны, кто лучше может знать возможности эксперта, чем он сам? С другой стороны, при самооценке компетентности скорее оценивается степень самоуверенности эксперта, чем его реальная компетентность. Тем более, что само понятие *«компетентность»* строго не определено. Можно его уточнять, выделяя составляющие, но при этом усложняется предварительная часть деятельности экспертной комиссии. Достаточно часто эксперт преувеличивает свою реальную компетентность. Например, большинство людей считают, что они хорошо разбираются в политике, экономике, проблемах образования и воспитания, семьи и медицины. На самом деле экспертов (и даже знающих людей) в этих областях весьма мало. Бывают уклонения и в другую сторону, излишне критичное отношение к своим возможностям.

При использовании метода взаимооценки, помимо возможности проявления личностных и групповых симпатий и антипатий, играет роль малая осведомленность экспертов о возможностях друг друга. В современных условиях достаточно хорошее знакомство с работами и возможностями друг друга может быть лишь у специалистов, много лет (не менее 3–4) работающих совместно, в одной комнате, над одной темой. Именно про такие пары можно сказать, что они «*вместе пуд соли съели*». Однако привлечение таких пар специалистов не очень-то целесообразно, поскольку их взгляды из-за схожести жизненного пути слишком похожи друг на друга.

Если процедура экспертного опроса предполагает непосредственное общение экспертов, необходимо учитывать еще ряд обстоятельств. Большое значение имеют их личностные (социально-психологические) качества. Так, один-единственный «*говорун*» может парализовать деятельность всей комиссии на совместном заседании. К срыву могут привести и неприязненные отношения членов комиссии, и сильно различающийся научный и должностной статус членов комиссии. В подобных случаях важно соблюдение регламента работы, разработанного РГ.

Необходимо подчеркнуть, что подбор экспертов в конечном счете — функция Рабочей группы, и никакие методики подбора не снимают с нее ответственности. Другими словами, именно на Рабочей группе лежит ответственность за компетентность экспертов, за их принципиальную способность решить поставленную задачу. Важным является требование к ЛПР об утверждении списка экспертов. При этом ЛПР может как добавить в комиссию отдельных экспертов, так и вычеркнуть некоторых из них — по собственным соображениям, с которыми членам РГ и ЭК знакомиться нет необходимости.

Существует ряд нормативных документов, регулирующих деятельность экспертных комиссий в тех или иных областях. Примером является Федеральный закон «Об экологической экспертизе» от 23.11.1995 № 174-ФЗ, в котором регламентируется процедура экспертизы «намечаемой хозяйственной или иной деятельности» с целью выявления возможного вреда, который может нанести рассматриваемая деятельность окружающей природной среде.

О разработке регламента проведения сбора и анализа экспертных мнений. Существует масса методов получения экспертных оценок. В одних с каждым экспертом работают отдельно, он даже не знает, кто еще является экспертом, а потому высказывает свое мнение независимо от авторитетов, «кланов» и отдельных коллег. В других экспертов собирают вместе для подготовки материалов для ЛПР, при этом эксперты обсуждают проблему друг с другом, принимают или отвергают аргументы друг друга, учатся друг у друга, и неверные или недостаточно обоснованные мнения отбрасываются. В одних методах число экспертов фиксировано и таково, чтобы статистические методы проверки согласованности мнений и затем (в случае достаточно хорошей согласованности мнений) их усреднения позволяли принимать обоснованные решения с точки зрения эконометрики. В других — число экспертов растет в процессе проведения экспертизы, например, при использовании метода «снежного кома» для формирования команды экспертов.

В настоящее время *не существует* общепринятой научно обоснованной классификации методов экспертных оценок и тем более — однозначных рекомендаций по их применению. Попытка силой утвердить одну из возможных точек зрения может принести лишь вред.

Однако для рассказа о многообразии экспертных оценок необходима какая-либо рабочая классификация методов. Одну из таких возможных классификаций мы даем ниже, перечисляя основания, по которым мы делим экспертные оценки.

Один из основных вопросов — что именно должна представить экспертная комиссия в результате своей работы — информацию для принятия решения ЛПР или проект самого решения? От ответа на этот методологический вопрос зависит организация работы экспертной комиссии, и он служит первым основанием для разбиения методов.

ЦЕЛЬ — СБОР ИНФОРМАЦИИ ДЛЯ ЛПР. Тогда Рабочая группа должна собрать возможно больше относящейся к делу информации, аргументов «за» и «против» определенных вариантов решений. Полезен следующий метод постепенного увеличения числа экспертов. Сначала первый эксперт приводит свои соображения по рассматриваемому вопросу. Составленный им материал передается второму эксперту, который добавляет свои аргументы. Накопленный материал поступает к следующему — третьему — эксперту... Процедура заканчивается, когда иссякает поток новых соображений.

Отметим, что эксперты в рассматриваемом методе только поставляют информацию, аргументы «за» и «против», но не вырабатывают согласованного проекта решения. Нет никакой необходимости стремиться к тому, чтобы экспертные мнения были согласованы между собой. Более того, наибольшую пользу приносят эксперты с мышлением, отклоняющимся от массового. Именно от них следует ожидать наиболее оригинальных аргументов.

ЦЕЛЬ — ПОДГОТОВКА ПРОЕКТА РЕШЕНИЯ ДЛЯ ЛПР. Математические методы в экспертных оценках применяются обычно именно для решения задач, связанных с подготовкой проекта решения. При этом зачастую некритически принимают догмы согласованности и одномерности. Эти догмы «кочуют» из одной публикации в другую, поэтому целесообразно их обсудить.

ДОГМА СОГЛАСОВАННОСТИ. Часто без всяких оснований считается, что решение может быть принято лишь на основе согласованных мнений экспертов. Поэтому исключают из экспертной группы тех, чье мнение отличается от мнения большинства. При этом отсеиваются как неквалифицированные лица, попавшие в состав экспертной комиссии по недоразумению или по соображениям, не имеющим отношения к их профессиональному уровню, так и наиболее оригинальные мыслители, глубже проникшие в проблему, чем большинство. Следовало бы выяснить их аргументы, предоставить им возможность для обоснования их точек зрения. Вместо этого их мнением пренебрегают.

Бывает и так, что эксперты делятся на две или более групп, имеющих единые *групповые* точки зрения. Так, известен пример деления специалистов при оценке результатов научно-исследовательских работ на две группы: «теоретиков», явно предпочитающих НИР, в которых получены теоретические результаты, и «практиков», выбирающих те НИР, которые позволяют получать непосредственные прикладные результаты (речь идет о конкурсе НИР в академическом Институте проблем управления (автоматики и телемеханики)). Иногда заявляют, что в случае обнаружения двух или нескольких групп экспертов опрос не достиг цели. Это не так! *Цель достигнута — установлено, что единого мнения нет.* Это весьма важно. И ЛПР при принятии решений должен это учитывать. Стремление обеспечить согласованность мнений экспертов любой целой может приводить к сознательному одностороннему подбору экспертов, игнорированию всех точек зрения, кроме одной, наиболее «полюбившейся» Рабочей группе (или даже «подсказанной» ЛПР).

Часто не учитывают еще одного чисто эконометрического обстоятельства. Поскольку число экспертов обычно не превышает 20–30, то формальная статистическая согласованность мнений экспертов (установленная с помощью тех или иных критериев проверки статистических гипотез) может сочетаться с реально имеющимся разделением экспертов на группы, что делает дальнейшие расчеты не имеющими отношения к действительности. Для примера обратимся к конкретным методам расчетов с помощью коэффициентов конкордации на основе коэффициентов ранговой корреляции Кендалла или Спирмена. Необходимо напомнить, что, согласно эконометрической теории, положительный результат проверки согласованности таким способом означает ни больше, ни меньше, как отклонение гипотезы о независимости и равномерной распределенности мнений экспертов на множестве всех ранжировок. Таким образом, проверяется нулевая гипотеза, согласно которой ранжировки, описывающие мнения экспертов, являются независимыми случайными бинарными отношениями, равномерно распределенными на множестве всех ранжировок. Отклонение этой нулевой гипотезы толкуется как согласованность ответов экспертов. Другими словами, мы падаем жертвой заблуждений, вытекающих из своеобразного толкования слов: проверка согласованности в указанном математико-статистическом смысле вовсе не является проверкой согласованности в смысле практики экспертных оценок. (Именно ущербность рассматриваемых математико-статистических методов анализа ранжировок привела группу специалистов к разработке нового математического аппарата для проверки согласованности — непараметрических методов, основанных на т.н. *люсианах* и входящих в современный раздел статистических методов — *статистику нечисловых данных*). Группы экспертов с близкими методами можно выделить эконометрическими методами кластер-анализа.

МНЕНИЯ ДИССИДЕНТОВ. С целью искусственно добиться согласованности стараются уменьшить влияние мнений экспертов-*диссидентов*, т.е. инакомыслящих по сравнению с большинством. *Жесткий* способ борьбы с диссидентами состоит в игнорировании их мнений, т.е. фактически в их исключении из состава экспертной комиссии. Отбраковка экспертов, как и отбраковка резко выделяющихся результатов наблюдений (выбросов), приводит к процедурам, имеющим плохие или неизвестные статистические свойства. Так, известна *крайняя неустойчивость* классических методов отбраковки выбросов по отношению к отклонениям от предпосылок модели. *Мягкий* способ борьбы с диссидентами состоит в применении *робастных (устойчивых) статистических процедур*. Простейший пример: если ответ эксперта — действительное число, то резко выделяющееся мнение диссидента сильно влияет на среднее арифметическое ответов экспертов и не влияет на их медиану. Поэтому разумно в качестве согласованного мнения рассматривать медиану. Однако при этом игнорируются (не достигают ЛПР) аргументы диссидентов. В любом из двух способов борьбы с диссидентами ЛПР лишается информации, идущей от диссидентов, а потому может принять необоснованное решение, которое впоследствии приведет к отрицательным последствиям. С другой стороны, представление ЛПР всего набора мнений снимает часть ответственности и труда по подготовке окончательного решения с комиссии экспертов и рабочей группы по проведению экспертного опроса и перекладывает эти ответственность и труд на плечи ЛПР.

ДОГМА ОДНОМЕРНОСТИ. Распространен довольно примитивный подход, согласно которому объект экспертизы всегда можно оценить *одним числом*. Странная идея! *Оценивать человека одним числом приходило в голову лишь на невольничьих рынках*. Вряд ли даже самые рьяные квалиметристы рассматривают книгу или картину как эквивалент числа — ее «рыночной стоимости».

Вместе с тем нельзя полностью отрицать саму идею поиска обобщенных показателей качества, технического уровня и аналогичных. Так, каждый объект можно оценивать по многим показателям качества. Например, легковой автомобиль можно оценивать по таким показателям: расход бензина на 100 км пути (в среднем); надежность (средняя стоимость ремонта за год); экологическая безопасность, оцениваемая по содержанию вредных веществ в выхлопных газах; маневренность; быстрота набора скорости 100 км/час после начала движения; максимальная достигаемая скорость; длительность сохранения в салоне положительной температуры при низкой наружной температуре (– 50 °С) и выключенном двигателе; дизайн (привлекательность и «модность» внешнего вида и отделки салона); вес, и т.д. Можно ли свести оценки по этим показателям вместе? Определяющей является конкретная ситуация, для которой выбирается автомашина. Максимально достигаемая скорость важна для гонщика, но, как нам представляется, не имеет большого практического значения для водителя рядовой частной машины, особенно в городе с суровым ограничением на максимальную скорость. Для такого водителя важнее расход бензина, маневренность и надежность. Для машин различных государственных служб, видимо, надежность важнее, чем для частника, а расход бензина — наоборот. Для районов Крайнего Севера важна теплоизоляция салона, а для южных районов — нет. Таким образом, важна конкретная (узкая) постановка задачи перед экспертами. Но такой постановки зачастую нет. А тогда «игры» по разработке обобщенного показателя качества — например, в виде линейной функции от перечисленных переменных — не могут дать объективных выводов. Альтернативой единственному обобщенному показателю является математический аппарат типа *многокритериальной оптимизации* — множества Парето и т.д.

В некоторых случаях все-таки можно глобально сравнить объекты — например, с помощью тех же экспертов получить упорядочение рассматриваемых объектов — изделий или проектов. Тогда можно ПОДОБРАТЬ коэффициенты при отдельных показателях так, чтобы *упорядочение с помощью линейной функции возможно точнее соответствовало глобальному упорядочению* (например, найти эти коэффициенты методом наименьших квадратов). Наоборот, в подобных случаях НЕ СЛЕДУЕТ оценивать указанные коэффициенты с помощью экспертов. Эта простая идея до сих пор не стала очевидной для отдельных составителей методик по проведению экспертных опросов и анализу их результатов. Они упорно стараются заставить экспертов делать то, что они выполнить *не в состоянии* — указывать веса, с которыми отдельные показатели качества должны входить в итоговый обобщенный показатель. Эксперты обычно могут сравнить объекты или проекты в целом, но не могут вычленить вклад отдельных факторов. *Раз организаторы опроса спрашивают, эксперты отвечают,* но эти ответы не несут в себе надежной информации о реальности...

ВТОРОЕ ОСНОВАНИЕ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ ПРОЦЕДУР — ЧИСЛО ТУРОВ. Экспертизы могут включать один тур, некоторое фиксированное число туров (два, три, …) или неопределенное число туров. Чем больше туров, тем более тщательным является анализ ситуации, поскольку эксперты при этом обычно много раз возвращаются к рассмотрению предмета экспертизы. Но одновременно увеличивается общее время на экспертизу и возрастает ее стоимость. Можно уменьшить расходы, вводя в экспертизу не всех экспертов сразу, а постепенно. Так, например, если цель состоит в сборе аргументов «за» и «против», то первоначальный перечень аргументов может быть составлен одним экспертом. Второй добавит к нему свои аргументы. Суммарный материал поступит к первому и третьему, которые внесут свои аргументы и контраргументы. И так далее — добавляется по одному эксперту на каждый новый тур. Наибольшие сложности вызывают процедуры с заранее неопределенным числом туров, например, «снежный ком». Часто задают максимально возможное число туров, и тогда неопределенность сводится к тому, придется ли проводить это максимальное число туров или удастся ограничиться меньшим числом.

ТРЕТЬЕ ОСНОВАНИЕ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ ПРОЦЕДУР — ОРГАНИЗАЦИЯ ОБЩЕНИЯ ЭКСПЕРТОВ. Рассмотрим достоинства и недостатки каждого из элементов шкалы: отсутствие общения — заочное анонимное общение — заочное общение без анонимности — очное общение с ограничениями — очное общение без ограничений. *При отсутствии общения* эксперт высказывает свое мнение, ничего не зная о других экспертах и об их мнениях. Он полностью независим, что и хорошо, и плохо. Обычно такая ситуация соответствует однотуровой экспертизе. *Заочное анонимное общение*, например, как в методе Дельфи, означает, что эксперт знакомится с мнениями и аргументами других экспертов, но не знает, кто именно высказал то или иное положение. Следовательно, в экспертизе должно быть предусмотрено хотя бы два тура. *Заочное общение без анонимности* соответствует, например, общению по Интернету. Все варианты заочной экспертизы хороши тем, что нет необходимости собирать экспертов вместе, следовательно, находить для этого удобное время и место. При очных экспертизах эксперты говорят, а не пишут, как при заочных, и потому успевают за то же время сказать существенно больше. *Очная экспертиза с ограничениями* весьма распространена. Это — собрание, идущее по фиксированному регламенту. Примером является военный совет в императорской русской армии, когда эксперты (офицеры и генералы) высказывались в порядке от младшего (по чину и должности) к старшему. Наконец, *очная экспертиза без ограничений* — это свободная дискуссия. Все очные экспертизы имеют недостатки, связанные с возможностями отрицательного влияния на их проведение социально-психологических свойств и клановых (партийных) пристрастий участников, а также неравенства их профессионального, должностного, научного статусов. Представьте себе, что соберутся вместе 5 лейтенантов и 3 генерала. Независимо от того, какая информация имеется у того или иного участника встречи, ход ее предсказать нетрудно: генералы будут говорить, а лейтенанты — помалкивать.

КОМБИНАЦИЯ РАЗЛИЧНЫХ ВИДОВ ЭКСПЕРТИЗЫ. Реальные экспертизы часто представляют собой комбинации различных описанных выше типов экспертиз. В качестве примера рассмотрим защиту студентом дипломного проекта. Сначала идет многотуровая очная экспертиза, проводимая научным руководителем и консультантами, в результате студент подготавливает проект к защите. Затем два эксперта работают заочно — это автор отзыва сторонней организации и заведующий кафедрой, допускающий работу к защите. Обратите внимание на различие задач этих экспертов и объемов выполняемой ими работы — один пишет подробный отзыв, второй росписью на титульном листе проекта разрешает его защиту. Наконец — очная экспертиза без ограничений (для членов государственной аттестационной комиссии). Дипломный проект оценивается коллегиально, по большинству голосов, при этом один из экспертов (научный руководитель) знает работу подробно, а остальные — в основном лишь по докладу. Таким образом, имеем сочетание многотуровой и однотуровой, заочных и очных экспертиз. Подобные сочетания характерны для многих реально проводящихся экспертиз.

11.3. Теория измерений и средние величины

Для дальнейшего более углубленного рассмотрения проблем экспертных оценок понадобятся некоторые понятия так называемой *репрезентативной теории измерений*, служащей основой теории экспертных оценок, прежде всего той ее части, которая связана с анализом заключений экспертов, выраженных в качественном (а не в количественном) виде.

Мнения экспертов часто выражены в *порядковой шкале* (подробнее о шкалах говорится ниже), т.е. эксперт может сказать (и обосновать), что один показатель качества продукции более важен, чем другой, первый технологический объект более опасен, чем второй, и т.д. Но он не в состоянии сказать, *во сколько раз* или *на сколько* более важен, соответственно, более опасен. Экспертов часто просят дать ранжировку (упорядочение) объектов экспертизы, т.е. расположить их в порядке возрастания (или убывания) интенсивности интересующей организаторов экспертизы характеристики. Ранг — это номер (объекта экспертизы) в упорядоченном ряду. Формально ранги выражаются числами 1, 2, 3, ..., но с этими числами нельзя делать привычные арифметические операции. Например, хотя 1 + 2 = 3, но нельзя утверждать, что для объекта, стоящего на третьем месте в упорядочении, интенсивность изучаемой характеристики равна сумме интенсивностей объектов с рангами 1 и 2. Так, один из видов экспертного оценивания — оценки учащихся. Вряд ли кто-либо будет утверждать, что знания отличника равны сумме знаний двоечника и троечника (хотя 5 = 2 + 3), хорошист соответствует двум двоечникам (2 + 2 = 4), а между отличником и троечником такая же разница, как между хорошистом и двоечником (5 – 3 = 4 – 2). Поэтому очевидно, что для анализа подобного рода качественных данных необходима не всем известная арифметика, а другая теория, дающая базу для разработки, изучения и применения конкретных методов расчета. Это и есть РТИ. Надо иметь в виду, что в настоящее время термин «теория измерений» применяется для обозначения целого ряда научных дисциплин: классической метрологии, РТИ, некоторых других направлений, например, алгоритмической теории измерений.

Сначала РТИ развивалась как теория психофизических измерений. Основоположник РТИ американский психолог С.С. Стивенс основное внимание уделял шкалам измерения. Характерен следующий этап развития РТИ. Один из томов выпущенной в США в 1950-х гг. «Энциклопедии психологических наук» назывался «Психологические измерения». Значит, составители этого тома расширили сферу применения РТИ с психофизики на психологию в целом. А в основной статье в этом сборнике под названием, обратите внимание, «Основы теории измерений», изложение шло на абстрактно-математическом уровне, без привязки к какой-либо конкретной области применения. В этой статье упор был сделан на «гомоморфизмах эмпирических систем с отношениями в числовые» (в эти математические термины здесь вдаваться нет необходимости), и математическая сложность возросла по сравнению с работами С.С. Стивенса.

Уже в одной из первых отечественных статей по РТИ (конец 1960-х гг.) было установлено, что баллы, присваиваемые экспертами при оценке объектов экспертизы, как правило, измерены в порядковой шкале. Отечественные работы, появившиеся в начале 1970-х гг., привели к существенному расширению области использования РТИ. Ее применяли к педагогической квалиметрии (измерению качества знаний учащихся), в системных исследованиях, в различных задачах теории экспертных оценок, для агрегирования показателей качества продукции, в социологических исследованиях, и др.

В качестве двух основных проблем РТИ наряду с *установлением типа шкалы* был выдвинут поиск алгоритмов анализа данных, результат работы которых не меняется при любом допустимом преобразовании шкалы (т.е. является *инвариантным* относительно этого преобразования).

Основные шкалы измерения. В соответствии с РТИ при математическом моделировании реального явления или процесса следует прежде всего установить, *в каких типах шкал измерены* те или иные переменные. Тип шкалы задает группу допустимых преобразований. Допустимые преобразования не меняют соотношений между объектами измерения. Например, при измерении длины переход от аршин к метрам не меняет соотношений между длинами рассматриваемых объектов — если первый объект длиннее второго, то это будет установлено и при измерении в аршинах, и при измерении в метрах.

Укажем основные виды шкал измерения и соответствующие группы допустимых преобразований. В *шкале наименований* (другое название — н*оминальной* шкалы) **допустимыми** являются все взаимно-однозначные преобразования. В этой шкале числа используются лишь как метки. Примерно так же, как при сдаче белья в прачечную, т.е. лишь для различения объектов. В шкале наименований измерены, например, номера телефонов, автомашин, паспортов, студенческих билетов. Пол людей тоже измерен в шкале наименований, результат измерения принимает два значения — мужской, женский. Раса, национальность, цвет глаз, волос — номинальные признаки. Номера букв в алфавите — тоже измерения в шкале наименований. Никому в здравом уме не придет в голову складывать или умножать номера телефонов, такие операции не имеют смысла. Сравнивать буквы и говорить, например, что буква П лучше буквы С, также никто не будет. Единственное, для чего годятся измерения в шкале наименований — это различать объекты. Во многих случаях только это от них и требуется. Например, шкафчики в раздевалках для взрослых различают по номерам, т.е. числам, а в детских садах используют рисунки, поскольку дети еще не знают чисел.

В *порядковой шкале* числа используются для установления порядка между объектами. Простейшим примером являются оценки знаний учащихся. Символично, что в средней школе применяются оценки 2, 3, 4, 5, а в высшей ровно тот же смысл выражается словесно — неудовлетворительно, удовлетворительно, хорошо, отлично. Этим подчеркивается «нечисловой» характер оценок знаний учащихся. В порядковой шкале **допустимыми** являются все строго возрастающие преобразования.

Установление типа шкалы, т.е. задания группы допустимых преобразований шкалы измерения — дело специалистов соответствующей прикладной области. Так, оценки привлекательности профессий мы, выступая в качестве социологов, считали измеренными в порядковой шкале. Однако отдельные социологи не соглашались с нами, полагая, что выпускники школ пользуются шкалой с более узкой группой допустимых преобразований, например, интервальной шкалой. Очевидно, эта проблема относится не к математике, а к наукам о человеке. Для ее решения может быть поставлен достаточно трудоемкий эксперимент. Пока же он не поставлен, целесообразно принимать порядковую шкалу, так как это гарантирует от возможных ошибок.

Оценки экспертов, как уже отмечалось, часто следует считать измеренными в порядковой шкале. Типичным примером являются задачи ранжирования и классификации промышленных объектов, подлежащих экологическому страхованию (см. ниже).

Почему мнения экспертов естественно выражать именно в порядковой шкале? *Как показали многочисленные опыты, человек более правильно (и с меньшими затруднениями) отвечает на вопросы качественного, например, сравнительного, характера, чем количественного.* Так, ему легче сказать, какая из двух гирь тяжелее, чем указать их примерный вес в граммах.

Используется много других известных примеров порядковых шкал. Так, например, в минералогии используется шкала Мооса, по которому минералы классифицируются согласно критерию твердости. А именно: тальк имеет балл 1, гипс — 2, кальций — 3, флюорит — 4, апатит — 5, ортоклаз — 6, кварц — 7, топаз — 8, корунд — 9, алмаз — 10. Порядковыми шкалами в географии являются — бофортова шкала ветров («штиль», «слабый ветер», «умеренный ветер» и т.д.), шкала силы землетрясений. В медицине порядковыми шкалами являются — шкала стадий гипертонической болезни (по Мясникову), шкала степеней сердечной недостаточности (по Стражеско — Василенко — Лангу), шкала степени выраженности коронарной недостаточности (по Фогельсону). Номера домов также измерены в порядковой шкале. При оценке качества продукции и услуг, в квалиметрии (буквальный перевод: измерение качества) популярны порядковые шкалы. А именно, единица продукции оценивается как годная или не годная. При более тщательном анализе используется шкала с тремя градациями: есть значительные дефекты — присутствуют только незначительные дефекты — нет дефектов.

При оценке экологических воздействий первая оценка — обычно порядковая: природная среда стабильна — природная среда угнетена (деградирует). Аналогично в эколого-медицинской шкале: нет выраженного воздействия на здоровье людей — отмечается отрицательное воздействие на здоровье. Порядковая шкала используется и в иных областях.

Порядковая шкала и шкала наименований — основные шкалы качественных признаков. Поэтому во многих конкретных областях результаты качественного анализа можно рассматривать как измерения по этим шкалам.

Шкалы количественных признаков — это шкалы интервалов, отношений, разностей, абсолютная. По шкале *интервалов* измеряют величину потенциальной энергии или координату точки на прямой. В этих случаях на шкале нельзя отметить ни естественное начало отсчета, ни естественную единицу измерения. Допустимыми преобразованиями в шкале интервалов являются линейные возрастающие преобразования, т.е. линейные функции. Температурные шкалы Цельсия и Фаренгейта связаны именно такой зависимостью: С° = 5/9 (Ф° — 32), где С° — температура по шкале Цельсия, а Ф° — температура по шкале Фаренгейта.

Из количественных шкал наиболее распространенными в науке и практике являются шкалы *отношений*. В них есть естественное начало отсчета — нуль, т.е. отсутствие величины, но нет естественной единицы измерения. По шкале отношений измерены большинство физических единиц: масса тела, длина, заряд, а также цены в экономике. Допустимыми преобразованиями шкале отношений являются подобные (изменяющие только масштаб). Другими словами, линейные возрастающие преобразования без свободного члена.

Время измеряется по шкале *разностей*, если год принимаем естественной единицей измерения, и по шкале интервалов в общем случае. Естественного начала отсчета указать на современном уровне знаний нельзя. Дату сотворения мира различные авторы рассчитывают по-разному, равно как и момент рождества Христова. Так, согласно новой статистической хронологии, Господь Иисус Христос родился в 1054 г. в Стамбуле (он же — Царьград, Византия, Троя, Иерусалим, Рим) по принятому ныне летоисчислению.

Только для *абсолютной* шкалы результаты измерений — числа в обычном смысле слова. Примером является число людей в комнате. Для абсолютной шкалы допустимым является только тождественное преобразование.

В процессе развития соответствующей области знания тип шкалы может меняться. Так, сначала температура измерялась по *порядковой* шкале (холоднее — теплее). Затем — по *интервальной* (шкалы Цельсия, Фаренгейта, Реомюра). Наконец, после открытия абсолютного нуля температуру следует считать измеренной по шкале *отношений* (шкала Кельвина). Надо отметить, что среди специалистов иногда имеются разногласия по поводу того, по каким шкалам следует считать измеренными те или иные реальные величины. Другими словами, процесс измерения включает в себя и определение типа шкалы (вместе с обоснованием).

Инвариантные алгоритмы и средние величины. Основное требование к алгоритмам анализа данных формулируется в РТИ так: *выводы, сделанные на* *основе данных, измеренных в шкале определенного типа, не должны меняться при допустимом преобразовании шкалы измерения этих данных.* Другими словами, выводы должны быть *инвариантны* по отношению к допустимым преобразованиям шкалы.

Таким образом, одна из основных целей теории измерений — борьба с субъективизмом исследователя при приписывании численных значений реальным объектам. Так, расстояния можно измерять в аршинах, метрах, микронах, милях, парсеках и других единицах измерения. Массу (вес) — в пудах, килограммах, фунтах. Цены на товары и услуги можно указывать в юанях, рублях, тенге, гривнах, латах, кронах, марках, долларах США и других валютах (при условии заданных курсов пересчета). Подчеркнем очень важное, хотя и вполне очевидное обстоятельство: выбор единиц измерения зависит от исследователя, т.е. субъективен. *Статистические выводы могут быть адекватны реальности только тогда, когда они не зависят от того, какую единицу измерения предпочтет исследователь, т.е. когда они инвариантны относительно допустимого преобразования шкалы.*

В качестве примера рассмотрим обработку мнений экспертов, измеренных в порядковой шкале. Пусть $Y_1, Y_2, ..., Y_n$ — совокупность оценок экспертов, «выставленных» одному объекту экспертизы (например, одному из вариантов стратегического развития фирмы), $Z_1, Z_2, ..., Z_n$ — второму (другому варианту такого развития).

Как сравнивать эти совокупности? Очевидно, самый простой способ — по средним значениям. А как вычислять средние? Известны различные виды средних величин: среднее арифметическое, медиана, мода, среднее геометрическое, среднее гармоническое, среднее квадратическое. Обобщением нескольких из перечисленных является среднее по Колмогорову.

Для чисел $X_1, X_2, ..., X_n$ среднее по Колмогорову вычисляется по формуле:

$$G\{(F(X_1) + F(X_2) + ... + F(X_n))/n\},$$

где F — строго монотонная функция (т.е. строго возрастающая или строго убывающая), G — функция, обратная к F.

Среди средних по Колмогорову — много хорошо известных персонажей. Так, если $F(x) = x$, то среднее по Колмогорову — это среднее арифметическое, если $F(x) = ln\ x$, то среднее геометрическое, если $F(x) = 1/x$, то среднее гармоническое, если $F(x) = x^2$, то среднее квадратическое, и т.д. (в последних трех случаях усредняются положительные величины). С другой стороны, такие популярные средние, как медиана и мода, нельзя представить в виде средних по Колмогорову.

Напомним, что общее понятие средней величины введено французским математиком первой половины XIX в. академиком О. Коши. Оно таково: средней величиной является любая функция $f(X_1, X_2, ..., X_n)$, такая, что при всех возможных значениях аргументов значение этой функции не меньше, чем минимальное из чисел $X_1, X_2, ..., X_n$, и не больше, чем максимальное из этих чисел. Среднее по Колмогорову — частный случай среднего по Коши. Медиана и мода не являются средними по Колмогорову. Однако они — средние по Коши.

При допустимом преобразовании шкалы значение средней величины, очевидно, меняется. Но выводы о том, для какой совокупности среднее больше, а для какой — меньше, не должны меняться (в соответствии с требованием инвариантности выводов, принятом как основное требование в РТИ). Сформулируем соответствующую математическую задачу поиска вида средних величин, результат сравнения которых устойчив относительно допустимых преобразований шкалы.

Пусть $f(X_1, X_2, ..., X_n)$ — среднее по Коши. Пусть среднее по первой совокупности меньше среднего по второй совокупности:

$$f(Y_1, Y_2, ..., Y_n) < f(Z_1, Z_2, ..., Z_n).$$

Тогда согласно теории измерений для устойчивости результата сравнения средних необходимо, чтобы для любого допустимого преобразования g (из группы допустимых преобразований в соответствующей шкале) было справедливо также неравенство

$$f(g(Y_1), g(Y_2), ..., g(Y_n)) < f(g(Z_1), g(Z_2), ..., g(Z_n)),$$

т.е. среднее преобразованных значений из первой совокупности также было меньше среднего преобразованных значений для второй совокупности. Причем сформулированное условие должно быть верно для любых двух совокупностей $Y_1, Y_2, ..., Y_n$ и $Z_1, Z_2, ..., Z_n$. И, напомним, для любого допустимого преобразования. Средние величины, удовлетворяющие сформулированному условию, назовем допустимыми (в соответствующей шкале). Согласно теории измерений только допустимыми средними величинами можно пользоваться при анализе мнений экспертов и иных данных, измеренных в рассматриваемой шкале.

С помощью математической теории, развитой А.И. Орловым в 1970-х гг., удается описать вид допустимых средних в основных шкалах: в шкале наименований в качестве среднего годится только мода; из всех средних по Коши в порядковой шкале в качестве средних можно использовать только члены вариационного ряда (порядковые статистики), в частности, медиану (при нечетном объеме выборки; при четном же объеме следует применять один из двух центральных членов вариационного ряда — как их иногда называют, левую медиану или правую медиану), но не среднее арифметическое, среднее геометрическое и т.д.; в шкале интервалов из всех средних по Колмогорову можно применять только среднее арифметическое; в шкале отношений из всех средних по Колмогорову устойчивыми относительно сравнения являются только степенные средние и среднее геометрическое.

Приведем численный пример, показывающий некорректность использования среднего арифметического $f(X_1, X_2) = (X_1 + X_2)/2$ в порядковой шкале. Пусть $Y_1 = 1$, $Y_2 = 11$, $Z_1 = 6$, $Z_2 = 8$. Тогда $f(Y_1, Y_2) = 6$, что меньше, чем $f(Z_1, Z_2) = 7$. Пусть строго возрастающее преобразование g таково, что $g(1) = 1$, $g(6) = 6$, $g(8) = 8$, $g(11) = 99$. Таких преобразований много. Например, можно положить $g(x) = x$ при x, не превосходящих 8, и $g(x) = 99(x–8)/3 + 8$ для x, больших 8. Тогда $f(g(Y_1), g(Y_2)) = 50$, что больше, чем $f(g(Z_1), g(Z_2)) = 7$. Как видим, в результате допустимого, т.е. строго возрастающего преобразования шкалы упорядоченность средних величин изменилась.

Приведенные результаты о средних величинах широко применяются, причем не только в теории экспертных оценок или социологии, но и, например, для анализа методов агрегирования датчиков в АСУ ТП доменных печей. Велико прикладное значение РТИ в задачах стандартизации и управления качеством, в частности, в квалиметрии. Здесь есть и интересные теоретические результаты. Так, например, любое изменение коэффициентов весомости единичных показателей качества продукции приводит к изменению упорядочения изделий по средневзвешенному показателю (эта теорема доказана В.В. Подиновским).

11.4. Методы средних баллов

В настоящее время распространены экспертные, маркетинговые, квалиметрические, социологические и иные опросы, в которых опрашиваемых просят выставить баллы объектам, изделиям, технологическим процессам, предприятиям, проектам, заявкам на выполнение научно-исследовательских работ, идеям, проблемам, программам, политикам и т.п., а затем рассчитывают средние баллы и рассматривают их как интегральные оценки, выставленные коллективом опрошенных. Какими формулами пользоваться для вычисления средних величин? Ведь средних величин, как мы знаем, очень много разных видов. Обычно применяют среднее арифметическое. Однако специалистам уже более 30 лет известно, что *такой способ некорректен*, поскольку баллы обычно измерены в порядковой шкале (см. выше). Обоснованным является использование медиан в качестве средних баллов. Однако полностью *игнорировать средние арифметические нецелесообразно из-за их привычности и распространенности*. Поэтому **целесообразно использовать одновременно оба метода — и метод средних арифметических рангов (баллов), и методов медианных рангов.** Такая рекомендация находится в согласии с концепцией устойчивости, рекомендующей использовать различные методы для обработки одних и тех же данных с целью выделить выводы, получаемые одновременно при всех методах. Такие выводы, видимо, соответствуют реальной действительности, в то время как заключения, меняющиеся от метода к методу, зависят от субъективизма исследователя, выбирающего метод обработки исходных экспертных оценок.

Пример сравнения восьми проектов. Рассмотрим конкретный пример применения только что сформулированного подхода. По заданию руководства фирмы анализировались восемь проектов, предлагаемых для включения в план стратегического развития фирмы. Они были обозначены следующим образом: Д, Л, М-К, Б, Г-Б, Сол, Стеф, К (по фамилиям менеджеров, предложивших их для рассмотрения). Все проекты были направлены 12 экспертам, назначенным Правлением фирмы. В приведенной ниже табл. 1 приведены ранги восьми проектов, присвоенные им каждым из 12 экспертов в соответствии с представлением экспертов о целесообразности включения проекта в стратегический план фирмы. При этом эксперт присваивает ранг 1 самому лучшему проекту, который обязательно надо реализовать. Ранг 2 получает от эксперта второй по привлекательности проект, ..., наконец, ранг 8 — наиболее сомнительный проект, который реализовывать стоит лишь в последнюю очередь). Анализируя результаты работы экспертов (т.е. табл. 1), члены Правления фирмы были вынуждены констатировать, что полного согласия между экспертами нет, а потому данные, приведенные в таблице, следует подвергнуть более тщательному математическому анализу.

Таблица 1

Ранги 8 проектов по степени привлекательности для включения в план стратегического развития фирмы

№ эксперта	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
1	5	3	1	2	8	4	6	7
2	5	4	3	1	8	2	6	7
3	1	7	5	4	8	2	3	6
4	6	4	2,5	2,5	8	1	7	5
5	8	2	4	6	3	5	1	7
6	5	6	4	3	2	1	7	8
7	6	1	2	3	5	4	8	7
8	5	1	3	2	7	4	6	8
9	6	1	3	2	5	4	7	8
10	5	3	2	1	8	4	6	7
11	7	1	3	2	6	4	5	8
12	1	6	5	3	8	4	2	7

Примечание. Эксперт № 4 считает, что проекты М-К и Б равноценны, но уступают лишь одному проекту — проекту Сол. Поэтому проекты М-К и Б должны были бы стоять на втором и третьем местах и получить баллы 2 и 3. Поскольку они равноценны, то получают средний балл (2+3)/ 2 = 5/ 2 = 2,5.

Метод средних арифметических рангов. Сначала был применен метод средних арифметических рангов. Для этого прежде всего была подсчитана сумма рангов, присвоенных проектам (см. табл. 1). Затем эта сумма была разделена на число экспертов, в результате рассчитан средний арифметический ранг (именно эта операция дала название методу). По средним рангам строится итоговая ранжировка (в другой терминологии — упорядочение), исходя из принципа — чем меньше средний ранг, чем лучше проект.

Наименьший средний ранг, равный 2,625, у проекта Б, — следовательно, в итоговой ранжировке он получает ранг 1. Следующая по величине сумма, равная 3,125, у проекта М-К, — и он получает итоговый ранг 2. Проекты Л и Сол имеют одинаковые суммы (равные 3,25), значит, с точки зрения экспертов они равноценны (при рассматриваемом способе сведения вместе мнений экспертов), а потому они должны бы стоять на 3 и 4 местах и получают средний балл (3+4) /2 = 3,5. Дальнейшие результаты приведены в табл. 2 ниже. Итак, ранжировка по суммам рангов (или, что то же самое, по средним арифметическим рангам) имеет вид:

$$Б < М\text{-}К < \{Л, Сол\} < Д < Стеф < Г\text{-}Б < К. \qquad (1)$$

Здесь запись типа «А<Б» означает, что проект А предшествует проекту Б (т.е. проект А лучше проекта Б). Поскольку проекты Л и Сол получили одинаковую сумму баллов, то по рассматриваемому методу они эквивалентны, а потому объединены в группу (в фигурных скобках). В терминологии математической статистики ранжировка (1) имеет одну связь.

Таблица 2

Результаты расчетов по методу средних арифметических и методу медиан для данных, приведенных в табл. 1

Показатель	Д	Л	М-К	Б	Г-Б	Сол	Стеф	К
Сумма рангов	60	39	37,5	31.5	76	39	64	85
Среднее арифметическое рангов	5	3,25	3,125	2,625	6,333	3,25	5,333	7,083
Итоговый ранг по среднему арифметическому	5	3,5	2	1	7	3,5	6	8
Медианы рангов	5	3	3	2,25	7,5	4	6	7
Итоговый ранг по медианам	5	2,5	2,5	1	8	4	6	7

Метод медиан рангов. Значит, итог расчетов — ранжировка (1), и на ее основе предстоит принимать решение? Но тут наиболее знакомый с современной эконометрикой член Правления вспомнил, что ответы экспертов измерены в порядковой шкале, а потому для них неправомерно проводить усреднение методом средних арифметических. Надо использовать метод медиан. Что это значит? Надо взять ответы экспертов, соответствующие одному из проектов, например, проекту Д. Это ранги 5, 5, 1, 6, 8, 5, 6, 5, 6, 5, 7, 1. Затем их надо расположить в порядке неубывания (проще было бы сказать — «в порядке возрастания», но поскольку некоторые ответы совпадают, то приходится использовать непривычный термин «неубывание»). Получим последовательность: 1, 1, 5, 5, 5, 5, 5, 6, 6, 6, 7, 8. На центральных местах — шестом и седьмом — стоят 5 и 5. Следовательно, медиана равна 5.

Медианы совокупностей из 12 рангов, соответствующих определенным проектам, приведены в предпоследней строке табл. 2. (При этом медианы вычислены по обычным правилам статистики — как среднее арифметическое центральных членов вариационного ряда.) Итоговое упорядочение по методу медиан приведено в последней строке таблицы. Ранжировка (т.е. упорядочение — итоговое мнение комиссии экспертов) по медианам имеет вид:

$$Б < \{М\text{-}К, Л\} < Сол < Д < Стеф < К < Г\text{-}Б. \qquad (2)$$

Поскольку проекты Л и М-К имеют одинаковые медианы баллов, то по рассматриваемому методу ранжирования они эквивалентны, а потому объединены в группу (кластер), т.е. с точки зрения математической статистики ранжировка (2) имеет одну связь.

Сравнение ранжировок по методу средних арифметических и методу медиан. Сравнение ранжировок (1) и (2) показывает их близость (похожесть). Можно принять, что проекты М-К, Л, Сол упорядочены как М-К < Л < Сол, но из-за погрешностей экспертных оценок в одном методе признаны равноценными проекты Л и Сол (ранжировка (3)), а в другом — проекты М-К и Л (ранжировка (2)). Существенным является только расхождение, касающееся упорядочения проектов К и Г-Б: в ранжировке (1) Г-Б < К, а в ранжировке (2), наоборот, К < Г-Б. Однако эти проекты — наименее привлекательные из восьми рассматриваемых, и при выборе наиболее привлекательных проектов для дальнейшего обсуждения и использования это расхождение не существенно.

Рассмотренный пример демонстрирует сходство и различие ранжировок, полученных по методу средних арифметических рангов и по методу медиан, а также пользу от их совместного применения.

11.5. Метод согласования кластеризованных ранжировок

Проблема состоит в выделении общего нестрогого порядка из набора кластеризованных ранжировок (на статистическом языке — ранжировок со связями). Этот набор может отражать мнения нескольких экспертов или быть получен при обработке мнений экспертов различными методами. В разделе разобран *метод согласования кластеризованных ранжировок, позволяющий «загнать» противоречия внутрь специальным образом построенных кластеров (групп), в то время как упорядочение кластеров соответствует всем исходным упорядочениям.*

В различных прикладных областях возникает необходимость анализа нескольких кластеризованных ранжировок объектов. К таким областям относятся, прежде всего, экология, инженерный бизнес, менеджмент, экономика, социология, прогнозирование, научные и технические исследования и т.д., особенно те их разделы, что связаны с экспертными оценками (см., например, [1, 2]). В качестве объектов могут выступать образцы продукции, технологии, математические модели, проекты, кандидаты на должность и др. Кластеризованные ранжировки могут быть получены как с помощью экспертов, так и объективным путем, например, при сопоставлении математических моделей с экспериментальными данными с помощью того или иного критерия качества. Описанный ниже был разработан метод в связи с проблемами химической безопасности биосферы и экологического страхования.

Рассмотрим метод построения кластеризованной ранжировки, согласованной (в раскрытом ниже смысле) со всеми рассматриваемыми кластеризованными ранжировками. При этом противоречия между отдельными исходными ранжировками оказываются заключенными внутри кластеров согласованной ранжировки. В результате упорядоченность кластеров отражает общее мнение экспертов, точнее, то общее, что содержится в исходных ранжировках.

В кластеры заключены объекты, по поводу которых некоторые из исходных ранжировок *противоречат* друг другу. Для их упорядочения необходимо провести новые исследования. Эти исследования могут быть как формально-математическими (например, вычисление медианы Кемени, упорядочения по средним рангам или по медианам и т.п.), так и требовать привлечения новой информации из соответствующей прикладной области, возможно, проведения дополнительных научных или прикладных работ.

Введем необходимые понятия, затем сформулируем алгоритм согласования кластеризованных ранжировок в общем виде и рассмотрим его свойства.

Пусть имеется конечное число объектов, которые мы для простоты изложения будем изображать натуральными числами 1, 2, 3, ..., k и называть «носителем». *Под кластеризованной ранжировкой, определенной на заданном носителе, понимаем следующую математическую конструкцию*. Пусть объекты разбиты на группы, которые будем называть кластерами. В кластере может быть и один элемент. Входящие в один кластер объекты будем заключать в фигурные скобки. Например, объекты 1, 2, 3, ..., 10 могут быть разбиты на 7 кластеров: {1}, {2, 3}, {4}, {5, 6, 7}, {8}, {9}, {10}. В этом разбиении один кластер {5, 6, 7} содержит три элемента, другой — {2, 3} — два, остальные пять — по одному элементу. Кластеры не имеют общих элементов, а объединение их (как множеств) есть все рассматриваемое множество объектов.

Вторая составляющая кластеризованной ранжировки — это строгий линейный порядок между кластерами. Задано, какой из них первый, какой второй, и т.д. Будем изображать упорядоченность с помощью знака <. При этом кластеры, состоящие из одного элемента, будем для простоты изображать без фигурных скобок. Тогда кластеризованную ранжировку на основе введенных выше кластеров можно изобразить так: $A = [1 < \{2, 3\} < 4 < \{5, 6, 7\} < 8 < 9 < 10]$. Конкретные кластеризованные ранжировки будем заключать в квадратные скобки. Если для простоты речи термин «кластер» применять только к кластеру не менее чем из 2-х элементов, то можно сказать, что в кластеризованную ранжировку A входят два кластера $\{2, 3\}$ и $\{5, 6, 7\}$ и 5 отдельных элементов.

Введенная описанным образом кластеризованная ранжировка является бинарным отношением на множестве $\{1, 2, 3, ..., 10\}$. Его структура такова. Задано отношение эквивалентности с 7-ю классами эквивалентности, а именно, $\{2, 3\}$, $\{5, 6, 7\}$, а остальные состоят из оставшихся 5 отдельных элементов. Затем введен строгий линейный порядок между классами эквивалентности. Введенный математический объект известен в литературе как *«ранжировка со связями»* (М. Холлендер, Д. Вулф), *«упорядочение»* (Дж. Кемени, Дж. Снелл), *«квазисерия»* (Б.Г. Миркин), *«совершенный квазипорядок»* (Ю.А. Шрейдер [3, с. 127, 130]). Учитывая разнобой в терминологии, мы ввели термин *«кластеризованная ранжировка»,* поскольку в нем явным образом названы основные элементы изучаемого математического объекта — кластеры, рассматриваемые на этапе согласования ранжировок как классы эквивалентности, и ранжировка — строгий совершенный порядок между ними (в терминологии [3, гл. IV]).

Следующее важное понятие — *противоречивость*. Оно определяется для четверки — две кластеризованные ранжировки на одном и том же носителе и два различных объекта — элементы того же носителя. При этом два элемента из одного кластера будем связывать символом равенства = , как эквивалентные.

Пусть A и B — две кластеризованные ранжировки. *Пару объектов* (a, b) *назовем «противоречивой» относительно A и B, если эти два элемента по-разному упорядочены в A и B, т.е. $a < b$ в A и $a > b$ в B (первый вариант противоречивости) либо $a > b$ в A и $a < b$ в B (второй вариант противоречивости).* Отметим, что в соответствии с этим определением пара объектов (a, b), эквивалентная хотя бы в одной кластеризованной ранжировке, не может быть противоречивой: $a = b$ не образует «противоречия» ни с $a < b$, ни с $a > b$.

В качестве примера рассмотрим две кластеризованные ранжировки $B = [\{1, 2\} < \{3, 4, 5\} < 6 < 7 < 9 < \{8, 10\}]$, $C = [3 < \{1, 4\} < 2 < 6 < \{5, 7, 8\} <$ $< \{9, 10\}]$.

Совокупность противоречивых пар объектов для двух кластеризованных ранжировок A и B назовем «ядром противоречий» и обозначим $S(A,B)$. Для рассмотренных выше в качестве примеров трех кластеризованных ранжировок A,

B и C, определенных на одном и том же носителе {1, 2, 3,..., 10}, имеем $S(A, B)$ = [(8, 9)], $S(A, C)$ = [(1, 3), (2, 4)], $S(B, C)$ = [(1, 3), (2, 3), (2, 4), (5, 6), (8, 9)]. Как при ручном, так и при программном нахождении ядра можно в поисках противоречивых пар просматривать пары (1, 2), (1, 3), (1, 4), ..., (1, k), затем (2, 3), (2, 4), ..., (2, k), потом (3, 4), ..., (3, k) и т.д., вплоть до (k – 1, k).

Пользуясь понятиями дискретной математики, «ядро противоречий» можно изобразить *графом* с вершинами в точках носителя. При этом *противоречивые пары задают ребра этого графа.* Граф для $S(A,B)$ имеет только одно ребро (одна связная компонента более чем из одной точки), для $S(A, C)$ — 2 ребра (две связные компоненты более чем из одной точки), для $S(B, C)$ — 5 ребер (три связные компоненты более чем из одной точки {1, 2 , 3, 4}, {5, 6} и {8, 9}).

Каждую кластеризованную ранжировку, как и любое бинарное отношение, можно задать матрицей $||x(a, b)||$ из 0 и 1 порядка $k \cdot k$. При этом $x(a, b) = 1$ тогда и только тогда, когда $a < b$ либо $a = b$. В первом случае $x(b, a) = 0$, а во втором $x(b, a) = 1$. При этом хотя бы одно из чисел $x(a, b)$ и $x(b, a)$ равно 1. Из определения противоречивости пары (a, b) вытекает, что для нахождения всех таких пар достаточно поэлементно перемножить две матрицы $||x(a, b)||$ и $||y(a, b)||$, соответствующие двум кластеризованным ранжировкам, и отобрать те и только те пары, для которых $x(a, b)y(a, b) = x(b, a)y(b, a)=0$.

Предлагаемый алгоритм согласования некоторого числа кластеризованных ранжировок состоят из трех этапов. На первом *выделяются противоречивые пары* объектов во всех парах кластеризованных ранжировок. На втором формируются кластеры итоговой кластеризованной ранжировки (т.е. классы эквивалентности — *связные компоненты графов*, соответствующих объединению попарных ядер противоречий). На третьем этапе эти *кластеры (классы эквивалентности) упорядочиваются*. Для установления порядка между кластерами произвольно выбирается один объект из первого кластера и второй — из второго, порядок между кластерами устанавливается такой же, какой имеет быть между выбранными объектами в любой из рассматриваемых кластеризованных ранжировок. Корректность подобного упорядочивания, т.е. его независимость от выбора той или иной пары объектов, вытекает из соответствующих теорем, доказанных в статье [2]. Два объекта из разных кластеров согласующей кластеризованной ранжировки могут оказаться эквивалентными в одной из исходных кластеризованных ранжировок (т.е. находиться в одном кластере). В таком случае надо рассмотреть упорядоченность этих объектов в какой-либо другой из исходных кластеризованных ранжировок. Если же во всех исходных кластеризованных ранжировках два рассматриваемых объекта находились в одном кластере, то естественно считать (и это является уточнением к этапу 3 алгоритма), что они находятся в одном кластере и в согласующей кластеризованной ранжировке.

Результат согласования кластеризованных ранжировок A, B, C, ... обозначим $f(A, B, C, ...)$. Тогда $f(A, B) = [1 < 2 < 3 < 4 < 5 < 6 < 7 < \{8, 9\} < 10]$, $f(A, C) = [\{1, 3\} < \{2, 4\} < 6 < \{5, 7\} < 8 < 9 < 10]$, $f(B, C) = [\{1, 2, 3, 4\} < \{5, 6\} < 7 < \{8, 9\} < 10]$, $f(A, B, C) = f(B, C) = [\{1, 2, 3, 4\} < \{5, 6\} < 7 < \{8, 9\} < 10]$. В случае $f(A, B)$ дополнительного изучения с целью упорядочения требуют только объекты 8 и 9. В случае $f(B, C)$ объекты 1, 2, 3, 4 объединились в один кластер, т.е. кластеризованные ранжировки оказались настолько противоречивыми, что процедура согласования не позволила провести достаточно полную декомпозицию задачи нахождения итогового мнения экспертов.

Замечание. В $f(A, C)$ объекты 5 и 7 объединены в кластер потому, что и в A, и в C они равноценны. Таким образом, кластер {5, 7} имеет другие свойства, чем остальные — в нем нет элементов, образующих противоречивые пары, наоборот, его элементы равноценны.

Рассмотрим некоторые свойства алгоритмов согласования. Пусть $D = f(A, B, C, ...)$. Если $a<b$ в согласующей кластеризованной ранжировке D, то $a<b$ или $a=b$ в каждой из исходных ранжировок A, B, C, ... Построение согласующих кластеризованных ранжировок может осуществляться поэтапно. В частности, $f(A, B, C) = f(f(A, B), f(A, C), f(B, C))$. Ясно, что *ядро противоречий для набора кластеризованных ранжировок является объединением таких ядер для всех пар рассматриваемых ранжировок*. Построение согласующих кластеризованных ранжировок нацелено на выделение общего упорядочения в исходных кластеризованных ранжировках. Однако при этом некоторые общие свойства исходных кластеризованных ранжировок могут теряться. Так, при согласовании ранжировок B и C, рассмотренных выше, противоречия в упорядочении элементов 1 и 2 не было — в ранжировке B эти объекты входили в один кластер, т.е. $1 = 2$, в то время как $1 < 2$ в кластеризованной ранжировке C. Значит, при их отдельном рассмотрении можно принять упорядочение $1 < 2$. Однако в $f(B,C)$ они попали в один кластер, т.е. возможность их упорядочения исчезла. Это связано с поведением объекта 3, который «перескочил» в C на первое место и «увлек с собой в противоречие» пару (1, 2), образовав противоречивые пары и с 1, и с 2. Другими словами, связная компонента графа, соответствующего ядру противоречий, сама по себе не всегда является полным графом. Недостающие ребра при этом соответствуют парам типа (1, 2), которые сами по себе не являются противоречивыми, но «увлекаются в противоречие» другими парами.

Необходимость согласования кластеризованных ранжировок возникает, в частности, при разработке методики применения экспертных оценок в задачах экологического страхования и химической безопасности биосферы. Как уже говорилось, популярным является метод упорядочения по средним рангам, в котором итоговая ранжировка строится на основе средних арифметических рангов, выставленных отдельными экспертами [1]. Однако из теории измерений известно (см. выше, а подробнее — [4, глава 3]), что более обоснованным является использование не средних арифметических, а медиан. Вместе с тем метод средних рангов весьма известен и широко применяется, так что просто отбросить его нецелесообразно. Поэтому было принято решение об одновременном применении обеих методов. Реализация этого решения потребовала разработки методики согласования двух указанных кластеризованных ранжировок.

Рассматриваемый метод согласования кластеризованных ранжировок построен в соответствии с *методологией теории устойчивости* [4], согласно которой результат обработки данных, инвариантный относительно метода обработки, соответствует реальности, а результат расчетов, зависящий от метода обработки, отражает субъективизм исследователя, а не объективные соотношения.

11.6. Математические методы анализа экспертных оценок

При анализе мнений экспертов можно применять самые разнообразные статистические методы, описывать их — значит описывать всю прикладную статистику. Тем не менее, можно выделить основные широко используемые в настоящее время методы математической обработки экспертных оценок — это проверка согласованности мнений экспертов (или классификация экспертов, если нет согласованности) и усреднение мнений экспертов внутри согласованной группы.

Поскольку ответы экспертов во многих процедурах экспертного опроса — не числа, а такие объекты нечисловой природы, как градации качественных признаков, ранжировки, разбиения, результаты парных сравнений, нечеткие предпочтения и т.д., то для их анализа оказываются полезными методы статистики объектов нечисловой природы.

Почему ответы экспертов часто носят нечисловой характер? Наиболее общий ответ состоит в том, что люди не мыслят числами. В мышлении человека используются образы, слова, но не числа. Поэтому требовать от эксперта ответ в форме чисел — значит насиловать его разум. Даже в экономике предприниматели, принимая решения, лишь частично опираются на численные расчеты. Это видно из условного (т.е. определяемого произвольно принятыми соглашениями, обычно оформленными в виде инструкций) характера балансовой прибыли, амортизационных отчислений и других экономических показателей. Поэтому фраза типа «фирма стремится к максимизации прибыли» не может иметь строго определенного смысла. Достаточно спросить: «Максимизация прибыли — за какой период?» И сразу станет ясно, что степень оптимальности принимаемых решений зависит от горизонта планирования (на экономико-математическом уровне этот сюжет рассмотрен в монографии [4]).

Эксперт может сравнить два объекта, сказать, какой из двух лучше (метод парных сравнений), дать им оценки типа «хороший», «приемлемый», «плохой», упорядочить несколько объектов по привлекательности, но обычно не может ответить, во сколько раз или на сколько один объект лучше другого. Другими словами, ответы эксперта обычно измерены в порядковой шкале, или являются ранжировками, результатами парных сравнений и другими объектами нечисловой природы, но не числами. *Распространенное заблуждение состоит в том, что ответы экспертов стараются рассматривать как числа, занимаются «оцифровкой» их мнений, приписывая этим мнениям численные значения — баллы, которые потом обрабатывают с помощью методов прикладной статистики как результаты обычных физико-технических измерений.* В случае произвольности «оцифровки» выводы, полученные в результате обработки данных, могут не иметь отношения к реальности.

Проверка согласованности мнений экспертов и классификация экспертных мнений. Ясно, что мнения разных экспертов различаются. Важно понять, насколько велико это различие. Если мало — усреднение мнений экспертов позволит выделить то общее, что есть у всех экспертов, отбросив случайные отклонения в ту или иную сторону. Если велико — усреднение является чисто формальной процедурой. Так, если представить себе, что ответы экспертов равномерно покрывают поверхность бублика, то формальное усреднение укажет на центр дырки от бублика, а такого мнения не придерживается ни один эксперт. Из сказанного ясна важность проблемы проверки согласованности мнений экспертов.

Разработан ряд методов такой проверки. Статистические методы проверки согласованности зависят от математической природы ответов экспертов. Соответствующие статистические теории весьма трудны, если эти ответы — ранжировки или разбиения, и достаточно просты, если ответы — результаты независимых парных сравнений. Отсюда вытекает рекомендация по организации экспертного опроса: не старайтесь сразу получить от эксперта ранжировку или разбиение, ему трудно это сделать, да и имеющиеся математические методы не позволяют далеко продвинуться в анализе подобных данных. Например, рекомендуют проверять согласованность ранжировок с помощью коэффициента ранговой конкордации Кендалла — Смита. Но давайте вспомним, какая статистическая модель при этом используется. Проверяется нулевая гипотеза, согласно которой ранжировки независимы и равномерно распределены на множестве всех ранжировок. Если эта гипотеза принимается, то конечно, ни о какой согласованности мнений экспертов говорить нельзя. А если отклоняется? Тоже нельзя. Например, может быть два (или больше) центра, около которых группируются ответы экспертов. Нулевая гипотеза отклоняется. Но разве можно говорить о согласованности?

Эксперту гораздо легче на каждом шагу сравнивать только два объекта. Пусть он занимается парными сравнениями. *Непараметрическая теория парных сравнений (теория люсианов) позволяет решать более сложные задачи, чем статистика ранжировок или разбиений.* В частности, вместо гипотезы равномерного распределения можно рассматривать гипотезу однородности, т.е. вместо совпадения всех распределений с одним фиксированным (равномерным) можно проверять лишь совпадение распределений мнений экспертов между собой, что естественно трактовать как согласованность их мнений. Таким образом, удается избавиться от неестественного предположения равномерности.

При отсутствии согласованности экспертов естественно разбить их на группы сходных по мнению. Это можно сделать различными методами статистики объектов нечисловой природы, относящимися к кластер-анализу, предварительно введя метрику в пространство мнений экспертов. Идея американского математика Джона Кемени об аксиоматическом введении метрик (см. ниже) нашла многочисленных продолжателей. Однако методы кластер-анализа обычно являются эвристическими. В частности, невозможно с позиций статистической теории обосновать «законность» объединения двух кластеров в один. Имеется важное исключение — *для независимых парных сравнений (люсианов) разработаны методы, позволяющие проверять возможность объединения кластеров как статистическую гипотезу*. Это — еще один аргумент за то, чтобы рассматривать теорию люсианов как ядро математических методов экспертных оценок [1].

Нахождение итогового мнения комиссии экспертов. Пусть мнения комиссии экспертов или какой-то ее части признаны согласованными. Каково же итоговое (среднее, общее) мнение комиссии? Согласно идее Джона Кемени, следует найти среднее мнение как решение *оптимизационной задачи*. А именно, надо минимизировать суммарное расстояние от кандидата в средние до мнений экспертов. Найденное таким способом среднее мнение называют «медианой Кемени».

Математическая сложность состоит в том, что мнения экспертов лежат в некотором пространстве объектов нечисловой природы. Общая теория подобного усреднения построена в ряде работ, в частности, показано, что в силу обобщения закона больших чисел среднее мнение при увеличении числа экспертов (чьи мнения независимы и одинаково распределены) приближается к некоторому пределу, который естественно назвать *математическим ожиданием* (случайного элемента, имеющего то же распределение, что и ответы экспертов).

В конкретных пространствах нечисловых мнений экспертов вычисление медианы Кемени может быть достаточно сложным делом. Кроме свойств пространства, велика роль конкретных метрик. Так, в пространстве ранжировок при использовании метрики, связанной с коэффициентом ранговой корреляции Кендалла, необходимо проводить достаточно сложные расчеты, в то время как применение показателя различия на основе коэффициента ранговой корреляции Спирмена приводит к упорядочению по средним рангам.

Бинарные отношения и расстояние Кемени. Как известно, бинарное отношение A на конечном множестве $Q = \{q_1, q_2, ..., q_k\}$ — это подмножество *декартова квадрата* $Q^2 = \{(q_m, q_n), m,n = 1, 2, ..., k\}$. При этом пара (q_m, q_n) входит в A тогда и только тогда, когда между q_m и q_n имеется рассматриваемое отношение. Каждую кластеризованную ранжировку, как и любое бинарное отношение, можно задать матрицей $||x(a, b)||$ из 0 и 1 порядка $k \cdot k$. При этом $x(a, b) = 1$ тогда и только тогда, когда $a < b$ либо $a = b$. В первом случае $x(b, a) = 0$, а во втором $x(b, a) = 1$. При этом хотя бы одно из чисел $x(a, b)$ и $x(b, a)$ равно 1.

Как использовать связь между ранжировками и матрицами? Например, из определения противоречивости пары (a, b) (см. выше, пункт о теории измерений) вытекает, что для нахождения всех таких пар можно воспользоваться матрицами, соответствующими ранжировкам. Достаточно поэлементно перемножить две матрицы $||x(a, b)||$ и $||y(a, b)||$, соответствующие двум кластеризованным ранжировкам, и отобрать те и только те пары, для которых $x(a, b)y(a, b)=x(b, a)y(b, a)=0$.

В экспертных методах используют, в частности, такие бинарные отношения, как ранжировки (упорядочения, или разбиения на группы, между которыми имеется строгий порядок), отношения эквивалентности, толерантности (отношения сходства). Как следует из сказанного выше, каждое бинарное отношение A можно описать матрицей $||a(i, j)||$ из 0 и 1, причем $a(i, j) = 1$ тогда и только тогда, когда q_i и q_j находятся в отношении A, и $a(i, j) = 0$ в противном случае.

Определение. *Расстоянием Кемени между бинарными отношениями A и B, описываемыми матрицами* $\|a(i, j)\|$ и $\|b(i, j)\|$ *соответственно, называется число* $D (A, B) = \sum | a(i, j) — b(i, j) |$ *, где суммирование производится по всем i, j от* 1 *до k, т.е. расстояние Кемени между бинарными отношениями равно сумме модулей разностей элементов, стоящих на одних и тех же местах в соответствующих им матрицах.*

Легко видеть, что расстояние Кемени — это число несовпадающих элементов в матрицах $\|a(i, j)\|$ и $\|b(i, j)\|$.Расстояние Кемени основано на некоторой системе аксиом. Эта система аксиом и вывод из нее формулы для расстояния Кемени между упорядочениями содержится в книге [7], которая сыграла большую роль в развитии в нашей стране такого научного направления, как анализ нечисловой информации [4, 6]. В дальнейшем под влиянием Кемени были предложены различные системы аксиом для получения расстояний в тех или иных нужных для социально-экономических исследований пространствах, например, в пространствах множеств [4].

Медиана Кемени и законы больших чисел. С помощью расстояния Кемени находят итоговое мнение комиссии экспертов. Пусть $A_1, A_2, A_3, \ldots, A_p$ — ответы p экспертов, представленные в виде бинарных отношений. Для их усреднения используют **медиану Кемени** Argmin $\sum D(A_i, A)$, где Argmin — то или те значения A, при которых достигает минимума указанная сумма расстояний Кемени от ответов экспертов до текущей переменной A, по которой и проводится минимизация. Таким образом,

$$\sum D(A_i, A) = D(A_1, A) + D(A_2, A) + D(A_3, A) + \ldots + D(A_p, A).$$

Кроме медианы Кемени, используют **среднее по Кемени,** в котором вместо $D(A_i, A)$ стоит $D^2(A_i, A)$. Медиана Кемени — частный случай определения эмпирического среднего в пространствах нечисловой природы. Для нее справедлив закон больших чисел, т.е. эмпирическое среднее приближается при росте числа составляющих (т.е. p — числа слагаемых в сумме), к теоретическому среднему:

$$\text{Argmin} \sum D(A_i, A) \to \text{Argmin M}\, D(A_1, A).$$

Здесь М — символ математического ожидания. Предполагается, что есть основания рассматривать ответы p экспертов $A_1, A_2, A_3, \ldots, A_p$ как независимые одинаково распределенные случайные элементы (т.е. как случайную выборку) в соответствующем пространстве произвольной природы, например, в пространстве упорядочений или отношений эквивалентности. Систематически эмпирические и теоретические средние и соответствующие законы больших чисел изучены в ряде работ (см., например, [4, 6]).

Законы больших чисел показывают, во-первых, что медиана Кемени обладает *устойчивостью* по отношению к незначительному изменению состава экспертной комиссии; во-вторых, при увеличении числа экспертов она *приближается к некоторому пределу.* Его естественно рассматривать как *истинное мнение* экспертов, от которого каждый из них несколько отклонялся по случайным причинам. Рассматриваемый здесь закон больших чисел является обобщением известного в статистике «классического» закона больших чисел. Он основан на иной математической базе — теории оптимизации, в то время как «классический» закон больших чисел использует суммирование. Упорядочения и другие бинарные отношения нельзя складывать, поэтому приходится применять иную математику.

Вычисление медианы Кемени — задача целочисленного программирования. В частности, для ее нахождения используется различные алгоритмы дискретной математики, в частности, основанные на методе ветвей и границ. Применяют также алгоритмы, основанные на идее случайного поиска, поскольку для каждого бинарного отношения нетрудно найти множество его соседей.

Рассмотрим упрощенный пример вычисления медианы Кемени. Пусть дана квадратная матрица (порядка 9) попарных расстояний для множества бинарных отношений из 9 элементов A_1, A_2, A_3, ..., A_9 (см. табл. 3). Пусть требуется найти в этом множестве *медиану* для множества из 5 элементов $\{A_2, A_4, A_5, A_8, A_9\}$.

Таблица 3

Матрица попарных расстояний

0	2	13	1	7	4	10	3	11
2	0	5	6	1	3	2	5	1
13	5	0	2	2	7	6	5	7
1	6	2	0	5	4	3	8	8
7	1	2	5	0	10	1	3	7
4	3	7	4	10	0	2	1	5
10	2	6	3	1	2	0	6	3
3	5	5	8	3	1	6	0	9
11	1	7	8	7	5	3	9	0

В соответствии с определением медианы Кемени следует ввести в рассмотрение функцию

$$C(A) = \sum d(A_i, A) = d(A_2, A)+d(A_4, A)+d(A_5, A)+d(A_8, A)+d(A_9, A),$$

рассчитать ее значения для всех A_1, A_2, A_3, ..., A_9 и выбрать наименьшее. Проведем расчеты:

$$C(A_1) = d(A_2, A_1)+d(A_4, A_1)+d(A_5, A_1)+d(A_8, A_1)+d(A_9, A_1) = 2 + 1 +7 +3 +11 = 24,$$
$$C(A_2) = d(A_2, A_2)+d(A_4, A_2)+d(A_5, A_2)+d(A_8, A_2)+d(A_9, A_2) = 0 + 6 + 1 + 5 + 1 = 13,$$
$$C(A_3) = d(A_2, A_3)+d(A_4, A_3)+d(A_5, A_3)+d(A_8, A_3)+d(A_9, A_3) = 5 + 2 + 2 + 5 +7 = 21,$$
$$C(A_4) = d(A_2, A_4)+d(A_4, A_4)+d(A_5, A_4)+d(A_8, A_4)+d(A_9, A_4) = 6 + 0 + 5 + 8 + 8 = 27,$$
$$C(A_5) = d(A_2, A_5)+d(A_4, A_5)+d(A_5, A_5)+d(A_8, A_5)+d(A_9, A_5) = 1 + 5 + 0 +3 + 7 = 16,$$
$$C(A_6) = d(A_2, A_6)+d(A_4, A_6)+d(A_5, A_6)+d(A_8, A_6)+d(A_9, A_6) = 3 + 4 + 10 + 1 + 5 = 23,$$
$$C(A_7) = d(A_2, A_7)+d(A_4, A_7)+d(A_5, A_7)+d(A_8, A_7)+d(A_9, A_7) = 2 + 3 +1 + 6 + 3 = 15,$$
$$C(A_8) = d(A_2, A_8)+d(A_4, A_8)+d(A_5, A_8)+d(A_8, A_8)+d(A_9, A_8) = 5 + 8 + 3 + 0 +9 = 25,$$
$$C(A_9) = d(A_2, A_9)+d(A_4, A_9)+d(A_5, A_9)+d(A_8, A_9)+d(A_9, A_9) = 1 + 8 + 7 + 9 + 0 = 25.$$

Из всех вычисленных сумм наименьшая равна 13, и достигается она при $A=A_2$, следовательно, медиана Кемени — это множество $\{A_2\}$, состоящее из одного элемента A_2.

Литература

1. *Орлов, А.И.* Экспертные оценки / А.И. Орлов // Заводская лаборатория. — 1996. — Т. 62. — № 1. — С. 54–60.

2. *Горский, В.Г.* Метод согласования кластеризованных ранжировок / В.Г. Горский, А.И. Орлов, А.А. Гриценко // Автоматика и телемеханика. — 2000. — № 3. — С. 159–167.

3. *Шрейдер, Ю.А.* Равенство, сходство, порядок / Ю.А. Шрейдер. — Москва : Наука, 1971.

4. *Орлов, А.И.* Устойчивость в социально-экономических моделях. — Москва : Наука, 1979. — 296 с.

5. Менеджмент : учебное пособие / под редакцией Ж.В. Прокофьевой. — Москва : Знание, 2000. — 288 с.

6. *Орлов А.И.* Современная прикладная статистика / А.И. Орлов // Заводская лаборатория. — 1998. — Т. 64. — № 3. — С. 52–60.

7. *Кемени, Дж.* Кибернетическое моделирование: Некоторые приложения / Дж. Кемени, Дж. Снелл. — Москва : Советское радио, 1972. — 192 с.

8. *Орлов, А.И.* Эконометрика : учебное пособие / А.И. Орлов. — Москва : Экзамен, 2002. — 576 с.

9. *Орлов, А.И.* Менеджмент в техносфере / А.И. Орлов, В.Н. Федосеев. — Москва : Академия, 2003. — 404 с.

Контрольные вопросы и задачи

1. Почему необходимо применение экспертных оценок при решении технических, организационных, экономических, экологических и иных проблем?

2. Какие стадии экспертного исследования выделяет менеджер — организатор такого исследования?

3. По каким основаниям классифицируют различные варианты организации экспертных исследований?

4. Какова роль диссидентов в различных видах экспертиз?

5. Какой вид могут иметь ответы экспертов?

6. Чем метод средних арифметических рангов отличает от метода медиан рангов?

7. Почему необходимо согласование кластеризованных ранжировок и как оно проводится?

8. В чем состоит проблема согласованности ответов экспертов?

9. Как бинарные отношения используются в экспертизах?

10. Как бинарные отношения описываются матрицами из 0 и 1?

11. Что такое расстояние Кемени и медиана Кемени?

12. Чем закон больших чисел для медианы Кемени отличается от «классического» закона больших чисел, известного в статистике?

13. В табл. 4 приведены упорядочения 7 инвестиционных проектов, представленные 7 экспертами.

Таблица 4

Упорядочения проектов экспертами

Эксперты	Упорядочения
1	1 < {2,3} < 4 < 5 < {6,7}
2	{1,3} < 4 < 2< 5< 7 < 6
3	1 < 4 < 2 < 3 < 6 < 5 < 7
4	1 < {2, 4} < 3 < 5 < 7 <6
5	2 < 3 < 4 < 5 <1 <6 <7
6	1 < 3 < 2 < 5 < 6 < 7 < 4
7	1 < 5 < 3 < 4 < 2 < 6 < 7

Найдите:

а) итоговое упорядочение по средним арифметическим рангам;

б) итоговое упорядочение по медианам рангов;

в) кластеризованную ранжировку, согласующую эти два упорядочения.

14. Выпишите матрицу из 0 и 1, соответствующую бинарному отношению (кластеризованной ранжировке) 5 < {1, 3} < 4 < 2 < {6, 7}.

15. Найдите расстояние Кемени между бинарными отношениями — упорядочениями A = [3< 2 <1< {4,5}] и B = [1 < {2, 3} < 4 < 5].

16. Дана квадратная матрица (порядка 9) попарных расстояний (мер различия) для множества бинарных отношений из 9 элементов A_1 , A_2 , A_3, ..., A_9 (табл. 5). Найдите в этом множестве медиану для множества из 5 элементов {A_2, A_3 , A_5 , A_6 , A_9}.

Таблица 5

Попарные расстояния между бинарными отношениями

0	5	3	6	7	4	10	3	11
5	0	5	6	10	3	2	5	7
3	5	0	8	2	7	6	5	7
6	6	8	0	5	4	3	8	8
7	10	2	5	0	10	8	3	7
4	3	7	4	10	0	2	3	5
10	2	6	3	8	2	0	6	3
3	5	5	8	3	3	6	0	9
11	7	7	8	7	5	3	9	0

Темы докладов, рефератов, исследовательских работ

1. Роль экспертных методов в менеджменте.
2. Организация различных видов экспертных исследований.
3. Сравнение очных и заочных вариантов работы экспертов.
4. Методы средних баллов.
5. Согласование кластеризованных ранжировок.
6. Методы теории люсианов в экспертных оценках [1].
7. Классификация мнений экспертов и проверка согласованности.
8. Использование люсианов в теории и практике экспертных оценок.
9. Формирование итогового мнения комиссии экспертов.
10. Расстояние по Кемени и медиана Кемени в экспертных оценках.
11. Законы больших чисел в пространствах нечисловой природы.

ГЛАВА 12. СТАТИСТИЧЕСКИЕ МОДЕЛИ В МЕДИЦИНЕ[7]

Глава посвящена компьютерно-статистической революции в медицине, прежде всего порожденным этой революцией интеллектуальным инструментам научных медицинских исследований, основанным на современных статистических методах и моделях.

Всего лишь за несколько последних десятилетий в довольно значительной части медицинской науки произошла настоящая революция, изменившая не что иное, как самую основу практической медицины — мышление врача. Причем это изменение было столь глобальным, что во многом перевернуло применявшиеся ранее теории, едва ли не в корне изменило те принципы, согласно которым врачи всего мира долгие годы лечили своих пациентов. И в этом изменении, в этой революции основную роль сыграло развитие информационных технологий, основанных на интенсивном использовании высоких статистических технологий анализа медицинских данных. Именно появление систем, могущих с достаточной скоростью и без титанических затрат труда осуществить сложные математические расчеты, привело к тому, что принципы так называемой «доказательной медицины» вошли в широкую практику, а результаты выполненных на этих принципах исследований стали применяться во всем мире.

Что такое «доказательная медицина»? Об этом ниже рассказано на уровне, достаточном для практического применения ее результатов, для выхода на современный фронт научных исследований. В главе рассмотрены основные статистические методы и модели, используемые в настоящее время в научных медицинских исследованиях, как классические, так и высокие статистические технологии последних лет.

12.1. Новое компьютерно-статистическое мышление врача

Медицина — одна из самых консервативных наук. И это понятно — ведь в ней речь идет о жизни и здоровье человека. Новые медицинские теории, новые технологии лечения и выставления диагноза обычно принимались медицинским сообществом после длительного и бурного обсуждения.

Клиническое мышление. Как врач принимает решение о том, какой диагноз поставить своему пациенту, какое лечение следует назначить, каково может быть будущее больного при том или ином варианте лечения?

[7] Глава написана совместно с врачом-кардиологом А.А.Орловым.

На протяжении многих десятилетий ход мыслей хорошего врача представлялся так. Врач, осматривая пациента, должен был вспомнить из институтского курса, прочитанных книг и личного опыта, какому заболеванию соответствуют те признаки болезни — симптомы — которые имеются у пациента. А потом, вспомнив, как это заболевание развивается, применить те лекарства, которые действуют на механизм возникновения заболевания. Воздействовать на ход болезни на основе знания того, как, из-за чего она возникает и каким образом течет. Сведения же о механизмах возникновения и течения болезней преподавались в медицинских институтах как основа всего курса обучения: многочисленные предметы, такие, как «патофизиология», «патанатомия», как раз изучению таких механизмов и посвящены.

«*Познай, из-за чего болезнь начинается, как она развивается, как течет — и ты поймешь, как ее надо лечить, и к чему она приведет*» — вот «классический» принцип действия врача, который главенствовал во всей медицинской науке практически с момента ее зарождения. Именно познание, максимально глубокое познание механизмов развития болезни считалось путем к получению возможности излечивать эту болезнь. А само лечение основывалось на знании того, как то или иное химическое вещество, то или иное воздействие влияет именно на механизм развития болезни и мешает его ходу, не дает болезни развиваться дальше. Наиболее эффективными средствами лечения признавались те, которые наиболее сильно нарушали ход течения заболевания, приводя тем самым к исчезновению болезни как таковой.

И многовековой опыт подтверждал этот принцип. Познав, что болезнетворные микроорганизмы вызывают инфекционные болезни, врачи научились предотвращать инфекции, ранее иной раз выкашивающие целые города и страны. Познав, как эти микроорганизмы устроены, как они развиваются, врачи разработали вещества, их уничтожающие. Познав, что кровообращение в ноге или руке может прекратиться из-за возникновения в их основных сосудах сгустка крови, врачи придумали, как этот сгусток оттуда извлечь и тем самым очистить пути для крови, несущей кислород и питание тканям.

Изучение же механизмов велось различными способами. Ученые сравнивали здоровые и больные органы путем исследования тканей под микроскопом, в химических лабораториях, на основе наблюдений делали выводы. Болезни, причины возникновения которых были известны, вызывались у лабораторных животных — тем самым исследовалось течение патологического процесса. Новые лекарственные препараты также вначале испытывались в лаборатории на экспериментальных материалах, затем — на животных, затем — на добровольцах, небольших контингентах людей, после чего в случае успешного прохождения всех исследований выпускались в широкую продажу.

Количество людей, участвующих в исследованиях, например, действия новых лекарств или новых методов лечения, измерялось десятками и редко сотнями. Работать с большими группами было крайне сложно — чтобы провести обработку данных о состоянии здоровья тысяч человек с помощью статистических методов, требовался огромный труд множества ученых — и поэтому крупномасштабные исследования практически не проводились.

Основная задача обучения врача в медицинском институте виделась в том, чтобы сформировать у нового специалиста «клиническое мышление». Иными словами — научить врача прежде всего видеть весь ход болезни у своего пациента, отчетливо представлять себе весь механизм ее возникновения, все подробности хода и — главное! — лечить пациента, исходя из этого представления о том, какие процессы протекают в организме у больного. «Познай процесс — назначай лечение» — таким был девиз каждого врача, выходящего из стен медицинской академии или университета.

Компьютеры в медицине. Однако развитие информационных технологий изменило этот традиционный подход.

Примерно к середине XX в. в обиход крупных научных центров стали входить первые компьютеры. Практически в каждом крупном университете мира появились машины, способные за несколько секунд или минут выполнить расчеты, над которыми раньше пришлось бы трудиться десяткам людей в течение сотен или даже тысяч дней. Те же машины позволили значительно быстрее и проще накапливать информацию различного типа на своих магнитных носителях, а с развитием компьютерных сетей и объединением их в Интернет столь же простой и быстрой стала и передача данных между компьютерами, университетами, континентами.

И у ученых появилась идея — а почему бы ни использовать открывшиеся возможности? Заняться крупномасштабными исследованиями — например, собрать данные о десятке тысяч человек, страдающих каким-либо заболеванием и принимающих тот или иной лекарственный препарат, и сравнить их с данными о таком же десятке тысяч человек, имеющих тот же диагноз, но принимающих другое лекарство? Или даже с теми больными, кто вообще никаких лекарств не пьет? Или даже сделать еще более интересное исследование — растянутое по времени. Выделить несколько больших групп людей, различающихся по каким-нибудь критериям (например, по уровню артериального давления), и посмотреть, что с людьми в этих группах будет через пять-десять лет. Сколько из них умрет, сколько заболеет еще более тяжелыми и серьезными недугами, а сколько, наоборот, выздоровеет. Ведь тогда можно будет достоверно сказать, что, скажем, тот или иной фактор, имеющийся в жизни человека, увеличивает его риск получить какие-либо серьезные проблемы со здоровьем.

И подобные исследования стали проводиться. Для них были разработаны специальные правила — кое-что о них вы можете прочитать ниже. Ученые создали и технологии анализа накопленных данных — особенно большую роль в их разработке сыграли специалисты в области прикладной математической статистики. Статистические методы позволяют определить, например, действительно ли именно медицинское вмешательство изменяет состояние здоровья пациентов или видимое изменение объясняется всего лишь волей случая. Или выполнить иные сравнения, общей целью которых было выяснение наличия или отсутствия лишь одного факта — эффективности медицинского вмешательства.

Применение статистических методов на больших объемах данных требовало весьма значительного количества расчетов. Но тут на помощь ученым пришли компьютеры: выполняя за минуты числовые операции, на которые человеку потребовались бы месяцы и годы работы, они позволили получить результаты анализа данных весьма быстро, да и хранить данные о состоянии здоровья десятков и сотен тысяч людей стало нетрудно. Но итоги многих исследований оказались весьма интересными...

Примечание. Прежде чем говорить о результатах исследований, отметим, что компьютеры во много раз ускоряют расчеты, однако сбор исходных данных для расчетов по-прежнему весьма трудоемок. Многие используемые при подобных расчетах статистические методы, например, методы дискриминантного и кластерного анализов, появились сравнительно недавно, лишь в 30–50-е гг. XX в. или даже позже. Новые статистические методы активно разрабатываются и в настоящее время.

Результаты статистического анализа. Они действительно были интересными.

В состоянии после инфаркта миокарда весьма опасным осложнением является аритмия сердца. Нарушение регулярности сердечных сокращений, резкое их учащение или урежение, изменение характера сокращения может привести к нарушению снабжения организма кровью и, следовательно, смерти. Отсюда вытекает совершенно естественный вывод — с аритмией надо бороться! Во-первых, ликвидируя возникшие расстройства — а, во-вторых, не допуская появления новых, занимаясь профилактикой возникновения аритмии. На основании медицинских теорий физиологии сердца, на основании данных о том, как в человеческом сердце возникает электрический импульс, приводящий к сокращению сердечной мышцы, были разработаны лекарства, которые, согласно этим теориям, не должны были допускать расстройства сердечных сокращений. Одним из таких препаратов был флекаинид. Лекарство было испытано на животных и небольших группах добровольцев — его эффективность для предотвращения аритмий была доказана: флекаинид действительно уменьшал частоту аритмий у больных с инфарктом миокарда. И он был рекомендован для массового применения.

Однако нашлись ученые, пожелавшие посмотреть: а как флекаинид влияет не на возникновение аритмий, а на выживаемость больных инфарктом миокарда? Ведь, в самом деле, именно выживаемость больных важна для врача: врачу нужно спасти жизнь пациенту, а ликвидация аритмии суть средство к этому. И началось крупное исследование: ученые стали собирать данные о пациентах с инфарктом миокарда, одни из которых получали для лечения флекаинид, а другие — нет, и сравнивать, сколько пациентов в каждой из этих групп осталось в конце концов в живых. Бесспорно, ученые приняли меры к тому, чтобы между группами не было различия в составе больных — скажем, в одной больше здоровых, а в другой меньше: при большом числе обследуемых это обеспечить было нетрудно.

Примечание. Для проверки однородности двух групп больных (двух выборок) разработаны различные статистические методы (см. главу 5). Если же группы различны, то можно провести т.н. «стандартизацию выборки», чтобы исключить влияние различий.

Результаты ошеломили ученых. Флекаинид превосходно предотвращал аритмии после инфаркта, это было однозначно ясно. Однако вот смертность в группе пациентов, получавших флекаинид, оказалась в несколько раз выше, чем среди тех, кто такого лечения не получал. То есть препарат, уменьшая частоту возникновения аритмий, убивал пациентов.

Почему так оказалось — медицинская наука не знала. Из механизмов действия препарата не следовало ничего, что могло бы приводить к такому исходу. Никакая теория не давала объяснения. Однако факт оставался фактом — больные, получающие, казалось бы, совершенно необходимый им препарат, при прочих равных условиях умирали чаще, чем те, кто такой препарат не получал. И итогом исследования стал вполне закономерный запрет на использование флекаинида для профилактики аритмий после инфаркта миокарда — как бы он ни был хорош в теории, но он должен излечивать больных, а не убивать их.

Стали выясняться и другие подобные факты — причем относящиеся не только к действию лекарств на организм человека (тут, как ни крути, все же можно было предположить наличие каких-то неизвестных физиологических механизмов), но и к, казалось бы, более чем ясным ситуациям. Например, путем многочисленных исследований трупов ученые установили, что инсульт — нарушение снабжения кровью участка мозга — весьма часто возникает в том месте, где артерия, питающая участок мозга, имеет сужение вследствие поражения атеросклерозом. Решение проблемы казалось более чем ясным — следует провести к этой артерии после сужения обходной путь подачи крови от другой артерии, сделав хирургическую операцию! Так и стали делать. Операции были успешные — кровоснабжение в участке мозга, ранее снабжаемом суженной артерией, полностью восстанавливалось.

А потом ученые проанализировали накопленный объем данных о здоровье прооперированных пациентов. И выяснилось, что никакого увеличения продолжительности жизни по сравнению с традиционным лечением лекарственными препаратами эта операция не дает. Даже инсультов у пациентов, подвергшихся операции, меньше не стало. Операция, которая по всем медицинским канонам должна помогать, оказалась бесполезной или даже вредной.

И подобные итоги получались весьма часто. Нет, бесспорно, весьма большая часть клинико-статистических исследований давала результаты, соответствующие «классической» медицинской науке — эффективность многих лекарств без всяких проблем подтверждалась. Но наличие данных, полученных при обследовании многих тысяч пациентов, которые прямо противоречили этой самой «науке», не позволяло жить и думать по-старому. Ибо продолжать держаться за, казалось бы, безупречные научные теории, не обращая внимания на итоги практических исследований, означало бы обрекать на смерть тысячи пациентов...

Новые принципы. Результаты проведенных исследований заставили ученых пересмотреть свои взгляды на медицинскую науку. Тот факт, что, казалось бы, совершеннейшая медицинская теория, подтвержденная в лабораториях, нередко совершенно не находила подтверждения на практике, при исследованиях десятков тысяч человек, заставил ученых выработать новые принципы мышления врача. Вкратце их можно сформулировать следующим образом.

1. В большинстве случаев врач не может с абсолютно полной уверенностью судить о том, как скажется на здоровье пациента то или иное лечение или, наоборот, отсутствие лечения. Однако оценить вероятность различных исходов вполне возможно.

И действительно — история медицины знает случаи, когда, казалось бы, совершенно безнадежные больные излечивались от своего недуга по совершенно непонятным причинам. Например, описаны ситуации, когда бесследно исчезали раковые опухоли. Увы, немало случаев и обратных: когда, казалось бы, совершенно правильное лечение не предотвращало гибель пациента. Однако и те, и другие варианты довольно редки; в то же время ситуации, когда лечение приводит к выздоровлению, а отказ от лечения — к ухудшению состояния здоровья, куда как более частые.

Поэтому оценивать возможные результаты того или иного лечения наиболее правильно именно с точки зрения их вероятности. Говорить не о том, что «вот это, по теории, должно привести к такому исходу, а вот это — к этому», а о том, что «при наличии этого фактора вероятность такого исхода такая-то». Повышенное артериальное давление не обязательно приведет к смерти человека: есть люди, живущие с высоким артериальным давлением по сто лет и больше — но вот риск смерти при высоком артериальном давлении куда как больше, чем при нормальном.

2. Эту вероятность лучше всего оценивать на основе опыта, накопленного при наблюдении за другими пациентами и их лечении. Более того — именно полученная таким образом оценка и будет наиболее верной.

В самом деле — именно опыт есть критерий теории, и лишь та теория верна, которая подтверждается практикой. А в том деле, в котором речь идет о жизни человека, опыт и только опыт может подсказать правильное решение. Недаром практически с начала развития медицинской науки ценились не столько «умные», сколько «опытные» врачи: пациенты во все времена шли прежде всего к тем врачам, которые на основании своего опыта знали, что надо делать с пациентом, чтобы он выздоровел, а не к тем, которые знали то же из теоретических изысканий.

Примечание. Для практического применения этой идеи надо уметь оценивать вероятность того или иного исхода. А также планировать объем наблюдений, необходимый для достижения заданной точности. Статистические методы для решения этих задач обсуждаются ниже.

3. Однако опыт каждого врача субъективен. Кроме того, даже опыт, накопленный в пределах одного медицинского учреждения, одной научной школы не обязательно будет объективным.

Однажды на медицинском конгрессе встретились врачи двух больниц: сельской и областной. И зашел у них спор о том, насколько хорош один известный антибиотик, насколько он хорошо излечивает больных пневмонией. Сельский врач был уверен в его эффективности: практически все больные с пневмонией, поступавшие в ту больницу, где он работал, быстро и хорошо этим антибиотиком вылечивались. А врач областной больницы очень сильно возмутился и сказал, что этот антибиотик бесполезен: он не смог вылечить им ни одного из больных с пневмонией, кто поступил в его отделение, и был вынужден давать этим больным другие лекарства.

Они так и не пришли к единому мнению — а все дело было в том, что в областную больницу привозили как раз тех больных, которым не смогли помочь сельские врачи, больных, у которых обсуждаемый антибиотик оказался неэффективным. Доля таких пациентов в сельских больницах была невелика, но все они в конце концов собирались в областной больнице — потому ее врач так критически об этом антибиотике и высказывался. И ведь оба были правы — на основании собственного опыта и опыта своих больниц. Поэтому, как видно, основываться на опыте не так просто...

4. Поэтому для того, чтобы накопить достоверные данные, необходимо выполнять исследования согласно определенным правилам, исключающим возникновение ошибок.

Такие правила известны. Разработаны они учеными-математиками, которые создали целую отрасль своей науки, посвященную в том числе и проблемам достоверности результатов анализа тех или иных данных — прикладную математическую статистику.

Примечание. Иногда эту отрасль науки называют анализом данных. В настоящее время анализ данных — весьма развитое научное направление. Некоторые входящие в него «статистические методы» будут рассмотрены ниже.

Правила описывают практически все аспекты исследований, которые хоть как-то могут повлиять на достоверность их результата: от методики набора пациентов для исследования до алгоритмов обработки накопленной информации. Позже мы рассмотрим некоторые из этих правил. Если же вас интересуют более подробные сведения о них, то в книгах, перечисленных в конце главы, вы найдете исчерпывающую информацию.

Примечание. Отметим здесь только один прием — двойной слепой метод. Он применяется, когда необходимо выяснить эффективность лекарственного препарата. Мешающим фактором является психологический — состояние некоторых больных улучшается просто от сознания того, что их лечат. Поэтому всем пациентам дают одинаковые по виду таблетки, одним — с лекарством, другим — пустышки (т.н. плацебо). Это — слепой метод. А двойной слепой —

когда сами лечащие врачи не знают, кто из их больных получает лекарство, а кто — плацебо. Знают только организаторы эксперимента.

Итак, отныне врач, назначая лечение, стал думать прежде всего не о том, как данное лекарство подействует на биохимические процессы в организме согласно имеющимся медицинским теориям, а о том, **будет ли данное лекарство эффективно** при той болезни, что есть у пациента. А думать об этом он мог **на основе анализа данных** крупномасштабных исследований состояния здоровья людей, проведенных его коллегами-врачами, а, может, и им самим. Изменился сам подход врача к назначению лечения — в основе этого подхода стали лежать не только научные концепции, а прежде всего накопленный опыт, который зачастую ни в какие концепции не вписывался.

Новый подход назвали *Evidence-based medicine* (англ.) — «медицина, основанная на доказательствах», или «доказательная медицина». Впервые сделано это было в 1990 г. учеными из канадского университета Мак-Мастера. В русском языке больше прижилось словосочетание «доказательная медицина», которое в настоящее время и используется для именования новых принципов лечения больных.

Знахарство XXI века. Если внимательно приглядеться, то описанный выше подход весьма похож на работу знахаря. Фактически врач из ученого, который назначает лекарства, исходя из теории о том, как оно действует, превратился в наследника деревенских знахарей, которые назначали лекарства исключительно на основе накопленных ими самими и переданных от предыдущих поколений знаний о том, «при каких хворях та или иная трава помогает».

Примечание. Изучение связи между явлениями и процессами, между случайными переменными — предмет ряда разделов прикладной статистики. Корреляционный анализ, регрессионный анализ, дисперсионный анализ — все эти статистические методы нацелены на обнаружение и изучение связи, восстановление зависимости по результатам наблюдений, измерений, испытаний, опытов. Мы поговорим об их роли в доказательной медицине в этой главе ниже.

Можно сказать, что медицина стала знахарством — изменился основной принцип этой науки. Однако знахарством не примитивным, а основанным на современных научных принципах и технических возможностях. Опыт, накапливаемый в современных исследованиях, можно считать уже полностью достоверным, полностью отражающим реальность. Древние философы говорили, что развитие идет по спирали: на каждом из ее витков происходит повторение старого, но на новом уровне развития. Так и в медицине: она, пройдя эпоху стройных научных теорий, вновь пришла к старому доброму принципу опоры на практический опыт — но опыт, уже накопленный не одним или даже десятью человеками, а всем медицинским сообществом, и не на сотне или тысяче пациентов, но на десятках и сотнях тысяч случаев. И проанализированный не на основании представлений одного — двух человек, но на основе отработанных и точных принципов научного статистического анализа.

Бесспорно, научные теории никуда не делись. Как бы ни был эффективен «знахарский» подход, но он практически не позволяет двигаться вперед: как искать новые лекарства, если не знать теорию их воздействия на организм? Как разрабатывать новые методы лечения, если не знать, на что они должны действовать, на какой этап болезни, на какой ее аспект? Новые теории, новые взгляды на устройство человеческого организма и его работу по-прежнему создаются, проверяются в лабораториях, обсуждаются, опровергаются и преобразуются в новые. Но вот в практику теоретические рекомендации идут уже лишь после проверки на опыте, на опыте реальном, а не теоретическом. И нередко оказывается так, что лекарство или лечебное воздействие, которое должно было бы в теории действовать так-то, на практике действует совсем иначе. Что поделать: человеческий организм весьма сложен, и силы нашего разума пока отнюдь не всегда хватает для познания этого удивительного творения...

И весьма большую роль в свершившейся революции сыграли информационные технологии. Именно появление и развитие компьютеров дало человечеству в руки те возможности, которые необходимы для проведения крупномасштабных исследований состояния здоровья людей и влияния на него различных факторов. Раньше для проведения подобных мероприятий потребовались бы месяцы или даже годы работы — подготовить данные на несколько десятков тысяч человек, рассчитать результаты по статистическим формулам вручную было архисложно, и потому никто этим и не занимался. Но компьютер, даже относительно слабый, справится с такой задачей за считанные минуты, в крайнем случае — часы, выделив из сотен тысяч данных измерений и оценок искомые закономерности, зависимости, влияния. Можно даже сказать, что именно развитие «доказательной медицины» и является самым что ни на есть «компьютерным» результатом развития компьютерных технологий. Вкупе, скажем, с расчетами орбит межпланетных станций. Ведь само слово «компьютер» означает «вычислитель», и именно возможность компьютеров быстро и без особых затрат сил и средств производить огромный объем вычислений, то есть, так сказать, выполнять то, ради чего они и были созданы, и обусловила быстрый рост числа выполненных исследований.

Новое медицинское мышление. Принципы «доказательной медицины» стали занимать все большее место в мышлении врача. Если ранее при назначении лечения врач отдавал предпочтение тем препаратам, которые «**должны помогать** согласно данным науки», то теперь — тем, которые **достоверно помогают** согласно накопленным данным (помогают с вероятностью, близкой к 1). В процессе совершенствования врача все большее место стало занимать освоение им результатов новых исследований. Если ранее «хороший врач» был обязан превосходно владеть медицинской теорией, знать, какие процессы протекают в организме при том или ином заболевании, то теперь все это, бесспорно, тоже требуется, однако современный врач прежде всего обязан следить за данными новых научных исследований, оперативно с ними знакомясь и применяя в своей практике (разумеется, с учетом индивидуальных особенностей каждого пациента).

Ни в коем случае не стоит думать, что «доказательная медицина» полностью исключила понимание механизма заболевания из процесса лечения больного. Никому не придет в голову проверять в исследованиях, скажем, основные принципы лечения ожогового шока или травм: то, что при наличии у травмированного человека кровотечения его необходимо остановить, а при обезвоживании организма ввести в кровеносную систему жидкость, ясно всем, и доказательств это не требует. Никто также не будет доказывать, что, например, пневмонию надо лечить, а не оставлять на «самостоятельное излечение»: исследования для такого доказательства обернутся множеством смертей больных, а подтвердят лишь очевидный факт. А вот определить, какие особенности лечения ожогового шока более эффективны, какими нитками лучше всего сшивать разорванные сосуды, какой жидкостью лучше всего проводить восстановление водного баланса организма, какой из антибиотиков наиболее эффективен при той или иной форме пневмонии, методы «доказательной медицины» вполне способны.

Медицина не стоит на месте. Появляются и публикуются новые данные исследований, создаются новые рекомендации, разрабатываются новые лекарства. И вполне возможно, что спираль развития этой науки в конце концов обернется и дальше. Будет создана новая, всеобъемлющая теория, объясняющая все процессы, происходящие в человеческом организме, и позволяющая абсолютно достоверно предсказать эффект любого лекарства, любого вмешательства в жизнедеятельность организма, которая вновь заставит врачей в своей работе прежде всего опираться не на опыт, а на эту теорию. До этого еще далеко — но кто знает, может быть, и в ее создании суждено сыграть свою роль наследникам современных компьютеров?

12.2. Методы «доказательной медицины»

Для того, чтобы данные, получаемые в результате исследований, отражали реальность наиболее адекватно, научные работники в области медицинских наук разработали правила, по которым эти исследования следует проводить.

Например, первое и едва ли не основное правило — это наличие так называемой «**контрольной группы»**, или «группы сравнения». Если надо оценить, насколько эффективно то или иное медицинское вмешательство — скажем, прием какого-либо лекарства или произведение хирургической операции — то всех пациентов, участвующие в исследовании, следует разделить как минимум на две группы — «группу вмешательства» (ее называют также экспериментальной группой или «группой лечения») и «группу сравнения». Пациенты из обеих групп должны быть одинаково обследованы, находиться в одинаковых условиях и вообще в идеале различаться между собой лишь одним — наличием или отсутствием вмешательства. В самом деле: в мире есть множество факторов, которые могут повлиять на пациентов так, что и не отличить: был ли тот или иной результат влиянием лечебного вмешательства или этих самых факторов. Например, если надо узнать, спасает ли то или иное лекарство жизни больных артериальной гипертонией, мало просто давать его группе больных несколько лет. Больные все равно будут умирать от разных причин, и чтобы определить, задержало ли лекарство эти смерти, нужно сравнить число таких исходов с числом смертей в другой группе, больные из которой данного лекарства не получали. Если смертность в группе вмешательства будет меньше, чем в группе контроля, то можно говорить о полезности этого вмешательства.

Примечание. В современной прикладной статистике разработано много методов, реализующих идею сравнения двух групп, выявления эффекта лечебного воздействия. В некоторых из них контрольная группа как фиксированный объект исчезает, например, при использовании скользящего контроля как варианта бутстрепа.

Технология распределения пациентов по группам получила название **рандомизации**. При хорошо выполненной рандомизации в каждой из групп доля пациентов с тем или иным свойством будет примерно одинакова: одинаков процент мужчин и женщин, одинаков процент лиц того или иного возраста, одинаково соотношение пациентов с различными степенями тяжести исследуемого заболевания. Способы проведения данного процесса описаны в специальной литературе, однако для полноценной рандомизации число пациентов должно быть достаточно большим — измеряться по меньшей мере тысячами.

Примечание. Плохая рандомизация «лечится» стандартизацией выборки. Сравнение групп можно проводить с помощью статистических методов проверки однородности. В научных медицинских исследованиях до сих пор часто применяют критерий Стьюдента для проверки однородности двух выборок. Однако он опирается на два предположения — о нормальном распределении результатов наблюдений и о равенстве дисперсий для двух групп. Ни то, ни другое предположение обычно не выполнено в научных медицинских исследованиях. Поэтому современная прикладная статистика рекомендует вместо критерия Стьюдента применять иные критерии, а именно, критерии Крамера-Уэлча, непараметрические критерии Вилкоксона, Смирнова, Лемана — Розенблатта (см. главу 5).

При лечении людей лекарственными препаратами весьма большое значение имеет психологический настрой человека на исцеление. Некоторые больные избавлялись от своего страдания, принимая совершенно бесполезные для них препараты: вера в то, что препарат поможет, исцеляла их куда как лучше самого препарата. Поэтому при проведении исследований лекарств нельзя просто давать его больным из группы вмешательства и не давать тем, кто входит в группу сравнения — в таком случае отличить истинное действие лекарства от его психологического эффекта будет практически невозможно. Необходимо исключить влияние психологического эффекта на результаты исследования — проще всего это сделать, если пациентам обеих групп давать совершенно одинаковые по внешнему виду, вкусу, запаху препараты, только в один из препаратов добавлять то вещество, влияние которого на человеческий организм и исследуется, а во второй — нет. В этом случае пациенты из обеих групп будут одинаково верить в то, что препарат им поможет, и эта вера не повлияет на конечный результат — оценку эффективности лечения лекарством. Препарат, который дается пациентам из контрольной группы и не содержит в себе собственно лекарства, принято называть **«плацебо»** — от латинского «пустышка», а исследование с использованием плацебо — **«слепым»**.

Однако даже при использовании плацебо могут быть проблемы. В самом деле: пациенты-то хоть и не знают, что они принимают — лекарство или «пустышку» — но врачам, проводящим исследование, это известно! И они вполне могут словом или жестом случайно дать понять пациенту, чем на самом деле его лечат, что, ясное дело, обесценит использование плацебо. Чтобы это предотвратить, исследование можно сделать **«двойным слепым»** — то есть не говорить об истинном содержимом препарата и лечащим врачам пациентов.

Оценка эффективности того или иного воздействия на организм человека можно оценивать по изменившимся параметрам жизнедеятельности. Например, по уровню артериального давления или концентрации в крови тех или иных веществ. Однако куда как более привлекательным представляется оценка по наличию или отсутствию в жизни пациента событий, имеющих для его жизни важное значение, — например, инфаркта или инсульта. Если в результате исследования получилось так, что у пациентов в группе вмешательства неблагоприятных событий было меньше, а благоприятных — больше, чем у пациентов из контрольной группы, то вмешательство, наверное, стоит назвать полезным, если же нет — то вредным. Такие события получили название **«конечных точек»** — во многих исследованиях оценка эффективности медицинского вмешательства ведется именно по ним.

Примечание. Доли благоприятных событий в группе вмешательства и в контрольной группе могут различаться по чисто случайным причинам. Необходимо установить, значимо ли различие. Для этого в прикладной статистике разработаны методы проверки однородности вероятностей (см. главу 1).

Некоторые данные доказательной медицины. Ниже вкратце приводятся некоторые факты, полученные методами «доказательной медицины», ценные тем, что их соответствие реальности можно считать полностью доказанным. Каждый из этих фактов проверен в многих десятках и сотнях тысяч наблюдений. Каждый их них — весомое доказательство практической пользы статистических методов.

Предупреждение для тех, кто не является врачом. Вы можете здесь ознакомиться с этими фактами, однако помните, что ***принимать любые лекарственные препараты можно только после консультации с врачом.***

1. *Чем выше у человека уровень артериального давления в покое, тем больше у него риск получить инсульт или инфаркт миокарда.* Эта зависимость плавная — чем выше артериальное давление, тем выше риск и, наоборот, чем ниже артериальное давление (в медицинских публикациях оно обычно обозначается аббревиатурой «АД»), тем ниже риск. Согласно общему мнению врачей мира, границей, отделяющей болезнь от здоровья, является уровень артериального давления, равный 140/90 мм рт. ст.

Примечание. Уточним краткое медицинское высказывание. Если верхнее артериальное давление превышает 140 мм рт. ст. и/или нижнее артериальное давление выше 90 мм рт. ст., то человек находится в «зоне болезни». Для такого заключения достаточно, чтобы хотя бы для одного из двух видов измерений граница была превышена. Реально о болезни можно говорить, если превышение является постоянным или почти постоянным.

Практический вывод прост — при повышенном уровне артериального давления человека надо лечить, добиваясь нормализации АД.

2. *Курение, повышенное артериальное давление и повышенный уровень холестерина в крови приводят к повышенному риску развития ишемической болезни сердца.* Если человек курит, если у него повышено артериальное давление в покое и высокий уровень холестерина в крови, то риск заболеть стенокардией или попасть в больницу с инфарктом, а то и умереть от него, у такого человека весьма высок. Не курите, следите за своим артериальным давлением и соблюдайте диету с пониженным содержанием продуктов, богатых холестерином — и не заболеете ишемической болезнью сердца.

3. *При отсутствии органического поражения сердца наличие такой аритмии сердца, как экстрасистолия, не влияет на риск смерти и появления какого-либо заболевания, а вот некоторые препараты, предотвращающие экстрасистолию, могут этот риск даже увеличить.* Экстрасистолия — это наличие у человека преждевременных сердечных сокращений; обычно она ощущается как «перебои в сердце» или неравномерность сердечных сокращений при прослушивании грудной клетки. Многие больные экстрасистолией часто пугаются своего заболевания и считают его опасным или даже смертельным — это далеко не так!

Экстрасистолию даже лечить не обязательно — наоборот, препараты, которые применяются для прекращения экстрасистолии, опасны для организма. Однако помните, что установить диагноз экстрасистолии и отличить экстрасистолию от других аритмий сердца может только врач; кроме того, вышеупомянутый факт верен только для тех, у кого никогда не было инфаркта миокарда.

4. *Назначение бета-блокаторов после инфаркта миокарда улучшает выживаемость больных.* Если после инфаркта миокарда человек будет принимать лекарства, относящиеся к группе так называемых «блокаторов бета-адренергических рецепторов» (например, атенолол, конкор, анаприлин), то его риск смерти в ближайшие 10 лет значительно снизится.

5. *Антиаритмические препараты 1С класса увеличивают смертность больных при сердечнососудистых заболеваниях.* Все лекарства, применяемые для лечения аритмий сердца, по механизму своего действия делятся на четыре класса — *1, 2, 3* и *4*, а в первом из этих классов выделяют еще три подкласса — *A*, *B* и *C*. Установлено, что препараты из подкласса *C* класса *1* (например, упомянутый выше флекаинид), увеличивают риск смерти у тех больных, которые их принимают, — несмотря на то, что уменьшают частоту возникновения аритмий.

6. *Так называемые «метаболические» препараты не дают значимого снижения смертности при лечении инфаркта миокарда.* При исследованиях процессов, происходящих в сердце, ученые выяснили, что при нарушении снабжения кровью сердечной мышцы — развитии инфаркта миокарда — практически во всей этой мышце очень сильно нарушается обмен веществ. Для того, чтобы степень выраженности этого нарушения свести к минимуму, ученые разработали специальные лекарства — например, одно из них называется «рибоксин» — и вполне естественно ожидали, что эти новые лекарства будут восстанавливать жизнедеятельность сердца после инфаркта. Это действительно происходило, но... на смертность пациентов применение метаболических препаратов никак не повлияло. То есть они оказались неэффективны — несмотря на, казалось бы, более чем правильный механизм действия.

7. *Большинство новых препаратов для лечения артериальной гипертонии не более эффективны, нежели использовавшиеся ранее.* Ученые провели исследование: сравнили эффективность старого доброго лекарства от артериальной гипертонии гипотиазида, которым гипертонию еще в семидесятых годах, и современного препарата арифона, действующего по тому же принципу, что и гипотиазид. И выяснили, что их эффект практически одинаков. Таким образом, несмотря на то, что гипотиазид принято считать устаревшим препаратом, он лечит гипертонию не хуже, чем самые современные аналогичные лекарства, а, кроме того, еще и весьма дешев — куда как дешевле модернизированных аналогов.

О публикациях по доказательной медицине. Данные новых исследований публикуются в международных медицинских журналах. В настоящее время уже сложился круг изданий, публикациям которых доверяют врачи всего мира: это, например, *Lancet, British Medical Journal, Journal of American Medical Association, New England Medical Journal, Annals of Internal Medicine, Journal of Clinical Investigations*. Есть журналы столь же высокого уровня по отдельным медицинским специальностям, например, *Hypertension, Cardiology*.

Широко известны отечественные журналы «Клиническая медицина», «Кардиология», «Российский кардиологический журнал», «Кардиоваскулярная терапия и профилактика», «Вестник Российской академии медицинских наук» и многие другие.

Современный врач для поддержания профессионального уровня обязан регулярно знакомиться с содержанием журналов, публикующих результаты новых исследованиях по его специальности — собственно, в настоящее время именно такое ознакомление и понимается под «совершенствованием врача».

Однако компьютерные технологии и тут облегчили задачу врача: появление и развитие Интернета и особенно *World Wide Web* позволило быстро доносить новые данные науки до самых отдаленных уголков планеты. В настоящее время все публикации наиболее крупных журналов мира и большинство статей, увидевших свет в серьезных медицинских изданиях различных стран, помещаются в медицинскую базу данных *Medline*, которая доступна через Сеть, а также распространяется на компакт-дисках по научным учреждениям и библиотекам. В Сети есть немало сайтов, на которых можно ознакомиться с новостями «доказательной медицины», узнать самые последние данные о результатах исследований эффективности лекарств, методов лечения, подходов к изменению образа жизни при самых различных заболеваниях.

Медицинские организации, основываясь на данных исследований, составляют рекомендации по диагностике и лечению тех или иных заболеваний — так, Всемирная Организация Здравоохранения регулярно извещает врачей мира, как, по общему мнению медицинского сообщества, стоит поступать с больными теми или иными заболеваниями, чтобы их вылечить. Такие рекомендации собирают воедино опыт множества исследований, и долгом каждого врача является знание их содержания и применение его на практике. В мире есть и некий «центр» доказательной медицины — так называемое «Кокрановское сотрудничество» (*Cochrane Collaboration*), организация, сотрудники которой взяли на себя обязательство регулярно выпускать обзоры новых исследований, проведенных в различных отраслях медицины. Собственно, и названа она по имени составителя первого обзора — Арчи Кокрана, который сделал это полвека назад. Кокрановское сотрудничество существует с 1992 г. и с тех пор снабжает своими обзорами врачей всего мира, в том числе и другие международные организации.

В век Интернета книги отнюдь не устарели. Порекомендуем только две из них [1, 2]. Эти книги по праву могут считаться кратким изложением основных принципов и технологий «медицины, основанной на доказательствах». Изучите их, если вы желаете разобраться в этом вопросе более глубоко. Первая книга содержит прежде всего описания технологии исследования и подходы к обработке медицинских данных, а вторая рассказывает в основном о мире современной научной медицинской литературы.

В Сети (в Интернете) есть сайты, на которых собираются сведения о новых исследованиях, проведенных согласно правилам «доказательной медицины».

12.3. Медико-статистические технологии

Перейдем к более подробному рассмотрению наиболее распространенных медико-статистических технологий, т.е. статистических технологий анализа медицинских данных.

Клинико-статистические исследования. Среди многих видов научных медицинских исследований важное место занимают клинико-статистические. Именно о них обычно идет речь в работах по доказательной медицине.

Под клинико-статистическим исследованием понимают специально организованный сбор и анализ медицинских данных о течении заболеваний у пациентов, динамике объективных и субъективных показателей их состояния, реакции на те или иные лечебные воздействия. Исследуются одна, две или более групп лиц (больных или здоровых), выводы делаются по группам в целом, а не по каждому конкретному пациенту. Цель таких исследований состоит в переносе выводов «с выборки на генеральную совокупность». Другими словами, клинико-статистическое исследование ориентировано на получение полезных рекомендаций, касающихся тех пациентов, которые попадут в поле зрения врачей после окончания исследования. Таким образом, имеется потенциальное противоречие интересов практикующего врача и научного работника, ведущего клинико-статистическое исследование. Первый заинтересован в том, чтобы оказать наилучшую возможную помощь каждому своему пациенту. Второй разрабатывает рекомендации для будущих больных, а для этого иногда, например, при использовании плацебо, вынужден лишить своих пациентов лечебного воздействия.

Интересными для теории и полезными для практики являются различные типы научных медицинских исследований. Различия между ними в настоящем разделе не принципиальны. Например, не будем выделять отдельно эпидемиологические исследования.

Сбор данных и «карта больного». По каждому отдельному пациенту подобная информация обычно содержится в его истории болезни. Поэтому при проведении клинико-статистического исследования сбор данных зачастую состоит в заполнении специальной формы («карты больного») на основе истории болезни. Такая работа сильно облегчается, если в медицинской организации ведутся автоматизированные истории болезни со стандартными процедурами заполнения. В идеале сбор данных для конкретного клинико-статистического исследования может быть заменен на выделение соответствующей подбазы данных из общей базы медицинской организации, содержащей автоматизированные истории болезни.

«Карта больного» является одним из видов анкет: слева вопросы, справа ответы. Начальные вопросы позволяют идентифицировать пациента (фамилия, имя, отчество, адрес, паспортные данные), другие нацелены на получение информации, используемой при лечении. Обычно фиксируют:

- пол, возраст (дату рождения), профессию (вид деятельности);
- данные о перенесенных заболеваниях (анамнез), вредных привычках (курение, употребление иных наркотических веществ);
- рост, вес, пульс, температура, диастолическое и систолическое артериальное давление, результаты ЭКГ, других обследований и анализов при поступлении в медицинское учреждение и в другие моменты времени;
- назначенное лечение и течение заболевания;
- даты начала заболевания, госпитализации, операции, выписки, смерти и т.д.

Внешне «карта больного» похожа на личный листок по учету кадров. Отличие в том, что она предназначена для получения выводов по группе пациентов, а не отдельно по каждому из них. Поэтому записи в каждой графе должны вестись так, чтобы на их основе можно было формулировать выводы по группе в целом. Другими словами, записи должны быть либо числами, либо кодами градаций из заранее составленных списков.

С формальной точки зрения «карта больного» напоминает социологические и маркетинговые анкеты. Только вопросы в них другие, соответствующие предметным областям. А вот методы обработки данных — те же. Поэтому читатель не должен удивляться тому, что в главе о доказательной медицине рассматриваются такие же методы анализа статистических данных, как и в публикациях по эконометрике.

Виды клинико-статистических исследований. Два основных вида — ретроспективные и проспективные. Первые связаны с анализом прошлого, вторые — будущего.

Исходный материал для ретроспективных исследований — истории болезней. Чего в них нет — восстановить нельзя. В прошлом нельзя ни опросить больных, ни провести дополнительные анализы и обследования. Это — недостаток.

Кроме того, трудно по историям болезни составить сопоставимые экспериментальную и контрольную группы, еще труднее обосновать эту сопоставимость. Если в прошлом врач выбрал для одних пациентов определенное лечебное воздействие, а для других — иной способ лечение, то, видимо, у него были для этого основания. Эти основания не всегда зафиксированы в историях болезни. Можно, конечно, подобрать экспериментальную и контрольную группы, сопоставимые по полу, возрасту, анамнезу, результатам анализов и обследований. Но, увы, нельзя быть уверенным, что в исследовании не упущен какой-то важный в прошлом фактор. Может быть, врач принимал решение на основе неформализуемой (экспертной) оценки состояния больного? Подобная оценка делается интуитивно, а потому и не может войти в такой официальный документ, как история болезни.

А преимущества ретроспективных исследований состоят в том, что, во-первых, нет необходимости ждать развития заболевания и проявления результата лечебного воздействия. Так что клинико-статистическое исследование может быть проведено быстро. При хорошо поставленной системе автоматизированных историй болезни — за один день. Во-вторых, отсутствуют этические проблемы, связанные, например, с разделением пациентов на экспериментальную и контрольную группу.

Отметим, что иногда удается изучить состояние пациентов в настоящее время. А то и спланировать будущее ведение больных. Но это уже переход к проспективному исследованию. Таким образом, кроме двух основных форм клинико-статистических исследований, есть промежуточные.

При планировании проспективного исследования есть возможность предусмотреть сбор всей необходимой информации, проведения всех нужных врачу-ученому анализов и обследований.

К недостаткам относится большая продолжительность исследования. В некоторых случаях представляется полезным знать продолжительность предстоящей жизни пациента после того или иного лечебного воздействия. Возможно, заканчивать подобное исследование придется уже следующим поколениям врачей. Впрочем, за несколько десятилетий изменятся взгляды на лечение рассматриваемых заболеваний, появятся новые виды анализов и обследований, так что отмеченное выше преимущество проспективного исследования сойдет на нет.

Преимуществом является также возможность обеспечить сопоставимость экспериментальной и контрольных групп, например, применить рандомизацию или «двойной слепой метод». С одной стороны, это дает возможность статистически достоверно (т.е. с вероятностью, близкой к 1) обосновать наличие (или отсутствие) эффекта. С другой стороны, часть пациентов будет получать плацебо, а не лекарство, т.е. будет лишено лечебного воздействия, что не всегда представляется корректным с этической точки зрения.

Формулировки вопросов. Как и в маркетинговых и социологических опросах (см. главу 1), в научных медицинских исследованиях в «картах больного (пациента)» используют три типа вопросов — закрытые, открытые и полузакрытые, они же полуоткрытые. При ответе на закрытый вопрос врач-исследователь может выбирать лишь из сформулированных составителями анкеты вариантов ответа. В качестве ответа на открытый вопрос его просят изложить свое мнение в свободной форме. Полузакрытые, они же полуоткрытые вопросы занимают промежуточное положение — кроме перечисленных в анкете вариантов, врач-исследователь может добавить свои соображения.

Преимущество открытых вопросов состоит в том, что врач-исследователь может свободно высказать свое мнение так, как сочтет нужным. Их недостаток — в сложности сопоставления мнений по поводу различных больных. Для такого сопоставления и получения сводных характеристик организаторы исследования вынуждены сами шифровать ответы на открытые вопросы, применяя разработанную ими схему шифровки.

Преимущество закрытых вопросов в том и состоит, что такую шифровку проводит сам врач-исследователь. Однако при этом организаторы опроса уподобляются древнегреческому мифическому персонажу Прокрусту. Как известно, Прокруст приглашал путников заночевать у него. Укладывал их на кровать. Если путник был маленького роста, он вытягивал его ноги так, чтобы они доставали до конца кровати. Если же путник оказывался высоким, и ноги его торчали — он обрубал их так, чтобы достигнуть стандарта: «рост» путника должен равняться длине кровати. Так и организаторы опроса, применяя закрытые вопросы, заставляют врача-исследователя «вытягивать» или «обрубать» свое мнение, чтобы выразить его с помощью приведенных в формулировке вопроса возможных ответов.

Ясно, что для обработки данных по группам и сравнения групп между собой нужны формализованные данные, и фактически речь может идти лишь о том, кто — заполняющий «карту» на конкретного больного врач-исследователь или организатор исследования при повторном просмотре всех карт — будет шифровать ответы.

Если в исследовании практически для всех вопросов варианты ответов можно перечислить заранее, то целесообразно широко использовать закрытые вопросы. В общем случае целесообразно применять в основном закрытые и полузакрытые вопросы. Как показывает опыт проведения научных медицинских исследований, такой подход обычно является правильным — лишь в небольшом числе карт оказываются вписаны свои варианты ответов. Вместе с тем демонстрируется уважение к мнению врача-исследователя, не выдвигается требование обязательного выбора из заданного множества ответов — врач-исследователь может добавить свое, но редко пользуется этой возможностью (как правило, не более чем в 5 % случаев).

Типовые ошибки при организации клинико-статистических исследований. Начинающие врачи-исследователи иногда недостаточно продумывают процедуры сбора и анализа информации.

Например, начинают бессистемно выписывать те или иные сведения, которые, по мнению, представляют интерес для клинико-статистического исследования. Накопив гору выписок, они приходят к специалисту по анализу данных и просят обработать. Тот смотрит и видит, что материал обработке не поддается. Масса пропусков — для одних пациентов информация по определенному вопросу «карты больного» есть, для других — нет. Ответы для разных пациентов не сопоставимы. Наборы анализов, обследований, лечебных воздействий — у каждого свои. В результате обработать удается примерно 1 % собранной информации. (Например, обычно удается установить численность изученной группы пациентов.) Все остальное пропадает. Большой труд потрачен зря.

Другая типичная ошибка — «текучесть» определений понятий. В начале работы начинающий врач-исследователь использует одну определенную формулировку, потом постепенно от нее отходит. Явление, хорошо известное студентам — в процессе экзамена представление о той или иной оценке постепенно меняется — экзаменатор становится снисходительнее.

Известно также, что разные научно-практические коллективы, в том числе в медицине, по-разному определяют понятия. Например, в 1950-х и 1960-х гг. психиатры отметили разительное несовпадение уровней госпитализации с диагнозом шизофрении и депрессивных расстройств в Великобритании и США. Так, уровень госпитализации с маниакально-депрессивными психозами в возрастной группе 55–64 лет в Англии и Уэльсе был в 20 раз выше, чем в США. И наоборот, шизофрения в Нью-Йорке оказалась в два раза более распространенной, чем в Лондоне. Причина различий оказалась простой — разные традиции диагностики, т.е. разное понимание понятий. Одно и то же заболевание в США может диагностироваться как шизофрения, а в Великобритании — как маниакально-депрессивный психоз [3].

Наиболее глубокая причина ошибок — несовершенство наших знаний. Хорошо известно, что обнаружить причину различия групп пациентов зачастую весьма трудно. Например, в середине XIX в. по уровню заболевания холерой различались богатые и бедные лондонцы. В Англии осенью и зимой холера не переходила в эпидемию, а в Шотландии — наоборот. Оказалось (и это было большое научное достижение), что заболеванию холерой способствует питье сырой воды, зараженной холерными вибрионами [3]. Сырую воду бедные лондонцы пили гораздо чаще, чем богатые. По традиции сырая вода широко употреблялась осенью и зимой в Шотландии, в то время в Англии в это время пили чай. Представим теперь, что в то время проводилось бы исследование, посвященное заболеваемости холерой. Были бы подобраны экспериментальная группа англичан и контрольная группа шотландцев, одинаковые во всем, кроме места жительства. И результат исследования состоял бы в том, что зимняя заболеваемость объясняется географическим положением пациентов. Аналогично было бы «доказано», что богатство приводит к снижению заболеваемости.

Планирование и организация сбора данных в научных медицинских исследованиях. Первое, с чего начинается как само научное медицинское исследование, так и рассказ о нем — это формулировка цели исследования. Например, в типовой научной работе [4] цель формулируется так:

«Изучить антигипертензивную эффективность и влияние на эндотелиальную функцию новой ретардной формы антагониста кальция дилтиазема (Антиазем® РР) у мужчин, больных мягкой и умеренной артериальной гипертонией».

Таким образом, в типовой формулировке цели научного медицинского исследования выделяются две составляющие. Во-первых, влияние какого фактора (воздействия) изучается. Во-вторых, какая совокупность пациентов рассматривается в исследовании. Первая составляющая редко упускается из внимания. Необходимость второй не столь очевидна.

Как используются результаты клинико-статистического и, более общо, научного медицинского исследования? В большинстве случаев выводы, полученные по сравнительно небольшому числу пациентов, распространяют на всех пациентов с подобными заболеваниями. В терминологии доказательной медицины — выводы с выборки переносят на генеральную совокупность. При таком взгляде становится ясной важность тщательного описания генеральной совокупности. Одно дело — пациенты поликлиники (начальные стадии заболевания, легкое течение, как следствие — амбулаторное лечение), другое — пациенты больницы.

Напомним ситуацию, описанную в начале главы. Сельский врач был уверен в его эффективности определенного антибиотика, поскольку практически все больные с пневмонией, поступавшие в ту больницу, где он работал, быстро и хорошо этим антибиотиком вылечивались. В то же время врач областной больницы вполне обоснованно считал, что этот антибиотик бесполезен: он не смог вылечить им ни одного из больных с пневмонией, кто поступил в его отделение, и был вынужден давать этим больным другие лекарства.

В чем причина такого расхождения мнений? Дело в том, что эти два врача лечили больных из двух разных генеральных совокупностей. Сельский врач основывал свои выводы на анализе лечения больных из генеральной совокупности жителей своего участка. В областную больницу привозили тех и только тех больных, которым не смогли помочь сельские врачи, т.е. больных, для которых обсуждаемый антибиотик оказался неэффективным. Это — совсем иная генеральная совокупность. Доля таких пациентов в сельских больницах была невелика, но все они, в конце концов, собирались в областной больнице — потому ее врач так критически об этом антибиотике и высказывался. И ведь оба были правы — на основании собственного опыта и опыта своих больниц. Поэтому, как видно, основываться лишь на опыте лечения конкретных пациентов не всегда полезно... Необходимо применять методы доказательной медицины.

Материалы и методы. Но вот цель исследования сформулирована. Что дальше? В стандартном содержании научной медицинской статьи следом за целью описывают «материалы и методы». Материалы, по крайней мере в клинико-статистических исследованиях, — это описание групп пациентов, прежде всего по полу, возрасту, анамнезу. Под методами понимают способы изучения течения заболевания. Например, в резюме статьи [5] сформулирована цель:

«Исследовать эффективность коррекции нарушений обменных процессов и тромбоцитарных дисфункций у больных артериальной гипертонией (АГ) с метаболическим синдромом (МС) сочетанным применением метформина и немедикаментозных средств».

Затем идет абзац:

«*Материалы и методы.* 25 больным АГ с МС назначался метформин, другим 25 — комплекс немедикаментозного лечения и метформин. Оценивалась динамика антропометрических показателей, липидного спектра крови, перекисного окисления липидов плазмы и тромбоцитов, их агрегация и внутрисосудистая активность».

Как уже отмечалось, при планировании исследования обычно выделяют экспериментальную и контрольную группы, одинаковые или близкие по всем показателям, кроме изучаемого фактора (воздействия). Более точны следующие формулировки, взятые из монографии по доказательной медицине [1].

«*Экспериментальная группа* — это группа, подвергающаяся вмешательству (лечению) в ходе исследования. Иначе называется *группой лечения*, или *группой вмешательства*.

Контрольная группа, или группа сравнения — группа испытуемых, получающих обычное лечение, или не получающих лечения, или получающих плацебо. Результаты измерений в контрольной группе сравниваются с результатами измерений в экспериментальной группе для оценки эффекта исследуемого метода лечения».

Все научные исследования, в том числе и медицинские, и особенно клинико-статистические, проводятся по одним и тем же правилам, выработанным многими поколениями исследователей. Опишем базовый вариант экспериментальной научной работы, следуя пособию проф. Д.А. Новикова [6]. Хотя эта книга, судя по названию, предназначена для начинающих исследователей в области педагогических наук, сформулированная в ней логика исследования и методы обработки данных «один к одному» пригодны для применения в медицине (см. также [7]).

Базовая схема эксперимента. Целью эксперимента обычно является эмпирическое подтверждение или опровержение гипотезы исследования и/или справедливости теоретических результатов.

Рассмотрим следующую *базовую модель эксперимента*. Пусть имеется некоторый медицинский *объект*, изменение *состояния* которого исследуется в ходе эксперимента. В качестве объекта может выступать отдельный индивид, группа, коллектив и т.д., например, множество пациентов, лечение которых проводится по новой методике. Состояние объекта измеряется[8] теми или иными показателями[9] (*характеристиками*) по *критериям*[10], отражающим его существенные характеристики. Примерами критериев являются исход заболевания, продолжительность лечения, число осложнений и т.д.

Эксперимент заключается в целенаправленном *воздействии* на объект, призванном изменить его определенным образом. Собственно, это воздействие — его состав, структура, свойства и т.д. — и есть результат теоретического исследования. Примерами воздействия являются применение новых лекарств, новых видов лечебного воздействия и т.д.

Следовательно, при проведении эксперимента необходимо обосновать, что состояние объекта изменилось, причем в требуемую сторону. Но этого недостаточно. Нужно обосновать, что изменения произошли именно в результате произведенного воздействия.

Действительно, на утверждение о том, что срок лечения сократился в результате использования нового вида лечебного воздействия, можно всегда возразить, — а, может быть, он сам сократился бы, без каких-либо нововведений, или в результате каких-либо других воздействий, совпавших по времени (например, повышения оплаты пребывания в лечебном заведении)? Аналогично, на утверждение о том, что объективно устанавливаемое состояние пациентов, прошедших лечение по новой методике, лучше состояния тех, кто лечился

---

[8] Измерение — «процесс определения какой-либо мерой величины чего-либо». Величина — «то (предмет, явление и т.д.), что можно измерить, исчислить». Мера — «единица измерения». Все определения здесь и далее взяты, если не оговорено особо, из словаря русского языка С.И. Ожегова. См. о современной теории измерений [8, гл. 3].

[9] Показатель — «то, по чему можно судить о развитии и ходе чего-либо».

[10] Критерий — «1) средство для вынесения суждения; стандарт для сравнения; правило для оценки; 2) мера степени близости к цели».

по традиционной методике, можно возразить следующим образом. А, может быть, состояние первых до начала применения новой методики была лучше, и, если бы новая методика не применялась, то оно в результате оказалось бы выше наблюдаемой?

Например, в России в 1990-е гг. резко снизилась обращаемость пациентов в лечебные заведения по поводу ряда заболеваний, особенно легких, не требующих госпитализации. Однако из этого не следует, что снизилась и заболеваемость, здоровье населения стало лучше. Судя по резкому возрастанию смертности в 1990-е гг., дело обстояло прямо противоположным образом — здоровье населения ухудшилось. А снижение обращаемости определялось социально-экономическими причинами — появлением безработицы, снижением уровня жизни населения, ухудшением работы учреждений здравоохранения. Все эти причины приводят к тому, что население при легких заболеваниях обращается не к врачу, а к «домашним» средствам, люди продолжают работать во время течения заболевания.

Итак, для того чтобы выделить в явном виде результат целенаправленного воздействия на исследуемый объект, необходимо взять аналогичный объект и посмотреть, что происходит с ним в отсутствии воздействий.

Как уже говорилось, традиционно эти два объекта в экспериментальных исследованиях называют соответственно *экспериментальной группой* (например, ведущейся по предложенной лечебной методике) и *контрольной группой* (например, ведущейся по традиционной методике).

На рис. 1 представлена в общем виде структура любого медицинского эксперимента. Двойными пунктирными стрелками отмечены процедуры сравнения характеристик объектов. С помощью рис. 1 обсудим все составляющие базовой модели медицинского эксперимента.

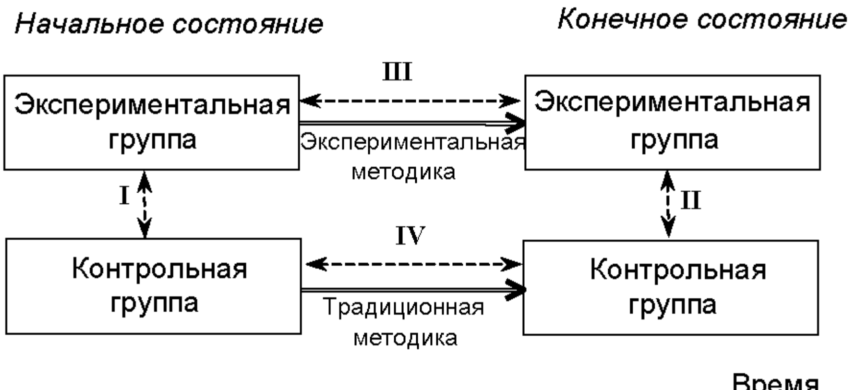


Рис. 1. Структура медицинского эксперимента

Констатации (в результате сравнения III — см. рис. 1) различий начального и конечного состояний (динамики) экспериментальной группы недостаточно — быть может, аналогичные изменения происходят и с контрольной группой, что может быть установлено сравнением IV. Поэтому **алгоритм** действий исследователя (т.е. способ решения исследовательских задач, точно предписывающий, как и в какой последовательности получить заключительный результат), заключается в следующем:

1) на основании сравнения I установить совпадение начальных состояний экспериментальной и контрольной группы;

2) реализовать воздействие на экспериментальную группу[11];

3) на основании сравнения II установить различие конечных состояний экспериментальной и контрольной группы.

Легко видеть, что, выполняя перечисленные шаги, мы, фактически, косвенным образом реализуем процедуру сравнения III, исключая влияние общих для экспериментальной и контрольной группы условий и воздействий.

Спрашивается, а где же место статистических методов доказательной медицины? Роль их заключается в том, чтобы корректно и достоверно обосновать совпадение или различие состояний контрольной и экспериментальной группы.

Многообразие схем эксперимента и статистические технологии. Эксперимент может следовать и более сложной, но укладывающейся в рамки описанной методологии, схеме — характеристики контрольных и экспериментальных групп могут измеряться и сравниваться неоднократно, в различные моменты времени.

Применяют и иные схемы клинико-статистических исследований. Например, в исследовании по схеме «до — после» нет контрольной группы. Изучается течение заболевания в одной группе пациентов, подвергающейся изучаемому вмешательству (в терминах рис. 1 — проводится сравнение III). Обычно такое исследование служит для первоначального анализа ситуации, является базой для постановки следующей работы по схеме рис. 1. Иногда же при проведении исследовании по схеме «до — после» предполагают, что любое улучшение, наблюдаемое после лечения, обусловлено именно лечением. Это предположение может оказаться ложным, что делает метод «до — после» весьма уязвимым [1].

Проведение сравнений III и IV позволяет выявить изменения в показателях пациентов экспериментальной и контрольной групп за время наблюдения. Таким образом, все четыре сравнения, указанные на рис. 1, широко использу-

[11] При выполнении данного шага необходимо быть уверенным, что и экспериментальная, и контрольная группы находятся в одинаковых условиях, за исключением целенаправленно изменяемых исследователем.

ются в научных медицинских исследованиях. Расчетные методы проведения сравнений показателей (между группами и в одной группе в разные моменты времени) рассмотрены ниже (см. также главу 5).

Отдельные методы сбора и анализа данных (как медицинских, так и данных иной природы) объединяются в статистические технологии. В настоящей книге рассматриваются лишь отдельные элементы статистических технологий, которые, тем не менее, позволяют решать подавляющую часть задач, возникающих при анализе данных в медицинских научных исследованиях. Понятие «статистические технологии» подробно обсуждается в эконометрике [8, гл. 11]. Приведенные в [8] соображения без труда переносятся в область доказательной медицины.

Выборочные научные медицинские исследования. Термин «выборочные исследования» применяют, когда невозможно изучить все единицы представляющей интерес совокупности. Приходится знакомиться с частью совокупности — с выборкой, а затем с помощью статистических методов и моделей переносить выводы с выборки на всю совокупность (см. главу 1).

Научные медицинские исследования можно разделить на три типа. К первому относятся немногочисленные работы, посвященные описанию отдельных интересных случаев. Ко второму — выборочные исследования, когда по результатам изучения сравнительно небольших групп больных (десятки или сотни человек) делаются попытки сформулировать выводы и рекомендации для гораздо более широких совокупностей. Третий тип — это обобщающие работы, построенные на основе ранее проведенных исследований. Как правило, именно они наиболее полезны для читателей, прежде всего для практикующих врачей. Однако без работ первых двух типов, содержащих всю первичную информацию, такие работы невозможны.

Методы выборочных исследований основаны на вероятностных моделях, описывающих получение ответов на вопросы, включенные в «карты больного». В случае ответов типа «есть» — «нет» наиболее распространенными являются две вероятностные модели — биномиальная и гипергеометрическая, подробно разобранные в главе 1. Там же установлено, что биномиальная и гипергеометрическая модели *весьма близки* (с практической точки зрения совпадают), когда объем генеральной совокупности по крайней мере в 10 раз превышает объем выборки. Отличия проявляются лишь при обсуждении вопроса о том, какую выборку считать *представительной* (в идущей с Запала терминологии — *репрезентативной*). Например, является ли таковой выборка, составленная из больных, находящихся в определенном отделении медицинского учреждения на фиксированный момент времени? В биномиальной модели вполне допустим ответ «да», в гипергеометрической — только «нет».

Биномиальная модель легче для теоретического изучения, поэтому будем ее рассматривать в дальнейшем. Однако при проведении реального научного медицинского исследования лучше формировать выборку, исходя из гипергеометрической модели, если такая возможность имеется. Например, при проведении эпидемиологического исследования с целью изучения качества жизни [9] это делают, выбирая опрашиваемых (участников исследования) из списка избирателей с помощью датчиков псевдослучайных чисел на компьютере или с помощью таблиц псевдослучайных чисел. Алгоритмы формирования выборки встраивают во все современные программные продукты, предназначенные для поддержки проведения подобных исследований.

Сколько больных достаточно для исследования? Другими словами, каков должен быть объем выборки? (По-английски — sample size, поэтому этот термин иногда неправильно переводят как «размер выборки».) Ответ на этот вопрос очевиден — чем больше, тем лучше. Обычно объем выборки определяется финансовыми возможностями организаторов научного медицинского исследования, числом больных с соответствующими нозологическими формами, находящимися в определенном медицинском учреждении во время исследования, и т.д. Лишь в отдельных случаях, например, когда требуется оценить определенный параметр с заданной точностью и выбрать между двумя полностью описанными гипотезами, необходимый объем выборки можно рассчитать на основе теории математической статистики.

Формирование и динамика выборки. Как реально подбирается выборка? В [1, с. 177] описывается конкретное научное медицинское исследование. В качестве популяции, из которой производится отбор в выборку, рассматриваются больные с инсулиннезависимым сахарным диабетом, наблюдающиеся в одной клинике. Предполагается, что клиника является типовой, и результаты могут быть перенесены на всех больных с рассматриваемой нозологической формой заболевания.

Критерии отбора (включения в выборку) таковы:

1. Возраст старше 40 лет.
2. Сахарный диабет диагностирован в возрасте старше 30 лет.
3. Требуется лекарственная коррекция гипергликемии.
4. Планируемая продолжительность наблюдения — более 2 лет.
5. Выполнение определенных ограничений (здесь не описываются), связанных с другими заболеваниями, инвалидностью и т.п.

Однако из тех больных, кто соответствует критериям включения в выборку, некоторая часть сразу откажется сотрудничать с исследователем. Некоторые не захотят лечиться предлагаемым способом. Другие станут возражать против того, чтобы выбор лечения определялся случайным образом (при применении двойного слепого методы — см. ниже), а не их лечащим врачом. Пациенты-отказники обычно отличаются от давших согласие по социально-экономическому статусу, тяжести заболевания и другим признакам.

За время исследования (в данном случае — за 2 года) состав выборки изменится. Отпадет ряд пациентов, которые не соблюдали график посещения врача-исследователя, полностью или частично не проходили регулярных обследований. Останутся лишь наиболее дисциплинированные и ответственные пациенты. Но и среди них будут выбывшие из исследования — из-за смерти, переезда, развития другого заболевания и т.п.

Таким образом, реальная выборка будет отличаться от намеченной. Выборка будет нуждаться в «ремонте». Если критериям отбора удовлетворяет достаточно много пациентов, то простейшее решение состоит в увеличении исходного объема выборки, чтобы к концу исследования обеспечить прослеживание динамики заболевания для заданного числа больных. Некоторые иные методы «ремонта» выборки рассматриваются ниже. Однако приходится, например, констатировать, что включенные в научное медицинское исследование пациенты являются (в среднем) более дисциплинированными и ответственными, чем больные в исходной популяции. Поэтому переносить полученные в научных медицинских исследованиях выводы и результаты в обычную клиническую практику следует с осторожностью.

Доверительное оценивание доли (вероятности). Зачем проводятся выборочные исследования? Чтобы получить необходимую информацию о генеральной совокупности. Для этого необходимо перенести выводы с выборки на генеральную совокупность. Как и с какой точностью можно это сделать?

Рассмотрим эту проблему для простейшего случая одного вопроса в «карте больного» с двумя возможными ответами — «есть» и «нет». Конечно, пары подсказок могут быть иными. Например, на вопрос об отношении к некоторому высказыванию пациент может отвечать «согласен» или «не согласен». Ситуация описывается одним параметром p — долей пациентов, выбирающих первую подсказку, во всей генеральной совокупности. Необходимо оценить этот параметр.

В главе 1 развита соответствующая теория. Предварительно требуется задать характеристику надежности переноса выводов с выборки на генеральную совокупность — доверительную вероятность γ, близкую к 1. В главе 1 построим доверительный интервал, «накрывающий» неизвестную нам величину p с вероятностью γ, близкой к 1. Чем больше вероятность γ, тем надежнее вывод. Но при этом вывод становится менее точным, длина интервала увеличивается. Надежность и точность меняются в различных направлениях при изменении доверительной вероятности.

Пусть функция $U(\gamma)$ удовлетворяет условию

$$\Phi(U(\gamma)) - \Phi(-U(\gamma)) = \gamma,$$

т.е.

$$U(\gamma) = \Phi^{-1}\left(\frac{1+\gamma}{2}\right),$$

где $\Phi(x)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. В главе 1 показано, что нижнюю доверительную границу целесообразно задать в виде

$$p_{нижн} = p^* - U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}},$$

в то время как верхняя доверительная граница такова:

$$p_{верх} = p^* + U(\gamma)\frac{\sqrt{p^*(1-p^*)}}{\sqrt{n}},$$

где p^* — выборочная доля, т.е. доля пациентов в выборке, выбравших первую подсказку.

Наиболее распространенным (в прикладных исследованиях) значением доверительной вероятности является $\gamma = 0{,}95$. Иногда употребляют термин «95 % доверительный интервал». Тогда $U(\gamma) = 1{,}96$.

Пример 1. Пусть в выборке обьема n=500 выбрали первую подсказку m=200 пациентов. Тогда p^*=0,40. Найдем доверительный интервал для $\gamma = 0{,}95$:

$$p_{нижн} = 0{,}40 - 1{,}96\frac{\sqrt{0{,}4\times 0{,}6}}{\sqrt{500}} = 0{,}40 - 0{,}043 = 0{,}357, \quad p_{верх} = 0{,}40 + 0{,}043 = 0{,}443.$$

Таким образом, хотя в достаточно большой выборке 40 % пациентов говорят «да», можно утверждать лишь, что во всей генеральной совокупности таких от 35,7 % до 44,3 % — крайние значения отличаются на 8,6 %.

Замечание. С достаточной для практики точностью можно заменить 1,96 на 2.

Удобные для использования в практической работе таблицы точности оценивания разработаны во ВЦИОМ (Всероссийском центре по изучению общественного мнения). Одна из них приведена ранее в главе 1.

Пример 2 [3, с.159]. Из 64 пациентов диагноз шизофрении был у 30. Что можно сказать о распространенности шизофрении среди всех психических заболеваний?

В рассматриваемом случае $p^* = 30/64 = 0{,}47$. Найдем доверительный интервал для $\gamma = 0{,}95$:

$$p_{нижн} = 0{,}469 - 1{,}96\frac{\sqrt{0{,}469 \times 0{,}531}}{\sqrt{64}} = 0{,}469 - 0{,}122 = 0{,}347,$$
$$p_{верх} = 0{,}469 + 0{,}122 = 0{,}591.$$

Таким образом, среди всех психических заболеваний шизофрения занимает от 34,7 до 59,1 %, что можно округлить до интервала от 35 до 60 %. Более точная оценка может быть найдена лишь на основе гораздо более обширной выборки.

Проверка однородности вероятностей — одна из базовых проблем, решаемых статистическими методами. Она часто обсуждается в литературе, а методы проверки однородности применяются при решении многих практических задач. Например, как сравнить две группы — мужчин и женщин, молодых и пожилых, и т.п.?

Если две группы не отличаются по ответам, значит, их можно объединить и осуществлять одни и те же лечебные воздействия. Если же две группы различаются, то и относиться к ним надо по-разному.

Обсуждаемая далее постановка задачи в вероятностно-статистических терминах такова. Рассматривается вопрос «карты больного» с двумя возможными ответами, например, «есть» и «нет». В первой группе из n_1 пациентов для m_1 человек ответ «есть», а во второй группе из n_2 больных для m_2 лиц ответ «есть». В вероятностной модели предполагается, что m_1 и m_2 — биномиальные случайные величины $B(n_1, p_1)$ и $B(n_2, p_2)$ соответственно.

Однородность двух групп означает, что соответствующие им вероятности равны, неоднородность — что эти вероятности отличаются. В терминах прикладной математической статистики задача ставится так: необходимо проверить гипотезу однородности

$$H_0: p_1 = p_2$$

при альтернативной гипотезе о наличии эффекта

$$H_1: p_1 \neq p_2$$

(Иногда представляют интерес и иные альтернативные гипотезы, прежде всего односторонние альтернативные гипотезы $H_1' : p_1 > p_2$ и $H_1'' : p_1 < p_2$.)

Оценкой вероятности p_1 является частота $p_1{}^*=m_1/n_1$, а оценкой вероятности p_2 является частота $p_2{}^*=m_2/n_2$. Даже при совпадении вероятностей p_1 и p_2 частоты, как правило, различаются. Как говорят, «по чисто случайным причинам».

Как показано в главе 1, правило принятия решения при проверке однородности двух выборок выглядит так:

1. Вычислить статистику

$$Q = \frac{p_1^* - p_2^*}{\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}}.$$

2. Сравнить значение модуля статистики $|Q|$ с граничным значением K. Если $|Q| \leq K$, то принять гипотезу однородности H_0. Если же $|Q| > K$, то заявить об отсутствии однородности и принять альтернативную гипотезу H_1 о наличии эффекта.

Граничное значение K определяется выбором уровня значимости статистического критерия проверки однородности. Граничное значение в зависимости от уровня значимости целесообразно выбирать из условия

$$K = K(\alpha) = \Phi^{-1}\left(\frac{1+\alpha}{2}\right).$$

Здесь $\Phi^{-1}(\cdot)$ — функция, обратная к функции стандартного нормального распределения. В клинико-статистических и медико-биологических наиболее распространен 5 % уровень значимости, т.е. $\alpha = 0{,}05$. Для него K = 1,96. Для других уровней значимости надо использовать другие коэффициенты $K(\alpha)$. Так,

$K(0{,}01) = 2{,}58$ для уровня значимости 1 % и $K(0{,}10) = 1{,}64$ для уровня значимости 10 %.

Пример 3. Пусть в первой группе из 500 мужчин оценка характеристики «есть» зафиксировано у 200, а во второй группе из 700 женщин ответ «есть» у 350. Есть ли разница по доле оценок «есть» между генеральными совокупностями, представленными этими двумя группами?

Для установления лучшего взаимопонимания с врачом-исследователем уберем из формулировки примера относящийся к теории статистики термин «генеральная совокупность». Получим следующую постановку.

Пусть из 500 мужчин определенное нарушение выявлено у 200, а из 700 женщин это нарушение зафиксировано у 350. Есть ли разница между мужчинами и женщинами по доле лиц, имеющих рассматриваемое нарушение?

В рассматриваемом примере нужные для расчетов величины таковы: $n_1 = 500, p_1^* = 200/500 = 0{,}4; n_2 = 700, p_2^* = 350/700 = 0{,}5.$ Вычислим статистику

$$Q = \frac{0{,}4 - 0{,}5}{\sqrt{\frac{0{,}4 \cdot 0{,}6}{500} + \frac{0{,}5 \cdot 05}{700}}} = \frac{-0{,}1}{\sqrt{\frac{0{,}24}{500} + \frac{0{,}25}{700}}} = \frac{-0{,}1}{\sqrt{0{,}00048 + 0{,}0003571}}$$

$$= \frac{-0{,}1}{\sqrt{0{,}0008371}} = \frac{-0{,}1}{0{,}029} = -3{,}45.$$

Поскольку $|Q| = 3{,}45 > 1{,}96$, то необходимо отклонить нулевую гипотезу и принять альтернативную. Таким образом, мужчины и женщины отличаются по рассматриваемому признаку — наличию рассматриваемого нарушения.

Необходимо отметить, что результат проверки гипотезы однородности зависит не только от частот, но и от объемов выборок. Предположим, что частоты (доли) зафиксированы, а объемы выборок растут. Тогда числитель статистики Q не меняется, а знаменатель уменьшается, значит, вся дробь возрастает. Поскольку знаменатель стремится к 0, то дробь возрастает до бесконечности и рано или поздно превзойдет любую границу. Есть только одно исключение — когда в числителе стоит 0. Следовательно, при строгом подходе к формулировкам вывод статистика должен выглядеть так: «различие обнаружено» (если принята альтернативная гипотеза) или «различие не обнаружено» (если нет оснований отклонять нулевую гипотезу). Во втором случае различие, возможно, было бы обнаружено при увеличении объемов выборок.

Как и для доверительного оценивания вероятности, во ВЦИОМ разработаны две полезные таблицы, позволяющие оценить вызванные чисто случайными причинами допустимые расхождения между частотами в группах (см. главу 1).

Итак, если частоты (выборочные доли) в группах достаточно сильно различаются, то делается вывод о различии долей в генеральных совокупностях. А если сравнительно мало различаются, то утверждают, что различие не обнаружено. А где граница? Допустимое расхождение $\Delta = \Delta(\alpha)$ между частотами нетрудно получить расчетным путем. Для этого достаточно воспользоваться формулой для статистики Q и определить, при каком максимальном расхождении частот все еще делается вывод о том, что верна гипотеза однородности. Следовательно, допустимое расхождение $\Delta = \Delta(\alpha)$ находится из уравнения

$$K(\alpha) = \frac{\Delta(\alpha)}{\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}}.$$

Таким образом,

$$\Delta(\alpha) = K(\alpha)\sqrt{\frac{p_1^*(1-p_1^*)}{n_1} + \frac{p_2^*(1-p_2^*)}{n_2}}.$$

Для данных примера 3 $\Delta = \Delta(\alpha) = 1{,}96 \times 0{,}029 = 0{,}057$, или 5,7 %, для уровня значимости 0,05.

В связи со сказанным возникает вопрос: каково типовое отличие частот в двух выборках из одной и той же совокупности? Соответствующая теория развита в главе 1. Она позволяет разработать метод контроля правильности (корректности) проведения повторных выборочных исследований. Если частоты *излишне устойчивы*, значения долей при повторных выборочных исследованиях слишком близки — это подозрительно! Возможно, нарушены правила проведения выборочных исследований, выборки не являются случайными, ответы фальсифицированы и т.д.

Пример 4. В работе [10] фраксипарин получали 75 пациентов (среди них 53 мужчины, 70,7 %), фраксифорте — тоже 75 пациентов (45 мужчин, 60 %). Различаются ли подгруппы по половому составу?

Нужные для расчетов величины таковы: $n_1 = 75$, $p_1^* = 53/75 = 0{,}707$; $n_2 = 75$, $p_2^* = 45/75 = 0{,}600$. Вычислим статистику критерия проверки однородности (обнаружения эффекта)

$$Q = \frac{0{,}707 - 0{,}600}{\sqrt{\frac{0{,}707 \cdot 0{,}293}{75} + \frac{0{,}6 \cdot 0{,}4}{75}}} = \frac{0{,}107}{\sqrt{\frac{0{,}207}{75} + \frac{0{,}24}{75}}} = \frac{0{,}107}{\sqrt{0{,}00276 + 0{,}0032}}$$
$$= \frac{0{,}107}{\sqrt{0{,}00596}} = \frac{0{,}107}{0{,}0772} = 1{,}386.$$

Поскольку $|Q| = 1{,}386 < 1{,}96$, то необходимо принять нулевую гипотезу и отклонить альтернативную. Различие не обнаружено. Таким образом, подгруппы не отличаются по половому составу.

Случай малых вероятностей. Рассмотренные выше методы анализа медицинских данных основаны на предположении о возможности использования биномиальной модели выборки. Кроме этого, необходимо, чтобы объем выборки n был достаточно большим — несколько десятков или сотен.

Наконец, третье условие. Необходимо, чтобы число m ответов «есть» (в схеме с двумя ответами «есть» — «нет») также было достаточно большим, сравнимым по величине с объемом выборки. Последнее эквивалентно предположению, что np — произведение числа опытов на вероятность рассматриваемого ответа при отдельном опыте — также достаточно велико. Напомним, что np — это среднее значение (математическое ожидание) для числа ответов m, т.е. $M(m) = np$. Другими словами, третье условие означает, что вероятность p не является слишком малой (не меньше 5–10 %). Не должна быть слишком малой и дополнительная вероятность $1 - p$. Следовательно, вероятность ответа «есть» должна лежать между 5 % и 95 % (а лучше — между 10 и 90 %).

Однако последнее условие выполнено отнюдь не всегда. Например, в области с миллионным населением наблюдаются сотни или тысячи заболеваний. Хотя их число достаточно велико, но оно составляет лишь доли процента от численности населения. Для описания подобных ситуаций используют не биномиальное распределение, а распределение Пуассона.

Случайная величина Y имеет распределение Пуассона, если

$$P(Y = y) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0,1,2,\ldots,$$

где λ — параметр распределения Пуассона (интенсивность заболеваемости или иной характеристики), e — основание натуральных логарифмов, e =

= 2,718281828..., и $P(Y=y)=0$ для всех прочих чисел y. Здесь $y!$ — произведение всех натуральных чисел от 1 до y, например, $2! = 1 \times 2 = 2$, $3! = 1 \times 2 \times 3 = 6$, $4! = 1 \times 2 \times 3 \times 4 = 24$ и т.д., при $y = 0$ обозначено $0! = 1$.

Для распределения Пуассона математическое ожидание и дисперсия совпадают с параметром интенсивности:

$$M(Y) = \lambda, D(Y) = \lambda. \tag{1}$$

Это распределение названо в честь французского математика С.Д. Пуассона (1781–1840), впервые получившего его в 1837 г. Распределение Пуассона является предельным случаем биномиального распределения, когда вероятность p осуществления события мала, но число испытаний n велико, причем $np = \lambda$. Точнее, справедливо предельное соотношение

$$\lim_{n\to\infty, np\to\lambda} P(Y = y \mid p, n) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0,1,2,...$$

Поэтому распределение Пуассона (в старой терминологии «закон распределения») часто называют также «законом редких событий».

Оценкой неизвестного параметра λ распределения Пуассона является наблюдаемое значение случайной величины Y. Доверительное оценивание может быть основано на асимптотической нормальности распределения Пуассона и соотношении (1). А именно, для доверительной вероятности 0,95 нижняя доверительная граница имеет вид

$$\lambda_{\text{нижн}} = Y - 1{,}96\sqrt{Y},$$

а верхняя доверительная граница такова:

$$\lambda_{\text{верх}} = Y + 1{,}96\sqrt{Y}.$$

Эти границы — приближенные (асимптотические), поскольку получены в предположении, что параметр интенсивности λ достаточно велик (несколько десятков или сотен). Более точные расчеты можно провести по правилам, приведенным в специальной литературе [11].

Пример 5. Пусть в муниципальном образовании численностью 125 тыс. человек наблюдается 438 случаев заболевания. Необходимо оценить заболеваемость (на 10 000 жителей).

Заболеваемость составляет

$$\frac{438}{125000}\times 10000 = 35{,}04$$

случаев на 10 000 жителей. Чтобы рассчитать доверительные границы, найдем сначала границы для общей заболеваемости (для всего муниципального образования):

$$\lambda_{нижн} = 438 - 1{,}96\sqrt{438} = 438 - 41 = 397,\ \lambda_{верх} = 438 + 41 = 479.$$

Нижняя доверительная граница для заболеваемости на 10 000 жителей составляет

$$\frac{397}{125000}\times 10000 = 31{,}76\ ,$$

а верхняя доверительная граница для заболеваемости на 10 000 жителей такова:

$$\frac{479}{125000}\times 10000 = 38{,}32\ .$$

Есть ли различия между двумя муниципальными образованиями, мало различающимися по численности, или между заболеваемостью на одной и той же территории, но в соседние годы? При моделировании распределениями Пуассона следует рассмотреть случайные величины Y_1 и Y_2, имеющие распределения Пуассона с параметрами λ_1 и λ_2 соответственно, и проверять статистическую гипотезу о равенстве параметров этих распределений, т.е. гипотезу однородности

$$H_0\colon \lambda_1 = \lambda_2$$

при альтернативной гипотезе о наличии эффекта

$$H_1\colon \lambda_1 \neq \lambda_2$$

(Иногда представляют интерес и иные альтернативные гипотезы, прежде всего односторонние альтернативные гипотезы $H_1'\colon \lambda_1 > \lambda_2$ и $H_1''\colon \lambda_1 < \lambda_2$.)

В прикладной математической статистике разработаны правила проверки гипотезы однородности для распределений Пуассона [8, п. 6.6]. В большинстве практических ситуаций, в которых наблюдаемые значения случайные величины Y_1 и Y_2 не являются малыми (другими словами, не менее нескольких десятков), достаточно воспользоваться статистическим критерием, основанным на асимптотической нормальности случайных величин, имеющих распределения Пуассона. В этом случае правило принятия решения при проверке однородности двух выборок выглядит так:

1. Вычислить статистику

$$R = \frac{|Y_1 - Y_2|}{\sqrt{Y_1 + Y_2}}.$$

2. Сравнить значение статистики R с граничным значением K. Если $R \le K$, то принять гипотезу однородности H_0. Если же $R > K$, то заявить об отсутствии однородности и принять альтернативную гипотезу H_1 о наличии эффекта.

Граничное значение K определяется выбором уровня значимости статистического критерия проверки однородности. В клинико-статистических и медико-биологических наиболее распространен 5 % уровень значимости, т.е. $\alpha = 0{,}05$. Для него $K = 1{,}96$.

Примечание. Если численности сравниваемых популяций сильно (в разы) отличаются, то необходимо применять несколько более сложные способы расчета, описанные в статистической литературе [11, п. 6.6]. При этом исходная гипотеза однородности состоит в равенстве заболеваемости на 10 000 жителей. В соответствующей статистической гипотезе утверждается, что отношение параметров двух распределений Пуассона равно определенной величине, известной исследователю (а именно, отношению численности популяций).

Пример 6. Пусть в некотором году на территории Ивановской области было 235 случаев заболевания, а в следующем году — 312 случаев.

Примем, что за год численность населения Ивановской области изменилась незначительно. Поскольку эта численность измеряется миллионами, то число заболеваний в определенном году моделируется распределением Пуассона. Рассчитаем значение статистики критерия проверки однородности двух распределений Пуассона:

$$R = \frac{|235 - 312|}{\sqrt{235 + 312}} = \frac{77}{23{,}39} = 3{,}29.$$

Поскольку полученное значение больше 1,96, то гипотезу однородности отклоняем на уровне значимости 0,05 (а также и на уровне значимости 0,01, поскольку $R > 2{,}58$). Заболеваемость значимо увеличилась.

Пример 7. Пусть выполнены все предположения примера 6, кроме одного — заболеваемость в 2003 г. составила 270 случаев. Изменятся ли выводы статистического анализа?

Рассчитаем значение статистики критерия

$$R = \frac{|235 - 270|}{\sqrt{235 + 270}} = \frac{35}{22{,}47} = 1{,}56\,.$$

Поскольку полученное значение меньше 1,96, то гипотезу однородности принимаем на уровне значимости 0,05. Различие между годами можно объяснить случайными причинами.

Расскажем об одном эпидемиологическом исследовании[12], в котором использовались методы, описанные в настоящем разделе.

Пример 8. Описторхоз является тяжелым глистным заболеванием, склонным к хроническому течению, вызывающим нарушения различных функций организма. Один из путей заражения — употребление в пищу сырой рыбы, например, в замороженном (строганина) или вяленом виде, поэтому описторхоз распространен в северных районах, но встречается практически во всех субъектах РФ. Он приводит к серьезным осложнениям вплоть до развития первичного рака печени. Например, в Тюменской области, где широко распространена описторхозная инвазия, количество заболеваний первичным раком печени в 7 раз превышает средний показатель по стране. За год в России регистрируется около 100 тыс. лиц, заболевших описторхозом.

Анализировались статистические данные о заболеваемости описторхозом по 61 административной территории. Одна из основных проблем — отличить территорию с наличием местных очагов заражения от территорий, где имеются лишь «завозные» случаи описторхоза. Другая проблема — прогноз (с помощью доверительных границ) числа заболеваний на будущее. Третья — типология территорий по степени зараженности описторхозом. Для решения поставленных задач использовались методы доказательной медицины, основанные на доверительном оценивании и проверке гипотез относительно параметров пуассоновских распределений. Число заболеваний на конкретной административной территории моделировалось распределением Пуассона. Изменение числа забо-

[12] Исследование выполнено нами совместно с сотрудниками Института медицинской паразитологии и тропической медицины им. Е.И. Марциновского Ю.А. Бочковым, Н.И. Войскунской и Г.П. Николаевским.

леваний с течением времени может быть связано с появлением очага заражения на рассматриваемой территории, но может объясняться и случайными отклонениями. Первый случай соответствует изменению параметра распределения Пуассона, а второй — его постоянству. Поэтому обосновать вывод о появлении очага заражения можно, отклонив статистическую гипотезу о равенстве параметров распределений Пуассона, описывающих число заболевших в соседние годы.

Полученные результаты представляли достаточно большой интерес для паразитологии. В частности, был сделан прогноз о наличии местной заболеваемости в Воронежской и Иркутской областях, после чего были обнаружены очаги заболевания в этих областях.

12.4. Высокие статистические технологии в научных медицинских исследованиях

Слепой метод. Как уже не раз говорилось, базовая схема проведения научного медицинского эксперимента основана на сравнениях экспериментальной и контрольной групп до и после воздействия. Могут сравниваться новая методика лечения, применяемая в экспериментальной группе, с традиционной методикой в контрольной группе. В ряде случаев, например, при изучении эффективности нового препарата, пациенты в контрольной группе получают плацебо.

Плацебо (placebo) — это лекарственная форма, которую пациент не может отличить от исследуемого препарата (по внешнему виду, цвету, вкусу, запаху и т.п.). Однако плацебо не оказывает специфического действия. Например, в качестве плацебо используют таблетки глюкозы или инъекции изотонического раствора хлорида натрия (физиологического раствора). В качестве плацебо может использоваться иное безразличное вмешательство, «воздействие без действия».

Часто наблюдается «плацебо-эффект» — изменение состояния больного (отмечаемое самим больным или лечащим врачом), связанное с самим фактом лечения, а не с биологическим действием препарата. Например, плацебо, которое больной воспринимает как лекарство, уменьшает послеоперационную боль, тошноту и зуд более чем у 30 % прооперированных больных [1]. Плацебо-эффект демонстрирует возможности воздействия психической сферы больного на течение заболевания и процесс лечения. Подобные эффекты наблюдаются не только в медицине, но и в других сферах деятельности, например, в менеджменте при управлении персоналом [12]. Сам факт внимания к работницам со стороны исследователей может повысить производительность труда, независимо от того, обоснованы или нет предлагаемые исследователями изменения в организации производства.

Таким образом, плацебо используется в научных медицинских исследованиях для имитации лечения с целью устранения систематической ошибки, порождаемой плацебо-эффектом.

Основная идея слепого метода. Итак, если больные, включенные в клинико-статистическое исследование, знают, кто из них какой вид лечения получает, их поведение может измениться и породить систематическую ошибку в итоговых научных выводах. Чтобы избежать этого эффекта, используют *слепой метод* (blinding). Более правильно было бы назвать его «методом масок» (masking), но термин «слепой метод» закрепился в массовом сознании медицинских работников.

Выделяют пять условий, соблюдение которых позволяет во многом избавиться от нежелательных эффектов [1].

1. Те, кто формирует выборку и распределяет пациентов по группам вмешательства (в базовом случае — направляет их в две группы, она из которых будет экспериментальной, а вторая — контрольной), не должны знать, какое лечение будет назначено каждому последующему больному, чтобы это не нарушало включение пациентов в исследование в порядке их поступления.

2. Пациенты не должны знать, какое именно лечение они получают. Тогда менее вероятны нарушения схемы лечения, субъективность при описании своего состояния.

3. Проводящие лечение врачи не должны знать, какое именно лечение назначено конкретному пациенту. Тогда не возникнут вызванные этим знанием различия в ведении больных.

4. Врачи-исследователи, оценивающие исходы лечения, не должны знать, какое лечение было назначено тому или иному пациенту.

5. Специалисты по доказательной медицине, с помощью статистических методов сравнивающие исходы лечения в группах, не должны знать, какая из них экспериментальная, а какая — контрольная.

Три основных слепых метода. Когда информация о реально используемом лечебном воздействии отсутствует только у пациента, говорят о простом слепом методе. Если также и у лечащих врачей и исследователей, оценивающих исходы, — о двойном слепом методе. Если рассматриваемой информации нет и у лиц, проводящих статистическую обработку результатов исследования, говорят о тройном слепом методе. Наиболее часто используется простой слепой метод, а в хорошо организованных исследованиях — двойной слепой метод.

Почему необходимо скрывать от пациента информацию о реально используемом лечебном воздействии, ясно из проведенного выше обсуждения плацебо-эффекта. Врачи, имеющие эту информацию, могут неосознанно или осознанно использовать ее при ведении больных и оценке результатов лечения. Специалисты по статистическим методам могут при выборе статистических процедур проверки гипотез или, например, отбраковки выбросов «подыграть» энтузиастам нового способа лечения.

Можно ли полностью реализовать слепой метод? Конечно, концепция слепого метода — это лишь идеальная схема. Реально и пациенты, и врачи достаточно часто догадываются о реально используемом воздействии. Имеющий биологическое действие препарат с течением времени проявляет свой лечебный потенциал, а также обнаруживаются и побочные эффекты. И то, и другое осознается и пациентами, и врачами. Например, в [1] описано двойное слепое рандомизированное исследование, проведенное с целью выяснить, предотвращает ли прием пропранолола повторный инфаркт миокарда. После окончания исследования, но до раскрытия кодов, пациентов и медицинский персонал попросили назвать, принимал ли пациент пропранолол или плацебо. Среди пациентов, принимавших пропранолол, правильно ответили 80 %, среди тех, кто принимал плацебо — 57 %. Примерно такой же результат показали врачи и медицинские сестры. Медицинский персонал мог исходить из частоты пульса. На чем основывали свои догадки пациенты — неизвестно.

Разработано много конкретных процедур организации научного медицинского исследования, реализующих основные идеи слепого метода. Конкретная применяемая процедура должна быть описана в документации исследования и в последующих публикациях.

Рандомизация и стандартизация выборки. Как распределять больных по группам — экспериментальной и контрольной? Ясно, что если отнести тяжелых больных в экспериментальную группу, а легких — в контрольную, то исходы заболевания будут тяжелее, скорее всего, именно в экспериментальной группе. Поэтому для эксперимента в больничных условиях, как правило, нельзя формировать контрольную группу из пациентов поликлиники. Группы должны быть похожи по всем основным демографическим и клиническим характеристикам. Этого можно добиться, применяя специальную процедуру формирования групп — рандомизацию.

Термин «рандомизация» происходит от *random* (англ.) — случайный. Следовательно, означает «внесение случайности». Точнее, рандомизация — это процедура, обеспечивающая случайное распределение больных в экспериментальную и контрольную группы. Случайным распределением достигается отсутствие различий между двумя группами и, таким образом, снижается вероятность систематической ошибки в клинических исследованиях вследствие различий групп по каким-либо признакам [1]. Рандомизация уравнивает вероятности воздействия на пациента не только тех факторов, которые, как мы предполагаем, могут влиять на итоговый результат лечебного воздействия, но и тех, о существовании которых мы даже не знаем.

Варианты метода рандомизации. Каждый пациент должен иметь одинаковые шансы попасть в группу воздействия и в группу без него. Назначение исследуемого или контрольного лечения производится при помощи процедуры, аналогичной подбрасыванию монеты, обеспечивающей равные шансы (вероятности) попасть в ту или иную группы. В настоящее время для получения случайности используют датчики псевдослучайных чисел, которые имеются в стандартном программном обеспечении современных компьютеров.

Пусть, например, надо разделить 100 пациентов на экспериментальную и контрольную группы. Можно по поводу каждого подбрасывать монетку (или генерировать случайное число с помощью ЭВМ), чтобы решить, в какую группу отнести. Если выпадает орел, то в экспериментальную группу, если цифра — то в контрольную.

Однако при этом численность экспериментальной группы будет случайной. Она имеет биномиальное распределение с параметрами $n = 100$ и $p = 0{,}5$. Стандартное отклонение (в другой терминологии — среднее квадратическое отклонение) численности экспериментальной группы от математического ожидания, равного 50, равно 5. И случайность численности групп, и возможное различие в объемах (на 10–15 человек, например, 43 в экспериментальной группе и 57 — в контрольной) — явные недостатки этого простого метода рандомизации.

Часто используют более сложные процедуры, обеспечивающие получение групп заданных численностей. Например, для этого с помощью датчика псевдослучайных чисел отбирают выборку объема 50 из генеральной совокупности 100 номеров — натуральных чисел от 1 до 100. Номера, вошедшие в выборку — это номера пациентов, включенных в экспериментальную выборку. Остальные 50 номеров соответствуют больным из контрольной выборки.

Описанный способ разделения пациентов на группы также имеет некоторые недостатки. Пусть, например, в генеральной совокупности 50 мужчин и 50 женщин. Число мужчин в экспериментальной группе является случайным и с большой вероятностью (точнее, с вероятностью 0,89) отличается от 25. Это число имеет биномиальное распределение с параметрами $n = 50$ и $p = 0,5$. Стандартное отклонение числа мужчин от математического ожидания, равного 25, составляет 3,54. Такой разброс может показаться излишне большим. Можно ли сравнивать группы, полагая их состав одинаковым, если в одной из них 38 % мужчин, а во второй — 62 %?

Поэтому применяют еще более сложные схемы рандомизации. Например, подбирают экспериментальную и контрольную группы так, чтобы они не только совпадали по численности, но и имели ровно по 25 мужчин и по 25 женщин. В теории выборочных исследований такие заранее заданные условия называют квотами. Разделить случайным образом 100 пациентов на две группы по 50 человек, соблюдая квоты, также можно с помощью датчиков псевдослучайных чисел, хотя это и сложнее, чем в предыдущем случае.

Другой недостаток квотного подхода к разделению заданного числа пациентов на две группы состоит в том, что эти группы нельзя рассматривать как независимые выборки. Например, численности мужчин в экспериментальной и контрольной группах соответственно нельзя рассматривать как независимые случайные величины, поскольку сумма этих численностей заранее известна. Следовательно, нельзя обрабатывать собранные в ходе исследования данные с помощью методов, опирающихся на модель двух независимых выборок (см. раздел 12.3 выше). Чтобы избежать этого недостатка, отбор пациентов в группы иногда проводят в два этапа. На первом формируют две совокупности кандидатов на участие в исследовании, отдельно для участия в экспериментальной группе и отдельно — в контрольной. А затем отбирают из них выборки «основного состава» групп.

Дополнительная проблема состоит в том, что за время исследования группы, как правило, сокращаются (см. раздел 12.3 выше). Тут-то могут оказаться полезными «кандидаты». Они могут быть переведены в «основной состав». Таким образом, начинать исследование целесообразно с несколько большим числом пациентов в группах, чем планируется закончить.

Стратификация и стандартизация. Однако по тем или иным причинам не всегда удается с помощью рандомизации выровнять состав групп. Возникает мысль провести «ремонт» выборок, чтобы они больше походили друг на друга. В научных медицинских исследованиях применяют метод «стандартизации выборок». Его применение не обязательно связано с формированием экспериментальной и контрольной групп. Рассмотрим основные идеи метода стандартизации выборок на примере из [1].

Пример 1. Пусть требуется сравнить послеоперационную летальность при операции аортокоронарного шунтирования в больницах *А* и *Б*. Всего в больнице *А* отмечено 48 смертей на 1 200 операций (4 %), а в больнице *Б* — 64 смерти на 2 400 операций (2,6 %). Судя по этому огрубленному показателю — проценту послеоперационной летальности по больнице в целом, — больница *Б* лучше больницы *А*.

Однако такой вывод был бы преждевременным. Вполне возможно, что совокупности пациентов двух больниц различаются по исходному состоянию, а потому и по прогнозу, а затем — и по частоте тех или иных исходов. Как учесть возможное исходное различие? Обычно проводят стратификацию — на основании возраста, функции миокарда, степени стеноза и других характеристик разделяют (стратифицируют) пациентов на группы (страты) с разным дооперационным риском (табл. 1), а затем сравнивают послеоперационную летальность в пределах каждой подгруппы.

Таблица 1

Пример стратификации: летальность в двух больницах с учетом дооперационного риска

Дооперационный риск	**Больница *А***			**Больница *Б***		
	Число больных	**Число летальных исходов**	**Летальность, %**	**Число больных**	**Число летальных исходов**	**Летальность, %**
Высокий	500	30	6	400	24	6
Средний	400	16	4	800	32	4
Низкий	300	2	0,67	1 200	8	0,67
Всего	1200	48	4	2 400	64	2,6

После того как пациенты были разделены по дооперационному риску, послеоперационная летальность оказалась одинаковой для двух больниц: 6,0, 4,0 и 0,67 % для пациентов с высоким, средним и низким риском соответственно.

Почему же общая смертность в больнице *Б* заметно меньше? Ответ очевиден — характеристики пациентов в двух больницах существенно различались. Высокий дооперационный риск имели 42 % пациентов в больнице *А* и только 17 % пациентов в *Б*. С низким риском, наоборот, пролечены 25 % больных в больнице *А* и 50 % — в больнице *Б*.

Рассмотренный пример показывает, что стратификация — действенный прием борьбы с поспешными выводами. Однако стратификация — только первый этап анализа. В примере он оказался достаточным, поскольку столбцы, описывающие летальность в зависимости от уровня риска, оказались одинаковыми для двух больниц. А как же сравнивать между собой больницы, у которых эти столбцы различны?

Основная идея состоит *в пересчете на стандартную популяцию*, короче, в *стандартизации выборки*. Распределения пациентов по степени дооперационного риска в двух больницах различны, поэтому надо перейти к некоторому единому распределению, *к стандартной популяции*.

Один из простейших способов выбора стандартной популяции — построение ее как объединения совокупностей пациентов рассматриваемых больниц. В табл. 2 приведены данные о стандартной популяции, полученной путем объединения популяций для больниц *А* и *Б* (табл. 1) и больницы *В*, данные о которой приведены в той же таблице.

Таблица 2

Переход к стандартной популяции

Дооперационный риск	Стандартная популяция			Больница *В*		
	Число больных	Число летальных исходов	Летальность, %	Число больных	Число летальных исходов	Летальность, %
Высокий	1700	94	5,5	800	40	5
Средний	2000	80	4	800	32	4
Низкий	2300	18	0,78	800	8	1
Всего	6000	192	3,2	2400	80	3,3

Метод стандартизации (перехода к стандартной популяции) состоит в расчете средневзвешенных показателей, в которых летальность по группам риска (по стратам) берется соответствующей рассматриваемому лечебному заведению, а веса групп риска — согласно стандартной популяции. Так, согласно табл. 2, вес страты с высоким дооперационным риском равен 1 700/6 000 = = 0,283, вес страты со средним дооперационным риском — 2 000/6 000 = 0,333, вес страты с низким дооперационным риском — 2 300/6 000 = 0,384. Следовательно, средневзвешенный показатель смертности для больниц *А* и *Б* равен

$$6 \times 0,283 + 4 \times 0,333 + 0,67 \times 0,384 = 1,698 + 1,332 + 0,257 = 3,287\ (\%),$$

а для больницы *В* аналогичный расчет дает

$$5 \times 0,283 + 4 \times 0,333 + 1 \times 0,384 = 1,415 + 1,332 + 0,384 = 3,131\ (\%).$$

Таким образом, по рассматриваемому показателю больница *В* имеет преимущество — для нее смертность меньше на 0,156 % по сравнению с больницами *А* и *Б*.

Проведенные расчеты дают возможность выделить вклад различных страт в средневзвешенный показатель. Так, для больниц *А* и *Б* вклад страты больных с высоким риском составляет 1,698, а для больницы *В* меньше — 1,415, поскольку смертность для этой категории больных в ней не 6 %, а 5 %. В соответствии с приведенными численными данными страта больных с высоким риском дает примерно половину смертности, хотя такие больные составляют 28 % по численности.

Стандартизацию можно рассматривать как метод «лечения» выборки, позволяющий скорректировать численности тех или иных страт выборки (например, страт по возрасту и полу) в соответствии с тем, как эти страты представлены в более широкой («стандартной») популяции (например, в населении региона).

Медицинские информационные системы. Развитие компьютерной науки и техники дало в руки врачей исследователей мощные интеллектуальные инструменты — информационные технологии. В каждом современном лечебном заведении работает медицинская информационная система (МИС). Она, конечно, используется для управления больницей или поликлиникой, но не только для этого.

Сведения о больном естественно хранить не на страницах бумажной истории болезни, а в памяти компьютера. Кроме вносимых с клавиатуры текстов, туда могут в автоматическом режиме поступать сведения о результатах анализов и исследований. Для этого необходимы соответствующие программные продукты, обеспечивающие, в частности, ведение автоматизированной истории болезни. А также компьютерная сеть, объединяющая различные подразделения лечебного заведения. Отметим пользу электронного документооборота, во многом снимающего с врача нагрузки на ведение медицинской документации. В частности, основные (по размеру) разделы выписок из историй болезни могут быть составлены автоматически.

Роль МИС в исследованиях. Компьютерные информационные технологии во многом облегчают проведение научных медицинских исследований. Выделим два этапа:

1. Сбор и систематизация исходных данных о пациентах.

2. Анализ медицинских данных с помощью различных статистических методов.

Первый этап обычно реализуется с помощью баз данных различных типов, входящих в состав медицинских информационных систем. Для конкретного научного медицинского исследования производят «выемку» соответствующей информации из баз данных общего назначения. Если, конечно, значения всех необходимых показателей уже введены. В противном случае в МИС включаются новые показатели, затем начинается сбор соответствующих данных. В технологии проведения ретроспективных научных медицинских исследований благодаря использованию МИС ликвидируется этап извлечения информации из историй болезни.

Методам анализа медицинских данных с помощью различных статистических методов посвящена следующая глава. Современные информационные технологии позволяют легко и быстро проводить многочисленные сравнения групп между собой, находить связи и зависимости между показателями, строить правила диагностики и прогноза.

Проблема множественных сравнений. Однако в легкости проведения расчетов таятся и опасности получения необоснованных выводов. Одна из наиболее известных опасностей именуется «проблемой множественных сравнений» [1, с. 247–249]. Она давно известна и в других областях знаний под названием «проблема множественных проверок статистических гипотез» [8, п. 11.1].

Чтобы дать представление о проблеме, рассмотрим пример [1]. Допустим, что было выполнено масштабное клинико-статистическое исследование с большим числом подгрупп пациентов и различными исходами заболевания. В частности, это могло быть клиническое испытание эффективности лечения ишемической болезни сердца, в которое было включено несколько разных групп больных. Выделялись группы с поражением одной, двух или трех коронарных артерий; с нарушениями ритма сердца и без них; с нарушением и без нарушения функции левого желудочка. Всего групп с различными сочетаниями этих состояний имеется, как следует из правил комбинаторики, $3 \times 2 \times 2 = 12$. Рассматривалось несколько исходов (смерть, инфаркт миокарда, стенокардия). Предположим, что эффективность лечения оценивается отдельно для каждой подгруппы и каждого исхода. Фактически это означает, что подгрупп имеется $12 \times 3 = 36$ (т.е. столько, сколько сочетаний группа × исход). И каждая подгруппа сравнивается с каждой. Таким образом, всего проводится $(36 \times 35)/2 = 630$ сравнений, т.е. 630 проверок статистических гипотез.

Для выявления сути дела предположим, что в действительности не существует связи между различными видами лечения и исходами в любой из подгрупп. На языке статистических методов это означает, что в действительности во всех описанных выше 630 задачах сравнения выполнена нулевая гипотеза. Если, как это принято в научных медицинских исследованиях, гипотезы проверяются на уровне значимости 0,05, то в среднем в 5 % случаев нулевая гипотеза отклоняется, принимается альтернативная — о наличии эффекта, хотя на самом деле эффекта нет. При 630 сравнениях в 5 % случаев, т.е. примерно в 630 × 0,05 = = 31 сравнении будет сделан ложный вывод о наличии эффекта. Если уровень значимости уменьшен до минимального из реально используемых значений, т.е. до 0,01, то число ложных заключений о наличии эффекта уменьшится примерно до 6.

Таким образом, при большом числе сравнений для некоторых из них просто вследствие случайности значения статистических критериев превысят критические значения, что приводит к заключению о наличии связей, даже если на самом деле связей между переменными не существует. Чем больше проводится сравнений, тем выше вероятность ложного выявления якобы статистически значимых различий, и тем больше число ошибочно выявленных связей.

Этот феномен и называется проблемой множественных сравнений (или множественных проверок). Следовательно, доказательность выводов клинико-статистических исследований зависит от того, планировалось ли до начала исследования проведение сравнений, описанных в публикации, или же опубликованные сравнения — те, в которых формально выявлено различие (обнаружена связь), из сотен и тысяч сравнений, проведенных при анализе данных.

К сожалению, из публикуемых результатов исследований не всегда можно узнать, сколько в действительности было сделано сравнений. Часто важные находки отбираются среди большого числа сравнений, не приведших к обнаружению связи.

При подготовке результатов научного медицинского исследования к публикации, естественно, происходит отбор материала. Из большого объема результатов применения статистических методов для анализа конкретных медицинских данных авторы отбирают для публикации лишь наиболее интересные. Необходимо подчеркнуть, что субъективное принятие решений о том, какие из проведенных расчетов важны, а какие нет, может привести к значительному искажению действительности. «Находки», отобранные для публикации среди массы результатов проверок статистических гипотез, приведших к заключению об отсутствии эффектов (связей), могут не подтвердиться при проведении дальнейших исследований. В рассмотренном выше примере следует ожидать обнаружения примерно 30 (при уровне значимости 0,05) или 6 (при уровне значимости 0,01) связей. Все они — ложные, будут отклонены при дальнейших исследованиях. Хорошо это или плохо для авторов гипотетического исследования, рассмотренного в примере?

Феномен множественных проверок важно учитывать не только при попарном сравнении большого числа групп и подгрупп. Он встречается и при применении многих других статистических методов доказательной медицины. Например, при выявлении множества информативных признаков в задачах регрессионного и дискриминантного анализов (т.е. при восстановлении зависимостей и построении правил диагностики). А также при анализе таблиц сопряженности, когда формально строят т.н. детерминации, т.е. находят такие наборы значений признаков, в которых конкретные принимаемые значения одних признаков полностью определяют значения других признаков (естественно, при фиксированных исходных данных). Пример детерминации: рыжий мужчина в возрасте 20–29 лет, инвалид, с легочной патологией и родовой травмой, обязательно является наркоманом. Казалось бы, интересная «находка». Однако если оказывается, что таких мужчин в выборке из 1 000 человек всего двое, интерес к «находке» сразу снижается.

Доказательная медицина и МИС. Как подробно обосновано в разделах 12.1 и 12.2, современная доказательная медицина немыслима без интенсивного использования информационных компьютерных технологий и медицинских информационных систем. Только с их помощью можно отслеживать течение заболеваний у тысяч больных, оценивать риски, строить диагностические и прогностические правила, применять продвинутые статистические методы и разрабатывать математические модели медико-биологических явлений и процессов.

Коротко говоря, МИС — инструмент врача-исследователя. Как и всякий инструмент, он имеет достоинства и недостатки. О достоинствах сказано много. К недостаткам относится, например, то, что до сих пор не решена проблема множественных сравнений.

Иногда надеются, что МИС — в форме т.н. «экспертных систем» — смогут заменить врача. Это — утопия. Изложим соображения по этому поводу академика РАН Д.А. Поспелова [11].

Классический подход в направлении исследований, известном как «искусственный интеллект», породил идею разработки «экспертных систем» (один из видов программных продуктов) в области медицины. Важная их составляющая — схемы вывода новых истинных утверждений из базовых, играющих роль аксиом, и ранее полученных утверждений. Схемы вывода предполагалось формулировать на основе опыта врачей, а в качестве исходных данных использовать информацию, полученную от пациентов, и результаты измерений и анализов. Была надежда, что экспертные системы смогут ставить диагнозы и назначать лечение.

Но уже к середине 1970-х Гг. постепенно выясняется, что классических логических моделей и схем вывода явно не хватает для того, чтобы строить достаточно богатые и практически значимые интеллектуальные системы. Логический подход в его классической форме требовал для каждой предметной области, в частности, для кардиологии или пульмонологии, наличия полного перечня исходных положений, которые можно было бы считать аксиомами этой предметной области. Их существование (сюда естественно включаются и априорно задаваемые правила вывода) обеспечивало замкнутость используемых моделей, позволяло ставить и решать круг проблем, связанных с полнотой, результативностью и непротиворечивостью используемых моделей и процедур.

Однако конкретные предметные области, в которых стремился действовать искусственный интеллект, оправдывая свою практическую значимость, в подавляющем большинстве случаев не давали возможностей построения аксиоматических систем. Знания о предметных областях, как правило, были неполными, неточными и лишь правдоподобными. Это приводило к различным нежелательным эффектам: сбоям в процессах получения результатов, возникновению фальсификатов ранее полученных утверждений, быстрому снижению достоверности утверждений, получаемых в результате последовательного процесса логического вывода.

Так возникла проблема замены формальной системы с присущими ей процедурами дедуктивного вывода иной, столь же мощной моделью, где отражались бы основные особенности поиска решения в плохо определенных предметных областях, которые описываются как открытые системы с обновляемыми знаниями об их строении и функционировании.

С конца 70-х гг. XX в. старая парадигма, опирающаяся на идею строгого логического вывода, начинает постепенно сменяться новой парадигмой, провозглашающей, что основной операцией при поиске решения должна быть правдоподобная аргументация. Однако, в отличие от завершенной структуры логического вывода, до сих пор не существует столь же стройной, научно разработанной теории правдоподобной аргументации. Хорошо известные специалистам эффекты, связанные с появлением парадоксов при рассуждениях, показывают, что переход к более богатой по сравнению с моделью логического вывода модели правдоподобной аргументации неизбежно приводит к большому количеству новых проблем, связанных с обоснованием подобной модели и изучением ее особенностей. По мнению Д.А. Поспелова [13], в ближайшие десятилетия усилия многих специалистов сосредоточатся именно в этой области исследований.

Таким образом, в настоящее время экспертные системы не могут претендовать на самостоятельное участие в процессе ведения больных. Исключением могут быть ситуации, когда врач недоступен — на корабле в открытом море, в отдаленном населенном пункте и т.п. Главное сейчас, как и раньше — знания и интуиция врача. А МИС и, в частности, экспертные системы, лишь помогают ему. Однако потенциальные возможности этих помощников весьма велики, и современный врач должен постоянно и умело пользоваться медицинскими информационными системами.

Анализ медицинских данных как научное направление — это применение статистических методов в медицине, прежде всего в научных медицинских исследованиях. Можно сравнить с эконометрикой — применением статистических методов в экономике и менеджменте. Как и в других научных областях, базовую роль играет описание медицинских данных, измеренных в тех или иных шкалах (о теории измерений см. раздел 11.3). Инструменты статистического описания данных, рассмотренные в главе 2 (таблицы, диаграммы, графики и др.), зачастую позволяют сформулировать основные выводы исследования, которые затем подтверждаются методами оценивания и проверки гипотез. Примером являются коллективные работы, посвященные клинико-статистическому анализу результатов лечения больных острой пневмонией [14, 15].

Какие статистические методы наиболее часто используются при анализе медицинских данных? Судя по наукометрическому анализу публикаций, проведенному в работах [7, 16], второе место по распространенности после методов описания данных занимают методы интервального оценивания доли, сравнения долей (глава 1), проверки однородности независимых и связанных выборок (глава 5).

На третьем месте — изучение связи между признаками с помощью коэффициентов корреляции и регрессионного анализа (глава 6). К сожалению, не всегда осознается принципиальная разница между активным и пассивным экспериментом [17]. Напомним, что активный эксперимент предполагает проведение измерений при специально выбранных (согласно плану эксперимента) значениях факторов, а потому позволяет предсказывать результаты лечебных воздействий. В то время как при пассивном эксперименте исследователь не вмешивается в ход процесса, он может прогнозировать его течение, но не может ничего сказать о результатах воздействия на процесс, поскольку не в состоянии изучить его внутренний механизм.

Далее отметим дискриминантный анализ, т.е. математические и статистические методы диагностики заболеваний, и кластер-анализ, позволяющий выделять варианты течения заболевания (глава 6). Анализ временных рядов (главы 7 и 9), в том числе методы обнаружения разладки (глава 10), применяют при изучении динамики характеристик, описывающих состояние пациента, например, при электрокардиографии (ЭКГ). Для обработки т.н. топокарт, получаемых при одновременном получении ЭКГ со многих отведений, необходимы методы статистики нечисловых данных [18, глава 11], прежде всего статистики случайных множеств [19, 20]. Перспективны методы математического, в том числе вероятностно-статистического, моделирования функционирования различных структур человеческого организма [21].

Индивидуальные и коллективные экспертные оценки (глава 11) широко используются как в научных медицинских исследованиях, так и при управлении ими [22].

На основе практического опыта можно констатировать, что статистические методы представляют собой неотъемлемую часть инструментария врача-исследователя, причем эта часть непрерывно расширяется и усовершенствуется [16, 23, 24]. Практически все методы современной прикладной статистики [18] могут быть с успехом использованы в научных медицинских исследованиях. Более того, основанная на использовании высоких статистических технологий доказательная медицина основана на принципиально новой парадигме врачебного мышления [1]. Компьютеры, информационные технологии и статистические методы — надежные помощники врачей. Однако никогда формальные расчеты не смогут заменить людей при принятии решений (хотя современную теорию принятия решений [25] необходимо использовать в медицине, как и во всех иных областях деятельности). Ответственность — всегда лежит на враче!

Автор искренне благодарен академику АН СССР и РАН Е.И. Чазову, под руководством которого работал в 4-м Главном управлении при Минздраве СССР в качестве заведующего математическим отделением Центральной научно-исследовательской лаборатории.

Литература

1. *Флетчер, Р.* Клиническая эпидемиология: основы доказательной медицины / Р. Флетчер, С. Флетчер, Э. Вагнер. — Москва : Медиа Сфера, 1998. — 352 с.

2. *Власов, В.В.* Введение в доказательную медицину / В.В. Власов. — Москва : Медиа Сфера, 2001. — 392 с.

3. *Голдстейн, М.* Как мы познаем. Исследование процесса научного познания / М. Голдстейн, И.Ф. Голдстейн ; сокращенный перевод с английского — Москва : Знание, 1984. — 256 с.

4. Антигипертензивная эффективность дилтиазема и его влияние на эндотелиальную функцию у мужчин с мягкой и умеренной артериальной гипертонией / Д.В. Небиеризде, А.С. Сафарян, В.А. Метельская [и др.]. — Кардиоваскулярная терапия и профилактика. — 2004. — Т. 3. — № 1. —С. 4–9.

5. *Громнацкий, Н.И.* Коррекция тромбоцитарного гемостаза, нарушения толерантности к глюкозе, дислипидемии и перекисного окисления липидов у больных метаболическим синдромом / Н.И. Громнацкий, И.Н. Медведев. — Кардиоваскулярная терапия и профилактика. — 2004. — Т. 3. — № 1. — С. 10–15.

6. *Новиков, Д.А.* Статистические методы в педагогических исследованиях (типовые случаи) / Д.А. Новиков. — Москва : МЗ-Пресс, 2004. — 67 с.

7. *Новиков, Д.А.* Статистические методы в медико-биологическом эксперименте (типовые случаи) / Д.А. Новиков, В.В. Новочадов. — Волгоград : Изд-во ВолГМУ, 2005. — 87 с.

8. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 3-е изд., испр. и доп. — Москва : Экзамен, 2004. — 576 с.

9. *Новик, А.А.* Руководство по исследованию качества жизни в медицине / А.А. Новик, Т.И. Ионова. — Санкт-Петербург : Нева ; Москва : ОЛМА-ПРЕСС Звездный мир, 2002. — 320 с.

10. Эффективность и безопасность низкомолекулярного гепарина надропарина в удвоенной концентрации, вводимого один раз в сутки, в сравнении со стандартным режимом дозирования надропарина при остром коронарном синдроме без подъема ST (результаты рандомизированного, контролируемого исследования SAFRAX в России) / С.В. Шалаев, М.Я. Руда, Н.А. Грацианский и др. // Кардиоваскулярная терапия и профилактика. — 2004. — Т. 3. — № 1. — С. 40–45.

11. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1983. — 416 с.

12. *Федосеев, В.Н.* Управление персоналом организации / В.Н. Федосеев, С.Н. Капустин. — Москва : Экзамен, 2003. — 368 с.

13. *Поспелов, Д.А.* Десять «горячих точек» в исследованиях по искусственному интеллекту / Д.А. Поспелов // Интеллектуальные системы (МГУ). — 1996. — Т. 1. — Вып. 1–4. — С. 47–56.

14. Результаты лечения больных острой пневмонией / А.Е. Рабухин, В.П. Сильвестров, А.В. Винокурова и др. // Актуальные вопросы клинической

и экспериментальной медицины. — Москва : 4-е Главное Управление при Минздраве СССР, 1978. — С. 132–138.

15. Об эффективности этапного лечения больных острой пневмонией (стационар — реабилитационный центр) / В.П. Сильвестров, А.И. Романов, Н.С. Леонтьева и др. // Актуальные вопросы клинической и экспериментальной медицины. — Москва : 4-е Главное Управление при Минздраве СССР, 1980.

16. *Орлов, А.И.* О применении статистических методов в медико-биологических исследованиях / А.И. Орлов // Вестник Академии медицинских наук СССР. — 1987. — № 2. — С. 88–94.

17. *Налимов, В.В.* Теория эксперимента / В.В. Налимов. — Москва : Наука, 1971. — 208 с.

18. *Орлов, А.И.* Прикладная статистика : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 672 с.

19. Кинетотопография в диагностике инфаркта миокарда / Г.А. Аксенова, Е.С. Кузьмина, А.И. Орлов, Н.К. Розова // Актуальные вопросы клинической и экспериментальной медицины. — Москва : 4-е Главное Управление при Минздраве СССР, 1979. — С. 24–26.

20. Кинетокардиография в определении зон асинергии у больных инфарктом миокарда / В.Г. Попов, Г.А. Аксенова, А.И. Орлов и др. // Клиническая медицина. — 1982. — Т. LX. — № 3. — С. 25–30.

21. Математическая модель роли концентрирующей функции почки в патогенезе хронической артериальной гипертензии / Л.В. Архаров, А.И. Орлов, Ю.Л. Перов // Тезисы докладов III симпозиума «Морфометрия и математическое моделирование патологических процессов» (Кутаиси, 1–3 октября 1980 г.). — Кутаиси : Изд-во ЦОЛИУвр., 1980. — С. 35–37.

22. *Раушенбах, Г.В.* Экспертные оценки в медицине. Научный обзор / Г.В. Раушенбах, О.В. Филиппов. — Москва : ВНИИММТИ Минздрава СССР, 1983. — 80 с.

23. Перспективы применения статистики объектов нечисловой природы в медико-биологических исследованиях / А.И. Орлов, Г.В. Раушенбах, О.В. Филиппов // Применение математических методов и ЭЦВМ в медико-биологических исследованиях. Тезисы Всесоюзного симпозиума (Ленинград, 14–15 декабря 1982 г.). — Ч. 1. — Ленинград : Министерство здравоохранения СССР, 1982. — С. 80–82.

24. *Орлов, А.И.* О некоторых математических задачах, возникающих при обработке медицинских данных / А.И. Орлов // Статистика. Вероятность. Эко-

номика. Ученые записки по статистике. — Т. 49. — Москва : Наука, 1985. — С. 323–326.

25. *Орлов, А.И.* Теория принятия решений : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 576 с.

26. *Орлов, А.И.* Статистические модели в медицине / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 124. — С. 954–983.

Контрольные вопросы и задачи

1. Почему доказательная медицина невозможна без интенсивного использования высоких статистических и информационных технологий?

2. Какие варианты слепого метода Вы знаете?

3. Является ли обоснованным использование критерия Стьюдента для проверки наличия эффекта?

4. В каких случаях использование критерия Стьюдента не приводит к ошибочным заключениям?

5. Пусть в некотором году на территории Ярославской области было выявлено 175 случаев определенного заболевания, а в следующем году — 247 случаев.

6. Почему необходима стандартизация выборки?

Темы докладов, рефератов, исследовательских работ

1. Сопоставьте традиционные научные медицинские исследования и работы по доказательной медицине.

2. Информационные технологии в доказательной медицине.

3. Методы обнаружения эффекта при сравнении независимых и связанных выборок.

4. Статистические методы диагностики.

5. Экспертные системы в медицине.

6. Статистические методы анализа топокарт.

7. Методы оценивания характеристик распределения по стандартизованной выборке.

ГЛАВА 13. СТАТИСТИЧЕСКИЕ МЕТОДЫ В СОЦИОЛОГИИ

Социология (от латинского societas, т.е. общество, и греческого logos, т.е. учение) — наука, изучающая общество. Среди общественных отношений важное место занимают экономические. Изучение предпочтений потребителей входит в состав маркетинговых исследований, поэтому не удивительно, что социологии относят маркетинг к сфере своих профессиональных интересов. Обращает на себя внимание общность методов, применяемых социологами и маркетологами — выборочные исследования (см. главу 1), экспертные оценки (см. главу 11). Впрочем, те же статистические методы используются и в медицинских научных исследованиях (см. главу 12), и в психологии, и во многих иных областях.

В социологических исследованиях применяются различные статистические методы и модели анализа данных. После анализа динамики развития статистического инструментария социологов кратко рассмотрим применение теории люсианов для анализа дихотомических данных. Сколько градаций признаков надо использовать в анкетах? Ответ можно получить с помощью методов анализа сгруппированных данных и принципа уравнивания погрешностей. В качестве следствия получаем полезные выводы для теории управления запасами (см. раздел 8.3). Методы социометрии рассматриваем применительно к управлению малыми группами. Прикладные вопросы обсуждаются в разделах, посвященных анализу данных о научных организациях и маркетингу образовательных услуг (на примере Вечерней математической школы при Московском математическом обществе).

13.1. Развитие статистического инструментария социологов

Российская социология бурно растет по всем количественным параметрам. Если в 1989 г. в России было 6 социологических факультетов, отделений, кафедр, то в 2003 г. — уже 105. Число студентов-социологов выросло более чем в 100 раз. Во всех вузах преподают социологию (она вошла в перечень «Общих гуманитарных и социально-экономических дисциплин» государственных образовательных стандартов высшего профессионального образования). Издается более 20 социологических журналов. Каждый год ВАК утверждает около 50 докторских диссертаций по социологии [1, с. 2–3].

Очевидно, глубину исследованиям придает использование развитого научного аппарата — методологии и методов сбора и анализа данных, математических моделей. На наш взгляд, принципиальный прорыв был осуществлен в нашей стране в 1970-е гг. Именно тогда в арсенале отечественных социологов появились теория измерений и нечеткие множества, математические методы классификации и многомерное шкалирование, непараметрическая статистика и статистика нечисловых данных.

В дальнейшие десятилетия шло естественное развитие научного аппарата. К сожалению, нельзя сказать, что в последние годы темпы этого развития усилились. Действующие лица 1970-х гг. выпустили учебники [2–5], но поток научных результатов в области математических методов в социологии не расширился по сравнению с 1970-ми гг. — периодом «бури и натиска». Из этого следует, в частности, что публикации тех лет [6–8] отнюдь не устарели, они представляют большой интерес для статистиков, социологов и маркетологов XXI в.

Итак, социология бурно развивается вширь, но весьма медленно — вглубь. Это вполне естественно. Прочитав в 1970 г. популярную книгу В.Э. Шляпентоха [9], автор настоящего учебника провел свое первое полевое исследование. Несмотря на простоту статистического инструментария, оно позволило решить управленческие задачи, стоявшие перед автором как директором Вечерней математической школы (ВМШ) при Московском математическом обществе. Итоги многолетней деятельности в ВМШ по маркетингу образовательных услуг подведены в монографии [10]. И сейчас наши ученики-маркетологи, готовя выпускные работы на степень магистра делового администрирования, обходятся полевыми исследованиями на столь же простом уровне (см. описание исследования «Потребители растворимого кофе» в главе 1 и [5, гл. 2]).

И лишь постепенно практики приходят к необходимости применять более сложные методы. Например, в крупном маркетинговом агентстве, опрашивающем за год около 0,5 млн потребителей, в котором автор этих слов работал консультантом, был создан специализированный отдел обработки данных, сотрудники которого ежедневно применяли различные алгоритмы статистической обработки данных, в частности, включенные в известный пакет SPSS.

Как показывает анализ тезисов докладов и выступлений на II Всероссийском социологическом конгрессе «Российское общество и социология в XXI веке: социальные вызовы и альтернативы» [11], большинство участников конгресса не проводит полевых исследований и не испытывает потребности в применении математических и статистических методов. Такие методы необходимы для продвинутых социологических исследований. Очевидно, с укреплением социологических центров в них возникают подразделения анализа данных, которые сначала пользуются стандартными статистическими пакетами, а затем востребуют и современные методы.

В социологии с успехом используются различные методы анализа данных и разнообразные математические модели [12]. Обсудим развитие методов обработки данных в социологических исследованиях за последние тридцать лет.

Математические методы выборочных исследований. Выборочные исследования — один из основных инструментов социологов. Для переноса выводов с выборки на всю интересующую исследователя совокупность необходимо использовать вероятностно-статистические методы и модели. Уже в 1970-х гг. в нашей стране активно разрабатывались продвинутые математические и статистические методы анализа данных социологических опросов (см., например, сборники [6, 7]). Отметим, что работы тех уже далеких лет, как правило, отнюдь не устарели и по-прежнему представляют интерес для специалистов по анализу социологических данных и математическому моделированию социальных процессов. Однако за тридцать прошедших лет в некоторых направлениях удалось существенно продвинуться. Основное содержание настоящего раздела — обсуждение развития ориентированных на социологию статистических методов и моделей за последние тридцать лет в нашей стране.

Одни и те же математические и статистические методы и модели могут с успехом применяться в самых разных областях науки и практики. Статистические методы и модели весьма эффективны в социологических, социально-экономических, управленческих, технических и технико-экономических исследованиях, медицине, истории, практически в любой прикладной отрасли и области знания. Очевидна связь между исследованиями, выполненными в рамках различных дисциплин. Например, на Втором Всероссийском социологическом конгрессе (2003) активно обсуждалась такая традиционно экономическая тематика, как маркетинговые и инновационные исследования [11]. Однако для специалиста вполне естественным является желание «замкнуться» внутри своей предметной области. Например, довольно странным выглядело бы предложение о преподавании на социологическом факультете в соответствии с учебником по эконометрике [5]. Удивление значительно возросло бы при констатации того, что этот учебник составлен в основном из статей, опубликованных в журнале «Заводская лаборатория» (в прошлом — орган Министерства черной металлургии). Действительно, есть ли что-либо общее у инженера-металлурга, менеджера, экономиста и социолога? Необходимо известное интеллектуальное развитие, чтобы понять, что все эти специалисты могут использовать одни и те же инструменты исследования — статистические методы и модели.

В рассматриваемой области основное событие последних тридцати лет — это становление научно-практической дисциплины «прикладная статистика», посвященной разработке и применению статистических методов и моделей (см. главу 15). На Западе, как мы убедились в процессе становления Всесоюзной статистической ассоциации, аналогичный процесс начался несколько раньше и протекает иначе из-за сложившихся традиций и отличия научно-организационных форм.

В статье [13] мы рассмотрели развитие идей и научной области, а не персоналии. В краткой, но весьма содержательной сводке [25] описаны основные научные результаты большого числа отечественных исследователей в области статистических методов анализа социологических данных, названы основные исследовательские коллективы Москвы, Петербурга, Новосибирска и многих иных городов, приведена обширная библиография (119 названий). Подробное же описание требует серии книг, а не статьи.

Специфика социологических данных. В социологии велика доля нечисловых данных. Среди измеряемых факторов от 70 до 90 % и более составляют факторы, имеющие нечисловую природу. Примерами нечисловых данных являются:

- значения качественных признаков, т.е. результаты кодировки объектов (например, ответов на вопросы социологической анкеты) с помощью заданного перечня категорий (градаций);

- результаты парных сравнений, т.е. последовательности из 0 и 1;

- результаты множественных сравнений и бинарные отношения — упорядочения (ранжировки), классификации (отношения эквивалентности), толерантности[13],

- множества (обычные или нечеткие),

- слова, предложения, тексты;

- вектора, координаты которых — совокупность значений разнотипных признаков, часть из них носит качественный характер, а часть — количественный;

- ответы на вопросы экспертной, маркетинговой или социологической анкеты, часть из которых носит количественный характер (возможно, интервальный), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты и т.д.

[13] Толерантность — это рефлексивное симметричное отношение. Отличается от классификации (отношения эквивалентности) возможным отсутствием транзитивности. Толерантностями естественно описывать отношения сходства или знакомства. Вероятностно-статистическая теория толерантностей содержится в монографии [8].

Все эти виды нечисловых данных (и многие другие) используются при проведении социологических и маркетинговых исследований. В частности, необходимость обрабатывать результаты измерений в шкалах, отличных от абсолютной (прежде всего в номинальной и порядковой шкалах), объясняет важную роль теории измерений (см. раздел 11.3) в социологии [3, 8]. Большое значение имеет выбор методов обработки данных, адекватных используемым шкалам измерения. Так, в [5, 8, 15] построена теория, связывающая способы усреднения данных со шкалами измерения.

Ясно из сказанного, что для социологии определяющее значение имеют методы статистического анализа нечисловых данных. В течение 1970-х гг. на основе запросов социологии [16], экономики, техники и медицины развивались конкретные направления статистики объектов нечисловой природы. Были установлены связи между конкретными видами таких объектов, разработаны для них вероятностные модели. Научные итоги этого периода подведены в монографии [8]. Следующий этап (1980-е гг.) — выделение статистики нечисловых данных в качестве самостоятельного направления в прикладной статистике. Предварительные итоги были подведены в сборнике научных статей [17], полностью посвященном статистике нечисловых данных. Этот сборник был подготовлен комиссией «Статистика нечисловых данных» Научного совета АН СССР по комплексной проблеме «Кибернетика» и Институтом социологии АН СССР.

В настоящее время статистика нечисловых данных с теоретической точки зрения достаточно хорошо развита, основные идеи, подходы и методы были разработаны и изучены математически, в частности, доказано достаточно много теорем. Наступило время перейти к применению полученных результатов на практике. Одним из примеров такого применения являются работы по социологии науки (см. раздел 13.5 и [18]). Современное состояние статистики нечисловых данных отражено в монографиях [8, 15].

13.2. Перспективы применения люсианов в социологии

Как уже отмечалось, среди перспективных для социологии статистических методов выделяется по своей практической значимости статистика нечисловых данных, поскольку полученная от респондентов или экспертов информация часто носит нечисловой характер [17]. Частью статистики нечисловых данных является теория люсианов.

Наиболее простой (по форме) вопрос к респонденту или к эксперту — это вопрос, на который возможны лишь два ответа — «да» или «нет». В вероятностной модели ответ на такой вопрос описывается испытанием Бернулли, т.е. случайной величиной, принимающей два значения, для определенности, 1 и 0. Событие, которому соответствует ответ, обозначенный как 1, условно называется «успехом». Распределение испытания Бернулли описывается одним числом — вероятностью успеха.

При опросе респондентов и при проведении экспертного опроса задают не один вопрос, а несколько. Пусть на каждый из них возможны лишь два ответа. Тогда совокупность ответов кодируется последовательностью из 1 и 0. Ясно поэтому, что среди вероятностно-статистических методов анализа социологических данных центральное место должны занимать методы теории люсианов.

Люсианы — это конечные (длины k) последовательности независимых испытаний Бернулли с, вообще говоря, разными вероятностями успеха [5, 15]. Распределение люсиана A задается вектором параметров $P = (p_1, p_2..., p_k)$, где p_i — вероятность того, что i-я координата люсиана A равна 1 (и с вероятностью $1 — p_i$ она равна 0), $i = 1, 2, ..., k$.

В теории люсианов центральное место занимают методы проверки трех гипотез — согласованности, однородности и независимости [15].

Пусть $A_1, A_2, ..., A_s$ — независимые (между собой) люсианы с векторами параметров $P_1, P_2, ..., P_s$ соответственно. *Гипотезой согласованности люсианов* называют гипотезу $P_1 = P_2 = ...= P_s$.

Случайные толерантности — частный случай люсианов [8]. Для других видов бинарных отношений, широко используемых в экспертных технологиях, — ранжировок и разбиений — под согласованностью понимают более частную гипотезу, предполагающую отрицание равномерности распределений (т.е. одинаковой вероятности появления каждой возможной ранжировки или разбиения), что соответствует замене гипотезы согласованности на гипотезу равномерности $P_1 = P_2 = ...= P_s = (1/2, 1/2, ..., 1/2)$.

Гипотеза согласованности, очевидно, более адекватна конкретным задачам обработки данных опросов или экспертных оценок, чем гипотеза равномерности. Поэтому реальные данные, обычно содержащие противоречия, целесообразно рассматривать как люсианы и проверять гипотезу согласованности, а не подбирать ближайшие ранжировки или разбиения с дальнейшей проверкой согласованности методами теории случайных ранжировок или разбиений, как иногда рекомендуется.

Пусть $A_1, A_2, ..., A_m$ и $B_1, B_2, ..., B_n$ — независимые в совокупности люсианы длины k, одинаково распределенные в каждой группе с параметрами $P(A)$ и $P(B)$ соответственно. *Гипотезой однородности* называют гипотезу $P(A) = P(B)$.

Пусть (A_i, B_i), i = 1. 2, ..., s — последовательность (фиксированной длины) пар люсианов. Пары предполагаются независимыми между собой. Требуется проверить гипотезу независимости A_i и B_i, т.е. независимость внутри пар. В ранее введенных обозначениях *гипотеза независимости* — это гипотеза $P(X_{ij}(A) = 1, X_{ij}(B) = 1) = P(X_{ij}(A) = 1)P(X_{ij}(B) = 1)$, где $i = 1, ..., s$; $j = 1, ..., k$, проверяемая в предположении $P_1(A) = P_2(A) = ... = P_s(A)$, $P_1(B) = P_2(B) = ... = P_s(B)$.

В последние десятилетия (с начала 70-х годов) в прикладной статистике все большее распространение получают постановки, в которых число неизвестных параметров растет вместе с объемом выборки. Результаты, полученные в подобных постановках, называют найденными «в асимптотике растущей размерности» или «в асимптотике А.Н. Колмогорова», перенося терминологию исследований по дискриминантному анализу на общий случай. Эта асимптотика естественна при обработке многих видов технических, организационно-экономических, социологических, медицинских данных. В теории люсианов используем асимптотику растущей размерности s = const, $k \to \infty$. При проверке однородности принимаем, что m и n постоянны, а $k \to \infty$. При этом число неизвестных параметров растет пропорционально объему данных.

Для проверки рассматриваемых гипотез применяем метод, основанный на использовании несмещенных оценок, построенных по малым выборкам [8, 15]. При этом исходим из вектора попарных расстояний между люсианами $\xi = \{d(A_p, A_q), 1 \le p < q \le s\}$, в котором пары индексов (p, q) упорядочены лексикографически, а расстояние $d(A_p, A_q)$ — число несовпадающих координат в векторах A_p и A_q. Расчетные формулы и алгоритмы приведены в литературе [8, 15].

Метод проверки гипотез. Пусть имеются k выборок, независимых между собой. Пусть при справедливости нулевой гипотезы по каждой из выборок можно построить несмещенную оценку $\xi_i \in R^p$ векторного нуля $0 \in R^p$, где $p \ge 1$, $i = 1, 2, ..., k$. Другими словами, пусть распределение i-й выборки описывается параметром θ_i, лежащим в произвольном пространстве, а нулевая гипотеза, очевидно, состоит в том, что $\theta_i \in \Theta_{0i}$, где Θ_{0i} — собственное подмножество множества $\{\theta_i\}$. Предполагается, что можно по i-й выборке вычислить статистику ξ_i такую, что

$$M\xi_i = 0 \qquad (1)$$

при всех $\theta_i \in \Theta_{0i}$. Очевидно, $\xi_i \equiv 0$ удовлетворяют (1). Однако для рассматриваемого метода необходимо, чтобы при всех $\theta_i \in \Theta_{0i}$ ковариационная матрица вектора ξ_i была ненулевой:

$$Cov(\xi_i) = M(\xi_i^T \xi_i) \ne 0\,. \qquad (2)$$

Следующее предположение — ковариационные матрицы статистик ξ_i, т.е. $Cov(\xi_i)$, также допускают несмещенные оценки S_i по тем же выборкам:

$$M(S_i) = Cov(\xi_i) \tag{3}$$

при всех $\theta_i \in \Theta_{0i}$.

Рассматриваемый метод основан на том, что поскольку случайные вектора ξ_i определяются по независимым между собой выборкам, то ξ_i независимы в совокупности, а потому случайный вектор

$$\xi = \sum_{i=1}^{k} \xi_i \tag{4}$$

является суммой независимых случайных векторов, имеет в силу (1) нулевое математическое ожидание, а его ковариационная матрица равна

$$C_k = \sum_{i=1}^{k} Cov(\xi_i)\ .$$

При справедливости многомерной центральной предельной теоремы (простейшее условие справедливости этой теоремы для ξ_i в случае люсианов — отделенность от 0 и 1 всех элементов матриц P_j, равномерная по s и k) вектор ξ является асимптотически нормальным, т.е. при $k \to \infty$ распределение ξ сближается (в смысле, раскрытом в главе 14) с многомерным нормальным распределением $N(0; C_k)$.

Однако эту сходимость нельзя непосредственно использовать для проверки исходной гипотезы, поскольку матрица C_k неизвестна статистику. Необходимо оценить эту матрицу по статистическим данным. В силу (3) в качестве оценки C_k стоит использовать

$$C_k^* = \sum_{i=1}^{k} S_i\ .$$

Простейшая формулировка условий справедливости такой замены — предположение о том, что к последовательности S_i можно применить закон больших чисел. А именно, пусть существует неотрицательно определенная матрица C такая, что при $k \to \infty$

$$\frac{1}{k}(C_k^* - C_k) \to 0, \quad \frac{1}{k} C_k \to C\ . \tag{5}$$

В силу результатов главы 14 из асимптотической нормальности ξ и соотношений (5) следует, что распределение статистики

$$\eta = \frac{1}{\sqrt{k}}\xi$$

сходится к нормальному распределению $N(0; C)$. При этом, если некоторый случайный вектор τ имеет распределение $N(0; C)$, то распределение случайной величины $q(\eta)$ сходится к распределению $q(\tau)$ для произвольной интегрируемой по Риману по любому кубу функции q: $R^p \to R^1$. Для проверки нулевой гипотезы предлагается пользоваться статистикой $q(\eta)$ при подходящей функции q, а процентные точки брать соответственно распределению $q(\tau)$. В этом и состоит рассматриваемый метод проверки гипотез о люсианах в асимптотике растущей размерности. Для реальных расчетов целесообразно использовать линейные или квадратические функции q от координат вектора η.

Отклонения от нулевой гипотезы приводят, как правило, к нарушению равенств (1) и (3). Случайный вектор η при этом обычно остается асимптотически нормальным.

Несмещенные оценки параметров асимптотического распределения вектора попарных расстояний. Применим описанный выше метод для проверки гипотезы согласованности люсианов. Исходные данные — люсианы

$$A_j = (X_{1j}, X_{2j}, \ldots, X_{kj}), j = 1, 2, \ldots, s.$$

В качестве i-й выборки возьмем совокупность испытаний Бернулли, стоящих на i-м месте в рассматриваемых люсианах:

$$X_{i1}, X_{i2}, \ldots, X_{is}. \tag{6}$$

При справедливости нулевой гипотезы в (6) стоят независимые испытания Бернулли с одной и той же вероятностью успеха p_i; при нарушении нулевой гипотезы согласованности независимость испытаний Бернулли сохраняется, но вероятности успеха могут различаться.

В качестве вектора ξ, на основе которого строятся статистики для проверки согласованности, будем использовать вектор попарных расстояний между люсианами

$$\xi = \{d(A_p, A_q), 1 \le p < q \le s\}, \tag{7}$$

в котором пары (p, q) упорядочены лексикографически,

$$d(A_p, A_q) = \sum_{i=1}^{k} \mu_i \,|\, X_{ip} - X_{iq} \,|, \quad \mu_i > 0. \tag{8}$$

Из вида расстояния в формуле (8) следует, что введенный в (7) вектор ξ является суммой независимых векторов, а именно, имеет вид (4) с

$$\xi_i = \mu_i\{|X_{ip} - X_{iq}|, 1 \le p < q \le s\}. \tag{9}$$

Следовательно, для применения описанного выше метода проверки гипотез о люсианах в асимптотике растущей размерности достаточно построить на основе вектора ξ_i из (9) несмещенную оценку 0 и найти несмещенную оценку ковариационной матрицы этой оценки.

Чтобы применить общую схему, необходимо начать с построения статистики β такой, чтобы при всех p_i имело место равенство

$$M(|X_{ip} - X_{iq}| - \beta) - 0, \ 1 \le p < q \le s.$$

Элементарный расчет дает:

$$M|X_{ip} - X_{iq}| = 2p_i(1 - p_i).$$

Как известно, несмещенная оценка многочлена

$$f(p) = \sum_{h=0}^{m} a_h p^h$$

по результатам m независимых испытаний Бернулли с вероятностью успеха p в каждом имеет вид

$$f^*(p) = \sum_{h=0}^{m} a_h \frac{\gamma^{[h]}}{m^{[h]}}, \tag{10}$$

где γ — общее число успехов в m испытаниях и использовано обозначение

$$n^{[h]} = n(n-1)...(n-h+1).$$

Ясно, что многочлены степени $m + 1$ и более высокой невозможно несмещенно оценить по результатам m испытаний.

В случае $f(p) = 2p(1 - p)$ в соответствии с (10) получаем несмещенную оценку

$$\beta = \frac{2}{m-1}\left(\gamma - \frac{\gamma^2}{m}\right). \tag{11}$$

Таким образом, можно применять общий метод проверки гипотез о люсианах в асимптотике растущей размерности с

$$\xi_i = \mu_i\,(\{|X_{ip} - X_{iq}|,\ 1 \le p < q \le s\} - \beta_i e),$$

где коэффициенты β_i определяются с помощью формулы (11) по γ_i — общему числу единиц, стоящих на i-м месте в люсианах $A_1, A_2, ..., A_s$, а e — вектор размерности $s(s - 1)/2$ с единичными координатами. Тогда несмещенная оценка 0, о которой идет речь в методе проверки гипотез по совокупности малых выборок, имеет вид

$$\xi = \{d(A_p, A_q),\ 1 \le p < q \le s\} - \sum_{i=1}^{k} \mu_i \beta_i e.$$

Для использования статистики типа η, распределение которой приближается с помощью нормального распределения

$$N\left(0;\ \frac{1}{k}\sum_{i=1}^{k} S_i\right),$$

необходимо уметь несмещенно оценивать ковариационные матрицы $Cov(\xi_i)$. Для этого достаточно найти математические ожидания элементов матрицы $M(\xi_i^T \xi_i)$ как функции (многочлены) от p_i, а затем использовать формулу (10) для получения несмещенных оценок.

Вычисление матрицы $M(\xi_i^T \xi_i)$ хотя и трудоемко, но не содержит каких-либо принципиальных трудностей. В [20] вычислены диагональные элементы рассматриваемой матрицы. Вычисление занимает около 2,5 книжных страниц (с. 299–301). Поэтому здесь приведен только окончательный итог.

Обозначим для краткости $p_i = p$. В [20] показано, что

$$D = D(|X_{ip} - X_{iq}| - \beta_i) = \left(2 - \frac{4}{s}\right)p(1-p) - 4\frac{(s-2)(s-3)}{s(s-1)}p^2(1-p)^2.$$

Если двухэлементные множества $\{p, q\}$ и $\{r, t\}$ не имеют ни одного общего элемента, то

$$C_1 = M(|X_{ip} - X_{iq}| - \beta_i)(|X_{ir} - X_{it}| - \beta_i) = -\frac{4}{s}p(1-p) + \frac{8(2s-3)}{s(s-1)}p^2(1-p)^2,$$

а если имеют ровно один общий элемент, то

$$C_2 = M(|X_{ip} - X_{iq}| - \beta_i)(|X_{ir} - X_{it}| - \beta_i) = \left(1 - \frac{4}{s}\right)p(1-p) - 4\frac{(s-2)(s-3)}{s(s-1)}p^2(1-p)^2.$$

С помощью формулы (10) получаем несмещенные оценки для D, C_1 и C_2 как многочленов от p:

$$D^* = \frac{2\gamma_i(s-\gamma_i)}{s^2(s-1)^2}\{(s-2)(s-1) - 2(\gamma_i - 1)(s-\gamma_i-1)\},$$

$$C_1^* = \frac{4\gamma_i(s-\gamma_i)}{s^2(s-1)}\left\{\frac{2(2s-3)(\gamma_i-1)(s-\gamma_i-1)}{(s-1)(s-2)(s-3)} - 1\right\},$$

$$C_2^* = \frac{\gamma_i(s-\gamma_i)}{s^2(s-1)^2}\{(s-4)(s-1) - 4(\gamma_i-1)(s-\gamma_i-1)\}.$$

С помощью трех чисел D^*, C_1^*, C_2^* выписывается несмещенная оценка матрицы ковариаций вектора ξ_i/μ_i, которую обозначим B_i. Тогда асимптотически нормальный вектор ξ имеет нулевое математическое ожидание и ковариационную матрицу, несмещенно и состоятельно (в смысле соотношений (5)) оцениваемую с помощью

$$Cov(\xi)^* = \sum_{i=1}^{k} \mu_i^2 B_i. \tag{12}$$

Асимптотическая нормальность доказывается, естественно, в схеме серий. Достаточным условием является существование положительной константы ε такой, что

$$\mu_i \geq \varepsilon,\ \frac{1}{\mu_i} \geq \varepsilon,\ p_i \geq \varepsilon,\ 1 - p_i \geq \varepsilon \tag{13}$$

при всех k и i, $1 \leq i \leq k$.

Поскольку D, C_1 и C_2 являются многочленами четвертой степени от p, то несмещенные оценки для них существуют при $s \geq 4$. Если же $s < 4$, то несмещенных оценок не существует. Поэтому указанным методом проверять согласованность можно лишь при числе люсианов $s \geq 4$.

Проверка согласованности люсианов. Пусть α — нормально распределенный случайный вектор размерности $s(s - 1)/2$ с нулевым математическим ожиданием и ковариационной матрицей, определенной формулой (12). Вместо распределения $f(\xi)$ для построения критериев проверки гипотез можно использовать распределение $f(\alpha)$. Более того, аналогичный результат верен при замене f на f_n (при слабых внутриматематических условиях регулярности, наложенных на последовательность функций f_n). Следовательно, для проверки гипотезы согласованности люсианов можно пользоваться любой статистикой $f_n(\xi)$, для которой могут быть вычислены на ЭВМ или заранее табулированы процентные точки распределения $f_n(\alpha)$, аппроксимирующего распределение $f_n(\xi)$.

В частности, можно использовать линейные статистики, представляющие собой скалярное произведение случайного вектора ξ и некоторого заданного детерминированного вектора коэффициентов a, т.е.

$$(\xi, a) = \sum_{i=1}^{k}\left(\mu_i \sum_{1 \leq j < t \leq s} a_{jt}(|X_{ij} - X_{it}| - \beta_i)\right). \tag{14}$$

Линейные статистики имеют нулевое математическое ожидание и дисперсию, очевидным образом выражающуюся через матрицу коэффициентов $\|a_{ij}\|$ и числа D, C_1 и C_2, а потому несмещенно и состоятельно оцениваемую с помощью с помощью выписанных выше оценок для D, C_1 и C_2.

Отметим, что $(\xi, a) = 0$ при $a_{ij} \equiv 1$, $1 \leq j < t \leq s$. Это следует как из непосредственного вычисления дисперсии (ξ, a), так и из того, что (ξ, a) в рассматриваемом случае выражается через достаточную статистику $(\gamma_1, \gamma_2, ..., \gamma_k)$ и является несмещенной оценкой нуля, а семейство биномиальных распределений полно, т.е. существует только одна несмещенная оценка нуля — тождественный нуль. Таким образом, сумма координат вектора ξ, т.е. непосредственный аналог коэффициента ранговой конкордации Кендалла — Смита из теории ранговой корреляции, тождественно равна 0.

Распределение статистики (14) при альтернативах изучено в работе [21]. Рассмотрим два частных случая.

Первый частный случай. Проверка согласованности двух определенных люсианов (ответов двух экспертов), j-го и t-го, может осуществляться с помощью статистики (14), в которой отличен от 0 только член с $a_{jt} = 1$. Оценкой дисперсии является D^*.

Второй частный случай. Пусть необходимо проверить согласованность люсианов с одним из них, скажем, с j-м (например, люсианы отражают мнения экспертов, а j-й из них является наиболее компетентным — по априорной оценке, или «лицом, принимающим решения», или его мнение сильно отличается от мнений остальных). Это можно сделать с помощью статистики (14), в которой

$$a_{jt} = 1,\ t - = j+1, j+2, ..., s;\ a_{tj} = 1,\ t = 1, 2, ..., j-1;$$
$$a_{qt} = 0,\ q \neq j,\ t \neq j,\ 1 \leq q < t \leq s.$$

Другими словами, она имеет вид

$$W = \sum_{t=1}^{s} d(A_j, A_t) - (s-1)\sum_{i=1}^{k} \mu_i \beta_i\,, \tag{15}$$

где расстояние d между люсианами определено в (8), а β_i — в (11) с заменой m на s и γ на γ_i. Используя полученные ранее несмещенные оценки элементов ковариационной матрицы, нетрудно показать, что несмещенная и состоятельная (в смысле формулы (5) выше) оценка дисперсии W имеет вид

$$D^*(W) = \sum_{i=1}^{k} \mu_i^2 \frac{\gamma_i(s-\gamma_i)}{s^2}\{(s-2)^2 - 4(\gamma_i - 1)(s - \gamma_i - 1)\}\,.$$

Тогда при выполнении некоторых внутриматематических условий регулярности, например, условий (13), распределение статистики

$$\frac{1}{\sqrt{D^*(W)}} W$$

сходится при $k \to \infty$, $s = \text{const}$ к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1 (при справедливости гипотезы (1) согласованности люсианов).

Исследования по теории люсианов, в том числе выполненные Г.В. Рыдановой, Т.Н. Дылько, Г.В. Раушенбахом, О.В. Филипповым, А.М. Никифоровым, рассмотрены в монографиях [5, 8, 15] и на сайте «Высокие статистические технологии» (http://orlovs.pp.ru).

Методы моделирования и анализа данных опросов и экспертных мнений, основанные на теории люсианов, заслуживают теоретического развития и широкого практического использования в социологии, а также в смежных областях — в маркетинговых исследованиях и при подготовке управленческих решений [19].

13.3. Асимптотика квантования и выбор числа градаций в социологических анкетах

Наличие погрешностей наблюдений приводит к тому, что результат наблюдения приходится рассматривать не как число, а как объект более сложной природы. «Размытость» реального результата наблюдения можно моделировать различными способами. В классической метрологии погрешность моделируется случайной величиной, математическое ожидание которой называют систематической ошибкой, а соответствующую центрированную случайную величину — случайной ошибкой. В статистике применяют группировку данных (см. раздел 3.5). Более общий подход состоит в использовании нечетких чисел [22]. Развиты методы статистики интервальных данных [15]. Наличие погрешностей наблюдений приводит к существенному изменению свойств статистических процедур.

Об учете ошибок в ответах респондентов. Рассмотрим один подход к определению шага квантования (группировки) при переходе от непрерывной шкалы к дискретной. Прикладная цель — выбор числа градаций в социологических анкетах [23]. В соответствии с методологией общей теории устойчивости (см. раздел 14.6) предлагается выбирать шаг так, чтобы ошибки, порожденные квантованием, были того же порядка, что и ошибки, присущие ответам респондентов (опрашиваемых). При конечном интервале изменения значений измеряемого признака шаг шкалы однозначно определяет число градаций. Оказывается, во многих вопросах закрытого типа достаточно указывать 3–6 градаций ответов (подсказок). Этот вполне конкретный вывод следует из теории, развитой в настоящем разделе.

Разработаем вероятностную модель, которую будем изучать. Пусть ν — истинное значение какого-либо признака. Респондент «измеряет» это значение с какой-то ошибкой, и его мнение отражает число $\xi = \nu + \alpha$, где α — случайная ошибка, присущая респонденту. Социолог просит респондента назвать не ξ, а одну из градаций, включенных в систему баллов, т.е. одно из чисел kh, $k = 0, \pm 1, \pm 2, \pm 3, \ldots$, где h — шаг шкалы. Естественно принять модель, согласно которой респондент называет градацию η, наиболее близкую к его мнению ξ. В результате социолог получает от респондента число $\eta = \xi + \beta$, где $\eta \in \{kh, k = 0, \pm1, \pm2, \pm3, \ldots\}$ и $-h/2 \le \beta < h/2$. Здесь β — методическая ошибка, порожденная инструментом социолога — системой баллов.

Очевидно, описанная процедура с ошибками двух типов — ошибкой измерения и ошибкой округления — применима для описания процесса измерения в различных областях, если итог выражается в равномерной шкале с конечным числом градаций. В частности, при оценке качества продукции в баллах, при выводе результата измерения на экран в цифровой форме, и т.д.

Рассмотрим теперь n респондентов, независимо друг от друга отвечающих на вопросы социолога. Респонденту с номером i соответствует истинное значение v_i (например, месячный доход или средние затраты времени на просмотр программ телевидения), о которых респондент имеет мнение $\xi_i = v_i + \alpha_i$ (или же: хочет сообщить социологу ξ_i, а не v_i). Используя систему баллов с шагом h, социолог получает $\eta_i = \xi_i + \beta_i$. Предположим, что социолога интересует среднее истинных значений по выборке, т.е.

$$\overline{v} = \frac{1}{n}\sum_{i=1}^{n} v_i .$$

Однако социолог может вычислить лишь

$$\overline{\eta} = \frac{1}{n}\sum_{i=1}^{n} \eta_i = \overline{v} + \overline{\alpha} + \overline{\beta}, \tag{1}$$

где

$$\overline{\alpha} = \frac{1}{n}\sum_{i=1}^{n} \alpha_i, \quad \overline{\beta} = \frac{1}{n}\sum_{i=1}^{n} \beta_i .$$

В (1) $\overline{\alpha}$ — ошибка, обусловленная природой респондентов, $\overline{\beta}$ — ошибка, связанная с использованием исследовательского инструмента — системы баллов. Уменьшение шага h (т.е. увеличение числа градаций) уменьшает суммарную ошибку, уменьшая дисперсию $\overline{\beta}$, однако дисперсия $\overline{\alpha}$ при этом не меняется, а потому дисперсия суммарной ошибки всегда больше дисперсии $\overline{\alpha}$. С другой стороны, увеличение числа градаций приводит к возрастанию сложности проведения опроса, к увеличению затрат на социологическое исследование. Поэтому чрезмерное увеличение числа градаций, т.е. чрезмерное уменьшение *h*, нецелесообразно. Согласно «принципу уравнивания погрешностей» в общей схеме устойчивости (см. главу 14.6) рациональным является такое *h*, при котором ошибки респондентов и ошибки исследовательского инструмента имеют одинаковый порядок. Чтобы вычислить указанные ошибки, необходимо построить вероятностную модель поведения совокупности респондентов.

Будем считать, что ξ_i являются независимыми случайными величинами, имеющими математические ожидания и дисперсии, причем

$$M(\xi_i) = \nu_i, \ \ i = 1,2,\dots,n,$$

т.е. мнение респондента, зависящее от случайных величин, является в среднем правильным (не содержит систематической ошибки). (Это предположение, разумеется, выполнено не всегда — влияние систематических погрешностей на статистические выводы продемонстрировано в статистике интервальных данных, см., например, [15, гл. 12].) Отметим, что β_i также являются независимыми случайными величинами (поскольку β_i — функции от случайных величин α_i). Ниже будет показано, что

$$\lim_{h \to 0} h^{-1} M(\beta_i) = 0,$$

кроме того, α_i и β_i при каждом *i* являются асимптотически независимыми при $h \to 0$. При малом *h* и большом *n* в силу многомерной центральной предельной теоремы и теорем о наследовании сходимости (см. разделы 14.2 и 14.3 соответственно) величины $\overline{\alpha}$ и $\overline{\beta}$ можно считать независимыми и нормально распределенными, а потому при указанных условиях случайная величина $\overline{\eta}$ асимптотически нормальна, причем

$$M(\overline{\eta}) \approx \overline{\nu}, \ \ D(\overline{\eta}) \approx D(\overline{\alpha}) + D(\overline{\beta}).$$

Условие «однопорядковости ошибок» приводит к рекомендации определять *h* из соотношения

$$D(\overline{\alpha}) = D(\overline{\beta}). \qquad (2)$$

В главе 3.5 установлено, что

$$\lim_{h\to 0}\frac{12D(\beta_i)}{h^2}=1 .$$

Поэтому $D(\overline{\beta})$ в (2) можно заменить при $h \to 0$ на $(n^{-1}h^2)/12$. Если предположить, что ошибки всех респондентов имеют одинаковую дисперсию, т.е. $D(\alpha_i)=\sigma^2$, то (2) можно преобразовать к эквивалентному при $h \to 0$ виду

$$\frac{\sigma^2}{n}=\frac{h^2}{12n}, \tag{3}$$

откуда

$$h=2\sqrt{3}\sigma\approx 3{,}47\sigma . \tag{4}$$

Соотношения (3) и (4) можно получить другим путем, исходя из уравнивания вероятностей неправильного упорядочения двух совокупностей (по двум рассматриваемым причинам). Опишем в терминах социологии соответствующую постановку задачи.

Пусть каждый респондент оценивает значения двух признаков. Так, в известных исследованиях В.Н. Шубкина [24] школьники оценивали привлекательность различных профессий. Пусть $v_i(1)$ и $v_i(2)$ — «истинные значения» рассматриваемых признаков для *i*-го респондента, $\xi_i(1)$ и $\xi_i(2)$ — то, что *i*-й респондент хотел бы сообщить социологу, т.е. истинные значения, искаженные из-за ошибок респондента; $\eta_i(1)$ и $\eta_i(2)$ — полученные социологом ответы, «привязанные» к градациям шкалы. Можно интерпретировать дисперсии $D(\xi_i(1))$ и $D(\xi_i(2))$ как показатели «нечеткости» мнений респондентов (ср. [22]), их изменчивости при повторном опыте через небольшой промежуток времени [25, 26].

Пусть $\overline{\nu}_1, \overline{\xi}_1, \overline{\eta}_1, \overline{\alpha}_1, \overline{\beta}_1$ — средние арифметические соответствующих показателей для первого признака, а $\overline{\nu}_2, \overline{\xi}_2, \overline{\eta}_2, \overline{\alpha}_2, \overline{\beta}_2$ — для второго. Предположим, социолога интересует, что больше — $\overline{\nu}_1$ или $\overline{\nu}_2$. Для ответа на этот вопрос он сравнивает $\overline{\eta}_1$ и $\overline{\eta}_2$, причем принимает, что $\overline{\nu}_1 - \overline{\nu}_2$ имеет тот же знак, что и $\overline{\eta}_1 - \overline{\eta}_2$. Такая процедура может привести к ошибочному заключению о знаке $\overline{\nu}_1 - \overline{\nu}_2$.

Опять можно выделить ошибки респондента и ошибки исследовательского инструмента. Последние имеют место, если

$$\overline{\xi}_1 \neq \overline{\xi}_2, \quad (\overline{\xi}_1 - \overline{\xi}_2)(\overline{\eta}_1 - \overline{\eta}_2) \leq 0 .$$

Пусть для определенности $\overline{\nu}_1 - \overline{\nu}_2 = \nu > 0$. Описанная выше процедура приводит к ошибке, если $\overline{\eta}_1 - \overline{\eta}_2 \leq 0$, т.е.

$$\overline{\alpha}_2 - \overline{\alpha}_1 + \overline{\beta}_2 - \overline{\beta}_1 \geq \nu . \tag{5}$$

Чтобы найти вероятность осуществления события (5), построим вероятностную модель. Примем все предположения, описанные выше для одного признака, и добавим, что оценивание различных признаков проводится независимо. Тогда $\overline{\alpha}_1, \overline{\alpha}_2, \overline{\beta}_1, \overline{\beta}_2$ при больших n и малых h можно рассматривать как четыре независимые в совокупности нормальные случайные величины с нулевыми математическими ожиданиями. Тогда $(\overline{\alpha}_2 - \overline{\alpha}_1 + \overline{\beta}_2 - \overline{\beta}_1)$ — нормальная случайная величина с нулевым математическим ожиданием и дисперсией $D(\overline{\alpha}_1) + D(\overline{\alpha}_2) + D(\overline{\beta}_1) + D(\overline{\beta}_2)$, а потому вероятность осуществления события (5) есть

$$1 - \Phi\left(\frac{\nu}{\sqrt{D(\overline{\alpha}_1) + D(\overline{\alpha}_2) + D(\overline{\beta}_1) + D(\overline{\beta}_2)}} \right),$$

где $\Phi(z)$ — функция распределения стандартной нормальной случайной величины с математическим ожиданием 0 и дисперсией 1. Условие «однопорядковости» порождающих ошибку причин приводит к равенству

$$D(\overline{\alpha}_1) + D(\overline{\alpha}_2) = D(\overline{\beta}_1) + D(\overline{\beta}_2) . \tag{6}$$

В случае $D(\overline{\alpha}_1) = D(\overline{\alpha}_2)$ и $D(\overline{\beta}_1) = D(\overline{\beta}_2)$ условие (6) переходит в (2), а при одинаковости дисперсий респондентов — в (4).

Процесс получения случайной величины η по случайной величине ξ естественно назвать «квантованием» или, если угодно, «дискретизацией». Выше описаны способы применения квантования в социологии. Есть и другие возможности его применения, связанные, например, с управлением запасами (в конце настоящего раздела мы получим результаты, примыкающие к рассмотрениям раздела 8.3) или с округлением чисел при вычислениях на ЭВМ.

Предельная теорема о квантовании. Перейдем к изучению асимптотических свойств квантования при $h \to 0$. Нам понадобятся вспомогательные утверждения — две леммы.

Лемма 1. Пусть функции $f : R^1 \to R^1$ и $g : [0;+\infty) \to [0;+\infty)$ таковы, что $|f(x)| \le g(|x|)$ при всех x, функция g строго убывает, причем существует

$$\int_0^{+\infty} g(x)dx = B < +\infty .$$

Рассмотрим последовательность точек числовой прямой a_k таких, что $a_k \in \Delta_k = [kh;(k+1)h),\ k = 0,\pm 1,\pm 2,\ldots,h > 0$. Тогда

$$h\sum_k | f(a_k) | < 2B + 2g(0)h .$$

Доказательство. Пусть сначала $k \ge 1$. Тогда

$$|f(a_k)| \le g(a_k) \le g(kh) < g(x) \tag{7}$$

при всех $x \in \Delta_{k-1}$. Проинтегрировав (7) по Δ_{k-1}, имеем

$$| f(a_k) | h < \int_{\Delta_{k-1}} g(x)dx . \tag{8}$$

Из (8) следует оценка

$$h\sum_{k\ge 1} | f(a_k) | < \int_0^{+\infty} g(x)dx = B . \tag{9}$$

Аналогично можно получить неравенство

$$h\sum_{k\le -2} | f(a_k) | < B . \tag{10}$$

Наконец,

$$| f(a_{-1}) | + | f(a_0) | \le 2g(0) . \tag{11}$$

Из (9), (10) и (11) следует требуемое.

Лемма 2. Пусть

а) функция $f: R^1 \to R^1$ дважды непрерывно дифференцируема на R^1;

б) $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$;

в) $\lim\limits_{|x|\to\infty} f(x) = 0$;

г) существует строго убывающая интегрируемая функция $g:[0;+\infty)\to[0;+\infty)$ такая, что $|f''(x)| \le g(|x|)$ при всех $x \in R^1$.

Тогда при всех $0 \le y \le h$ справедливо неравенство

$$\left|\sum_{-\infty<k<+\infty} \int\limits_{kh}^{kh+y} f(x)dx - \frac{y}{h}\right| < 2{,}25\left(g(0)h + \int\limits_{0}^{+\infty} g(x)dx\right)h^2.$$

Доказательство. Разложим функцию $f(x)$ в точке a по формуле Тейлора с остаточным членом в форме Лагранжа. Пользуясь этим разложением на отрезке $[a;\ a+y)$, получим равенство

$$\int\limits_{a}^{a+y} f(x)dx - \frac{y}{h}\int\limits_{a}^{a+h} f(x)dx = f'(a)\left\{\frac{y^2}{2} - \frac{yh}{2}\right\} + \left\{K(a,y) - \frac{y}{h}K(a,h)\right\}, \qquad (12)$$

где функция $K(a, z)$ такова:

$$K(a,z) = \int\limits_{a}^{a+z} \frac{(x-a)^2}{2} f''(\theta a + (1-\theta)x)dx,\ 0 \le \theta = \theta(x;a;f) \le 1. \qquad (13)$$

Полагая в (12) $a = kh$, $k = 0, \pm1, \pm2, \ldots$, суммируя по k и пользуясь условием б) настоящей леммы, приходим к разложению

$$\begin{aligned}&\sum_{-\infty<k<+\infty} \int\limits_{kh}^{kh+y} f(x)dx - \frac{y}{h} = \left(\frac{y^2}{2} - \frac{yh}{2}\right)\sum_{-\infty<k<+\infty} f'(kh) + \\ &+ \sum_{-\infty<k<+\infty}\left(K(kh,y) - \frac{y}{h}K(kh,h)\right)\end{aligned}. \qquad (14)$$

Оценим каждую из сумм в правой части (14) по отдельности. Согласно (13) при $0 \le y \le h$ имеем

$$|K(kh,y)| \le \frac{h^2}{2}\max_{x\in\Delta_k}|f''(x)|\,h. \qquad (15)$$

Из (15), леммы 1 и условия г) настоящей леммы следует, что вторая сумма в (14) не превосходит

$$2h^2\left(g(0)h+\int_0^{+\infty}g(x)dx\right). \qquad (16)$$

С помощью формулы Тейлора и условия в) нетрудно показать, что

$$h\sum_{-\infty<k<+\infty}f'(kh)=-\sum_{-\infty<k<+\infty}L(kh,h), \qquad (17)$$

где

$$L(a,h)=\int_a^{a+h}(x-a)f''(\theta_1 a+(1-\theta_1 x)dx,\ 0\le\theta_1=\theta_1(x;a,f)\le 1.$$

Поскольку

$$|L(kh,h)|\le h^2\left(\max_{x\in\Delta_k}|f''(x)|\right),$$

то с помощью условия г) и леммы 1 заключаем, что

$$\sum_{-\infty<k<+\infty}|L(kh,h)|\le h\left(2hg(0)+2\int_0^{+\infty}g(x)dx\right). \qquad (18)$$

Из (14), (17), (18) и неравенства

$$|y^2-yh|\le\frac{h^2}{4}$$

следует, что первая сумма в (14) не превосходит

$$0{,}25\left(\int_0^{+\infty}g(x)dx+hg(0)\right)h^2. \qquad (19)$$

Из (16) и (19) вытекает заключение леммы 2. Итак, лемма 2 полностью доказана.

Теорема 1. Пусть плотность $f(x)$ случайной величины ξ удовлетворяет условиям леммы 2. Пусть η — ближайший к ξ элемент системы баллов $\{kh,\ k =$ $= 0, \pm 1, \pm 2, \ldots\}$, т.е. $\eta = \xi + h\alpha_0$, где $\eta \in \{kh,\ k = 0, \pm 1, \pm 2, \ldots\}$ и $-0{,}5 \le \alpha_0 < 0{,}5$. Тогда ξ и α_0 асимптотически независимы при $h \to 0$, т.е.

$$\lim_{h\to 0}\sup_{x,y} | P\{\xi < x, \alpha_0 < y\} - P\{\xi < x\}P\{\alpha_0 < y\} | = 0. \tag{20}$$

Доказательство. Пусть целая часть числа $\frac{x}{h} + 0{,}5$ равна z. Имеем

$$\begin{aligned} P\{\xi < x, \alpha_0 < y\} &= \sum_{-\infty < k \le z} P\left\{\left(k - \frac{1}{2}\right)h \le \xi < \left(k + \frac{1}{2}\right)h, \alpha_0 < y\right\} + \\ &+ P\left\{\left(z - \frac{1}{2}\right)h \le \xi < \min(x < hy + zh)\right\} \end{aligned}. \tag{21}$$

Незначительное изменение доказательства леммы 2 дает, что сумма в формуле (21) есть

$$\left((y + \frac{1}{2}\right)P\left\{\xi < \left(z - \frac{1}{2}\right)h\right\} + O(h). \tag{22}$$

Действительно,

$$\sum_{-\infty < k \le z} P\left\{\left(k - \frac{1}{2}\right)h \le \xi < \left(k + \frac{1}{2}\right)h, \alpha_0 < y\right\} = \sum_{-\infty < k \le z} \int_{(k-0,5)h}^{kh+yh} f(x)dx. \tag{23}$$

Оценка (16) для суммы в (23) сохраняется. Вместо (17) имеем

$$h\sum_{-\infty<k\le z} f'\left(\left(k - \frac{1}{2}\right)h\right) = f\left(\left(z - \frac{1}{2}\right)h\right) - \sum_{-\infty<k\le z} L\left(\left(k - \frac{1}{2}\right)h, h\right) \tag{24}$$

и соответственно вместо (19) —

$$f\left(\left(z - \frac{1}{2}\right)h\right)\frac{h}{8} + \frac{1}{4}\left(\int_0^{+\infty} g(x)dx + hg(0)\right)h^2, \tag{25}$$

откуда и следует (22), причем равномерно по y и z. При этом использовалось, что согласно условиям а) и в) леммы 2

$$\max_{-\infty<x<+\infty} f(x) < +\infty. \tag{26}$$

А именно, из (26) следует равномерность по z. Кроме того, из (26) вытекает, что

$$P\left(\left(z-\frac{1}{2}\right)h \le \xi < \min(x, hy+hz)\right) = O(h), \tag{27}$$

причем равномерно по x, y. С помощью (22) и (27) можно привести (21) к виду

$$P(\xi < x, \alpha_0 < y) = \left(y+\frac{1}{2}\right)P(\xi < x) + O(h) \tag{28}$$

При этом использовано то, что

$$P\{\xi < x\} - P\left\{\xi < \left(z-\frac{1}{2}\right)h\right\} = O(h).$$

Наконец, из леммы 2 следует, что

$$\sup_{-0,5\le y\le 0,5}\left|P(\alpha_0 < y) - \left(y+\frac{1}{2}\right)\right| = O(h^2). \tag{29}$$

Из (28), (29) и отмеченной выше равномерности оценок следует (20), причем можно добавить, что используемый в (20) супремум есть $O(h)$. Теорема 1 доказана.

Формула (29) представляет самостоятельный интерес: она показывает, что нормированная ошибка округления α_0 имеет при $h \to 0$ асимптотически равномерное распределение на отрезке [– 0,5; + 0,5], причем сходимость весьма быстрая. Этот факт будет использован ниже при применении полученных в настоящем разделе результатов в задачах управления запасами (логистике).

Теорема 2. Пусть плотность $f(x)$ случайной величины ξ удовлетворяет условиям а) — г) леммы 2, причем математическое ожидание $M(\xi)$ существует.

Тогда для определенной в теореме 1 нормированной ошибки округления α_0 имеем

$$\lim_{h \to 0} M(\xi \alpha_0) = 0 . \tag{30}$$

Доказательство. Теорема 1 показывает, что распределение случайного вектора (ξ, α_0) сходится к распределению случайного вектора (ξ, δ) при $h \to 0$, где δ — независимая от ξ случайная величина с равномерным распределением на отрезке [– 0,5; + 0,5].

Отсюда следует, что при любом фиксированном N

$$\lim_{h \to 0} M(\xi_N \alpha_0) = M(\xi_N \delta) , \tag{31}$$

где

$$\xi_N = \begin{cases} \xi, & |\xi| < N, \\ N, & |\xi| \geq N. \end{cases} \tag{32}$$

Для доказательства достаточно сослаться на двумерный аналог утверждения, известного как первая теорема Хелли [27 с. 344–346], вторая теорема Хелли [28, с. 174–175], лемма Хелли — Брея [29, с. 193–194] (см. также [15, раздел 7.3]).

Из независимости ξ и δ следует, что ξ_N и δ также независимы. Поскольку $M(\delta) = 0$, то $M(\xi_N \delta) = 0$.

Для завершения доказательства достаточно показать, что

$$\lim_{h \to 0, N \to +\infty} M\{(\xi - \xi_N)\alpha_0\} = 0 . \tag{33}$$

Действительно, поскольку $|\eta \alpha_0| \leq \frac{1}{2} |\eta|$ при любом η, то

$$M\{(\xi - \xi_N)\alpha_0\} \leq \frac{1}{2} M(|\xi - \xi_N|) = \frac{1}{2} \int_{|x| \geq N} (|x| - N) dP(\xi < x) \leq \frac{1}{2} \int_{|x| \geq N} |x| dP(\xi < x) \tag{34}$$

Остается заметить, что последний член цепочки неравенств (34) стремится к 0 при $N \to +\infty$ в силу существования математического ожидания $M(\xi)$. Теорема 2 доказана.

Теорема 3. Пусть выполнены условия теоремы 2 и, кроме того, дисперсия $D(\xi) \neq 0$ существует. Тогда при $h \to 0$ коэффициент корреляции между ξ и α_0 стремится к 0.

Доказательство. В силу теоремы 1 (формула (29)) и цитированного в теореме 2 двумерного аналога теоремы Хелли

$$\lim_{h\to 0} M(\alpha_0) = 0, \; \lim_{h\to 0} D(\alpha_0) = \frac{1}{12} = D(\delta). \tag{35}$$

В силу теоремы 2

$$\lim_{h\to 0} M(\xi\alpha_0) = 0. \tag{36}$$

Из (35) и (36) следует, что

$$\lim_{h\to 0} M\{(\xi - M(\xi))(\alpha_0 - M(\alpha_0))\} = 0. \tag{37}$$

Для завершения доказательства теоремы 3 достаточно использовать (37) и еще раз сослаться на второе соотношение в (35) и на то, что $D(\xi)$ существует и не равна 0.

Доказаны все утверждения, необходимые для обоснования проведенных в начале раздела рассуждений, с помощью которых был получены формулы (3) и (4). Нет необходимости подробно расписывать применение многомерной центральной теоремы (раздел 14.2) и теорем о наследовании сходимости (раздел 14.3) к векторам $(\overline{\alpha}, \overline{\beta})$ и $(\overline{\alpha}_1, \overline{\alpha}_2, \overline{\beta}_1, \overline{\beta}_2)$лучения этих формул.

Рекомендации по выбору числа градаций в социологических анкетах. Пусть от респондента требуется оценить интенсивность какого-либо качества в процентах (т.е. рассматриваемые величины меняются в интервале от 0 до 100 %). С помощью формулы (4) можно найти требуемое число градаций *k*. А именно,

$$k \approx \frac{100}{h} \approx \frac{100}{3{,}47\sigma} = \frac{28{,}8}{\sigma}.$$

Поскольку *k* — натуральное число, можно рекомендовать к практическому использованию формулу

$$k = \left[\frac{28{,}8}{\sigma}\right] + 1, \tag{38}$$

где квадратными скобками обозначена операция взятия целой части числа. Так, при σ = 10 (%) в соответствии с (38) k =3, при σ =5 (%) соответственно k = 6. Серединными токами интервалов являются 17, 50, 83 % в первом случае и 8,3, 25 % и т.д. — во втором.

Дисперсию ошибки респондента можно оценить при пилотаже (пробном исследовании) ли найти иными способами, имеющимися в многочисленных работах по методике социологических опросов [25]. Для многих социологических данных [26] среднее квадратическое отклонение мнения респондента σ имеет порядок 5–10 % и более Аналогичный порядок точности для экономико-статистических данных приводит и Оскар Моргенштерн [30].

Наши результаты показывают, что вор многих социологических анкетах целесообразно использовать не более 3–6 градаций. Разумеется, эти формально-математические аргументы не могут служить единственным основанием при выборе числа градаций. Какова же их польза? Они могут служить ориентиром и, что, по нашему мнению, более важно, обосновывают корректность эмпирически применяемых мощностей множества градаций (в данном случае мощность конечного множества — это число его элементов). По данным [31], массивы данных в социологии на 92 % состоят из значений качественных признаков, что и объясняет больше значение статистики нечисловых данных для социологических исследований. Присоединяясь в целом к высказанному в этой работе [31] пожеланию шире использовать в социологии многомерный статистический анализ и другие методы, основанные на признаках, измеренных в количественных шкалах (интервалов, отношений и др.), необходимо подчеркнуть, что квантование шкалы является эмпирически найденным приемом, повышающим обоснованность выводов статистического анализа социологических данных путем повышения устойчивости к ошибкам респондентов. Выше дано теоретическое обоснование этому эмпирическому приему.

Теория асимптотического квантования была опубликована в конце 70-х гг. (см. прежде всего [23], [8, раздел 2.6]). Они использовались, например, в [32]. Ф.Н. Ильясов [33] провел их экспериментальную проверку, в результате от 13 градаций в исходной шкале перешел к шкале с 6 градациями.

Есть и иные способы определения необходимого числа градаций [34].

Применения в управлении запасами. Идея «квантования» имеет применения не только в социологии. Применять ее можно не только для выбора числа градаций. Так, весьма интересны два применения идеи «квантования» в теории управления запасами — в двухуровневой модели (полезны теоремы 1–3) и в классической модели Вильсона с учетом отклонений от нее (демонстрируется польза «квантования» как способа повышения устойчивости). В очередной раз видим взаимопроникновение статистических методов, возникших для анализа данных из различных предметных областей, в данном случае, из социологии и логистики. Имеем очередное доказательство того, что статистические методы — единая научно-практическая область, которую нецелесообразно делить по областям применения.

Введение в теорию управления запасами дано в разделе 8.3. Напомним, что в математических терминах двухуровневая однопродуктовая модель работы склада формулируется следующим образом.

Пусть $\{\xi_n, n=1,2,...\}$ — последовательность неотрицательных независимых одинаково распределенных случайных величин (последовательность величин заявок со стороны потребителей). Пусть $\tau(t)$ — неубывающий кусочно-постоянный случайный процесс на $[0;+\infty)$, независимый от последовательности $\{\xi_n\}$ и имеющий скачки величиной 1 в точках $\tau_1 < \tau_2 < ... < \tau_k < ...$, причем $\tau(0)=0$ (процесс $\tau(t)$ описывает моменты возникновения заявок со стороны потребителей). Положим

$$s(t) = \xi_1 + \xi_2 + ... + \xi_{\tau(t)}, \quad t \geq 0$$

(величина $s(t)$ показывает накопленный спрос за время t). Пусть R<0 и $R+Q$>0 — произвольные числа (уровни, упоминаемые в названии модели). Пусть

$$y(t) = \left[\frac{s(t)+R}{Q}\right]Q + Q - s(t),$$

где квадратными скобками обозначена целая часть числа. Здесь $y(t)$ — уровень запаса продукта на складе. Предполагается, что продукт отпускается в соответствии с заявками потребителей. Когда уровень запаса на складе опускается до граничного значения R<0, мгновенно поступает новая партия продукта величиной Q. Если уровень запаса на складе при удовлетворении очередной заявки становится меньше 0, но остается больше R<0, то фиксируется образование дефицита. Рассматриваемая заявка (точнее, ее необслуженная часть) и возникающие в следующие моменты заявки накапливаются вплоть до того момента, когда суммарный дефицит достигнет граничного уровня R, а затем мгновенно удовлетворяются после прихода очередной поставки величиной Q. Таким образом,

$$R < y(t) \leq R + Q$$

при всех t.

Пусть c_1, c_2, c_3 — положительные константы, $T > 0$. Введем функционал

$$w(T) = \frac{1}{T}\left\{c_1\int_0^T y^+(t)dt + c_2\int_0^T y^-(t)dt + c_3\left[\left[\frac{s(T)+R}{Q}\right]\right]\right\},$$

где $y^+ = \max(y,0)$ и $y^- = y^+ - y$. Требуется минимизировать $M(w(T))$, выбрав соответствующие R и Q.

В логистических терминах c_1 — расходы, связанные с хранением единицы продукта в течение единицы времени, c_2 — потери, порожденные дефицитом (отсутствием на складе) единицы продукции в течение единицы времени, c_3 — расходы на доставку партии продукта на склад, $w(T)$ — средние (на единицу времени) издержки на функционирование склада (точнее, складской подсистемы, связанной со снабжением потребителей рассматриваемым продуктом). Сформулированная модель рассматривалась рядом авторов [35, 36, 37, 38, 39, 40] при различных предположениях. У них же можно найти экономическую интерпретацию и развернутое обсуждение рассматриваемых математических объектов и констант (см. также раздел 8.3).

Будем изучать асимптотику при $T \to \infty$. Сначала найдем предельное распределение $y(t)$.

Теорема 4. Пусть случайная величина ξ_1 имеет первые пять моментов и интегрируемую характеристическую функцию, при некотором C

$$P(\tau(t) = 0) < CM\left(\frac{1}{\tau(t)+1}\right)$$

при всех $t \geq 0$. Тогда существует константа D такая, что

$$\left|P(y(t) \leq y) - \frac{y-R}{Q}\right| < DM\left(\frac{1}{\tau(t)+1}\right)$$

при всех y, t таких, что $R \leq y \leq R + Q$ и $t \geq 0$.

Доказательство. Нетрудно показать, что

$$P(y(t) \leq y) = \sum_{n=0}^{+\infty} P(nQ - y \leq s(t) < nQ - R). \tag{39}$$

Нам окажется полезным вытекающее из (39) представление

$$P(y(t) \le y) = \sum_{k=0}^{+\infty} A(k, y) P(\tau(t) = k), \tag{39-1}$$

где

$$A(k, y) = \sum_{n=0}^{+\infty} P(nQ - y \le s(t) < nQ - R \mid \tau(t) = k).$$

При $k \ge 1$ положим

$$T_k = \frac{\xi_1 + \xi_2 + \ldots + \xi_k - kM(\xi_1)}{\sqrt{k}\sqrt{D(\xi_1)}}.$$

Тогда

$$A(k, y) = \sum_{n=0}^{+\infty} P\left(\frac{nQ - y - kM(\xi_1)}{\sqrt{k}\sqrt{D(\xi_1)}} \le T_k < \frac{nQ - R - kM(\xi_1)}{\sqrt{k}\sqrt{D(\xi_1)}} \right).$$

Начнем с оценки $A(k, y)$.

Лемма 3. В условиях теоремы 4 существует константа d_1 такая, что при всех y таких, что $R \le y \le R + Q$ и $k = 1, 2, \ldots$, выполнено неравенство

$$\left| A(k, y) - \frac{y - R}{Q} \right| < \frac{d_1}{k}.$$

Доказательство. Пусть $f_k(x)$ — плотность случайной величины T_k, существующая в силу условий теоремы 4. Из них же в соответствии с теоремой 2 [41, с. 611] вытекает, что

$$f_k(x) - g_k(x) = \gamma_k(x) = O(k^{-3/2}) \tag{40}$$

равномерно по x. Здесь

$$g_k(x) = \eta(x)\left(1 + \frac{\alpha}{\sqrt{k}}(x^3 - 3x) + \frac{\beta}{k}(x^4 - 6x^2 + 3)\right), \tag{41}$$

где $\eta(x)$ — плотность стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1, α и β — константы, зависящие от моментов случайной величины ξ_1.

Для $A(k, y)$ справедливо представление

$$A(k,y)=\sum_{n=0}^{+\infty}\int_{a(n)}^{a(n)+b}g_k(x)dx+\sum_{n=0}^{+\infty}\int_{a(n)}^{a(n)+b}\gamma_k(x)dx\,, \tag{42}$$

где

$$a(n)=n\frac{Q}{\sqrt{k}\sqrt{D(\xi_1)}}-\frac{y+kM(\xi_1)}{\sqrt{k}\sqrt{D(\xi_1)}},\quad b=\frac{y-R}{\sqrt{k}\sqrt{D(\xi_1)}}\,. \tag{43}$$

К первому слагаемому в (42) применим лемму 2, второе слагаемое оценим с помощью (40).

Пусть m — такое целое число, что $a(m)\le 0$, но $a(m+1)>0$. Легко проверить, что для функции $f(x)=g_k(x-a(m))$ выполнены все условия леммы 2. Полагая (слева — обозначения леммы 2, справа — обозначения из (43))

$$h=\frac{Q}{\sqrt{k}\sqrt{D(\xi_1)}},\ \ y=b\,,$$

заключаем, что при некоторой константе d_2, зависящей, как видно из (41) и (43), лишь от Q и моментов ξ_1, и произвольном k

$$\left|\sum_{-\infty<n<+\infty}\int_{a(n)}^{a(n)+b}g_k(x)dx-\frac{y-R}{Q}\right|<\frac{d_2}{k}\,. \tag{44}$$

Отметим теперь, что

$$\left|\sum_{n<0}\int_{a(n)}^{a(n)+b}g_k(x)dx\right|\le\int_{-\infty}^{c(k)}|\,g_k(x)\,|\,dx\,,$$

где

$$c(k)=-\sqrt{k}\,\frac{M(\xi_1)}{\sqrt{D(\xi_1)}}-\frac{y}{\sqrt{k}\sqrt{D(\xi_1)}}\,.$$

Из (41) нетрудно вывести, что

$$\int_{-\infty}^{c(k)} | g_k(x) | \, dx = O\left(\frac{1}{k}\right). \tag{45}$$

Из (44) и (45) следует оценка

$$\left| \sum_{n=0}^{+\infty} \int_{a(n)}^{a(n)+b} g_k(x) dx - \frac{y-R}{Q} \right| = O\left(\frac{1}{k}\right). \tag{46}$$

Займемся теперь второй суммой в (42). Положим

$$N_1 = \min\{n : a(n) > -\sqrt{k}\}, \; N_2 = \max\{n : a(n) + b < \sqrt{k}\}.$$

Тогда в силу (40)

$$\left| \sum_{N_1 \le n \le N_2} \int_{a(n)}^{a(n)+b} \gamma_k(x) dx \le \int_{-\sqrt{k}}^{\sqrt{k}} | \gamma_k(x) | \, dx \right| = O\left(\frac{1}{k}\right). \tag{47}$$

Справедливо неравенство

$$\left| \sum_{n=0}^{N_1-1} \int_{a(n)}^{a(n)+b} \gamma_k(x) dx + \sum_{n=N_2+1}^{+\infty} \int_{a(n)}^{a(n)+b} \gamma_k(x) dx \le \int_{B(k)} f_k(x) dx + \int_{B(k)} | g_k(x) | \, dx \right|, \tag{48}$$

где $B(k) = (-\infty; -\sqrt{k} + b) \bigcup (\sqrt{k} - b; +\infty)$. Первый интеграл в правой части (48) равен $P\{| T_k | \ge \sqrt{k} - b\}$. С помощью неравенства Чебышева и (43) заключаем, что

$$P\{| T_k | \ge \sqrt{k} - b\} \le \frac{1}{(\sqrt{k} - b)^2} = O\left(\frac{1}{k}\right). \tag{49}$$

Из явного вида $g_k(x)$, по аналогии с (45), заключаем, что второй интеграл в правой части (48) также есть $O\left(\frac{1}{k}\right)$. Из (47), (48), (49) и последнего утверждения следует, что

$$\sum_{n=0}^{+\infty} \int_{a(n)}^{a(n)+b} \gamma_k(x) dx = O\left(\frac{1}{k}\right). \tag{50}$$

Из (46) и (50) вытекает заключение леммы 3, поскольку, как нетрудно проверить, просматривая проведенные выше рассуждения, все оценки являются равномерными по y. Лемма 3 полностью доказана.

Продолжение доказательства теоремы 4. Из (39-1) и заключения леммы 3 следует, что

$$P(y(t)\le y)=\frac{y-R}{Q}+P(\tau(t)=0)\left(A(0,y)-\frac{y-R}{Q}\right)+d_1\sum_{k\ge1}\frac{\theta_k}{k}P(\tau(t)=k)\,,$$

где $|\theta_k|<1,\ k=1,2,\ldots$ Поскольку

$$\left|\sum_{k\ge1}\frac{\theta_k}{k}P(\tau(t)=k)\right|\le\sum_{k\ge1}\frac{2}{k+1}P(\tau(t)=k)\le 2M\left(\frac{1}{\tau(t)+1}\right),$$

то заключение теоремы 4 справедливо при $D = 2d_1 + C$. Теорема 4 доказана.

Теорема 5. В условиях теоремы 4 и при $M(\tau(T))<+\infty$ имеем для минимизируемого функционала

$$M(w(T))=H(R,Q)+\delta_T(R,Q)\,,$$

где

$$H(R,Q)=(c_1+c_2)\frac{R^2}{2Q}+c_1R+\frac{c_1Q}{2}+\frac{c_3M(\tau(T))M(\xi_1)}{QT}\,, \qquad (51)$$

причем существуют константы d_3 и d_4 такие, что

$$|\delta_T(R,Q)|\le\frac{d_3}{T}+\frac{d_4}{T}\int_0^T M\left(\frac{1}{\tau(t)+1}\right)dt\,. \qquad (52)$$

Доказательство. Поскольку накопленный спрос есть сумма случайного числа случайных слагаемых, то по тождеству Вальда $M(s(T))=M(\xi_1)M(\tau(T))$. Следовательно, для математического ожидания третьего слагаемого средних издержек $w(T)$ имеем

$$M\left(c_3\left[\frac{s(T)+R}{Q}\right]\right)=\frac{c_3M(\tau(T))M(\xi_1)}{QT}+\frac{c_3}{T}\omega\,, \qquad (53)$$

где

$$\frac{R}{Q}-1<\omega\le\frac{R}{Q}. \tag{54}$$

При изучении математического ожидания первого слагаемого средних издержек $w(T)$ полезно проинтегрировать по частям:

$$M(y^{+}(t))=yP(y(t)\le y)\,|_{0}^{R+Q}-\int_{0}^{R+Q}P(y(t)\le y)dy\,. \tag{55}$$

Из (55) и заключения теоремы 4 следует, что

$$M(y^{+}(t))=R+\frac{Q}{2}+\frac{R^2}{2Q}+\lambda\,, \tag{56}$$

где

$$|\,\lambda\,|\le(R+Q)DM\left(\frac{1}{\tau(t)+1}\right). \tag{57}$$

Аналогично получаем асимптотическое выражение для математического ожидания второго слагаемого в формуле для $w(T)$:

$$M(y^{-}(t))=\frac{R^2}{2Q}+\lambda'\,, \tag{58}$$

где

$$|\,\lambda'\,|\le|\,R\,|\,DM\left(\frac{1}{\tau(t)+1}\right). \tag{59}$$

Из (53)–(59) следует утверждение теоремы 5, при этом можно считать, что в (52)

$$d_3=c_3\left(\frac{|\,R\,|}{Q}+1\right),\;\; d_4=c_1(R+Q)D+c_2\,|\,R\,|\,D\,. \tag{60}$$

Теорема 5 доказана.

Нетрудно показать, что $H(R,Q)$ из теоремы 5 (см. (51)) достигает минимума по области $\{(R, Q): Q > 0\}$ в точке (R_0, Q_0), где

$$R_0 = -\sqrt{\frac{2c_1c_3M(\xi_1)M(\tau(T))}{(c_1+c_2)c_2T}}, \quad Q_0 = \sqrt{\frac{2(c_1+c_2)c_3M(\xi_1)M(\tau(T))}{c_1c_2T}}. \tag{61}$$

Теорема 6. Пусть выполнены условия теоремы 5 и, кроме того,

$$\overline{\lim_{T\to+\infty}} \frac{M(\tau(T))}{T} < +\infty. \tag{62}$$

Предположим, что минимум $M(w(T))$ достигается, причем в точке (R_T, Q_T). Тогда существует константа d_5 такая, что при всех T

$$|R_T - R_0| + |Q_T - Q_0| < d_5\sqrt{\frac{1}{T} + \frac{1}{T}\int_0^T M\left(\frac{1}{\tau(T)+1}\right)dt}.$$

Доказательство. Нетрудно показать, что в силу (62) минимум $M(w(T))$ не может достигаться вне некоторого компакта. Далее, анализируя доказательства теорем 4 и 5, можно убедиться, что полученные в них оценки являются равномерными по произвольному компакту в области $\{(R, Q): Q > 0\}$. Поскольку

$$H(R_T,Q_T) \le H(R_0,Q_0) + \delta_T(R_0,Q_0) - \delta_T(R_T,Q_T) = H(R_0,Q_0) + \varepsilon,$$

то (R_T, Q_T) лежит в области $\{(R,Q): H(R,Q) \le H(R_0,Q_0) + \varepsilon\}$. Эта область является эллипсом, и, как легко видеть [42],

$$R_T - R_0 = O(\sqrt{\varepsilon}),\ Q_T - Q_0 = O(\sqrt{\varepsilon}). \tag{63}$$

Из (52), (60) и (63) следует требуемое. Теорема 6 доказана.

Следствие. Пусть выполнены условия теоремы 6 и, кроме того,

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T M\left(\frac{1}{\tau(t)+1}\right)dt = 0. \tag{64}$$

Тогда

$$\lim_{T\to+\infty} (|R_T - R_0| + |Q_T - Q_0|) = 0$$

и

$$\lim_{T\to+\infty} (M(w(t,R_T,Q_T)) - M(w(t,R_0,Q_0))) = 0,$$

где величину $M(w(T, R_T, Q_T))$ рассчитывают по формулам для математического ожидания средних издержек работы склада $M(w(T))$ при параметрах плана поставок $R = R_T$, $Q = Q_T$, а величину $M(w(T, R_0, Q_0))$ — по тем же формулам при $R = R_0$, $Q = Q_0$.

Таким образом, следствие указывает условия, при которых при большом горизонте планирования T возможно использовать асимптотически оптимальный план поставок с параметрами $R = R_0$, $Q = Q_0$, заданными формулами (61).

Отметим, что для справедливости соотношения (64) необходимо и достаточно, чтобы

$$\lim_{t\to+\infty} M\left(\frac{1}{\tau(t)+1}\right) = 0,$$

но недостаточно выполнения условия

$$\lim_{t\to+\infty} M(\tau(t)) = +\infty.$$

В работах [35, 39] рассматривался пуассоновский поток $\tau(t)$ заявок от потребителей. Тогда накопленный спрос $s(t)$ асимптотически (при $t \to +\infty$) нормален. В этих работах предлагалось вероятности в правой части (39) заменить на предельные и полученным рядом аппроксимировать функцию распределения $P(y(t) \le y)$ при больших t. Чтобы сделать расчеты возможными, т.е. из эвристических побуждений, функция Лапласа в этом ряду заменялась [35] на ступенчатую функцию с единственным скачком в 0 единичной высоты. С помощью результатов настоящего раздела нетрудно подсчитать, что такая замена приводит к появлению лишних членов

$$2c_1\frac{\nu}{\mu}(R+Q) + 2c_2\frac{\nu}{\mu}R$$

в аналогичном $H(R, Q)$ выражении (обозначения [35]). Таким образом, мы отказались от предположения пуассоновости, которое редко выполняется на практике, и получили в общем случае достаточно простые формулы, попутно исправив ошибки предшественников.

Отказ от пуассоновости спроса впервые осуществлен нами в [36]. Дальнейшая работа проводилась совместно с А.В. Воскресенским [42, 43]. Настоящее изложение соответствует работе [44], в которой предшествующие результаты были усилены.

Классическая модель Вильсона управления запасами разобрана в разделе 8.3. Там показано, что квантование используемых экономических показателей позволяет исключить расхождение между собой методик, разработанных различными коллективами специалистов, но при этом обеспечивает сокращение издержек складского хозяйства не менее чем в 2 раза.

13.4. Социометрическое исследование — инструмент менеджера

В любой фирме, на любом предприятии в дополнение к официальным организационным структурам создаются неформальные, основанные на отношениях между людьми. Менеджеру необходимо учитывать с своей работе неформальные связи. Выявить их можно с помощью социометрии [45, 46].

Социометрическая техника применяется для диагностики межличностных и межгрупповых отношений в целях их изменения, улучшения и совершенствования. С помощью социометрии можно изучать типологию социального поведения людей в условиях групповой деятельности, судить о социально-психологической совместимости членов конкретных групп.

Вместе с официальной или формальной структурой общения, отражающей рациональную, нормативную, обязательную сторону человеческих взаимоотношений, в любой социальной группе всегда имеется психологическая структура неофициального или неформального порядка, формирующаяся как система межличностных отношений, симпатий и антипатий. Особенности такой структуры во многом зависят от ценностных ориентаций участников, их восприятия и понимания друг друга, взаимооценок и самооценок. Неформальная структура группы зависит от формальной структуры в той степени, в которой индивиды подчиняют свое поведение целям и задачам совместной деятельности, правилам ролевого взаимодействия. С помощью социометрии можно оценить это влияние.

Общая схема действий при социометрическом исследовании заключается в следующем. После постановки задач исследования и выбора объектов измерений формулируются основные гипотезы и положения, касающиеся возможных критериев опроса членов групп. Здесь не может быть полной анонимности, иначе социометрия окажется малоэффективной. Требования экспериментатора раскрыть свои симпатии нередко вызывает внутренние затруднения у опрашиваемых и проявляется у некоторых людей в нежелании участвовать в опросе. Поэтому для проведения социометрического исследования целесообразно привлекать постороннюю специализированную организацию

Когда вопросы или критерии социометрии выбраны, они заносятся на специальную карточку или предлагаются в устном виде по типу интервью. Каждый член группы отвечает на них, выбирая тех или иных членов группы в зависимости от большей склонности, предпочтительности их по сравнению с другими, симпатий, доверия и т.д.

При этом социометрическая процедура может проводиться в двух формах. Первый вариант — непараметрическая процедура. В данном случае испытуемому предлагается ответить на вопросы социометрической карточки без ограничения выборов испытуемого. Второй вариант — параметрическая процедура с ограничением числа выборов. Испытуемым предлагают выбрать строго фиксированное число из всех членов группы. Социометрическая карточка или социометрическая анкета составляется на заключительном этапе разработки программы. В ней каждый член группы должен указать свое отношение к другим членам группы по выделенным критериям. Критерии определяются в зависимости от программы данного исследования. Когда социометрические карточки заполнены и собраны, начинается этап их математической обработки.

Вначале следует построить простейшую социоматрицу взаимных выборов. Анализ социоматрицы по каждому критерию дает достаточно наглядную картину взаимоотношений в группе. Могут быть построены суммарные социоматрицы, дающие картину выборов по нескольким критериям. Основное достоинство социоматрицы — возможность представить выборы в числовом виде. Проведение опроса и представление первичной информации является предварительным этапом для дальнейшего анализа социометрических данных.

Рассмотрим применение социометрической технологии на примере изучения неформальных соотношений в студенческой группе. Прикладная значимость работы определялась тем, что из студентов этой группы были сформированы две подрядные бригады, выполнявшие задания Института высоких статистических технологий и эконометрики по изучению динамики потребительских цен [5].

Респондентам были заданы вопросы (они предлагались в устной форме, без ограничения числа выборов):

1. У кого вы берете конспекты лекций?

2. С кем вы сидите на семинаре?

3. С кем вы советуетесь при выполнении контрольных работ, курсовых работ и т.п.?

4. С кем вы часто разговариваете по телефону?

5. У кого вы узнаете об учебной информации: об изменениях в расписании, о собрании группы и т.п.?

6. С кем вы больше всего общаетесь: а) во время занятий; б) вне института?

7. С кем вы обычно ходите обедать?

8. Если вы что-то не поняли на семинаре, то у кого просите объяснение?

9. Если бы вся ваша группа была членами одной фирмы, и перед вами встала задача выбрать руководителя, то кого бы вы предложили на эту должность?

10. Если бы у всех ваших одногруппников были свои фирмы, а у вас нет, то к кому из них вы пошли бы работать?

11. Если бы вы все были членами одной фирмы, как вы думаете, кого группа выбрала бы президентом?

12. У кого из вашей группы есть, по вашему мнению, талант руководителя (организатора)?

13. После окончания института, при устройстве на работу вы узнаете, что там уже работает ваш одногруппник. Кого вы хотели бы увидеть?

13-1) проранжируйте всех своих одногруппников от 1 до 22 (от самого привлекательного до самого антипатичного);

13-2) напротив каждой фамилии укажите одну из букв а, б, в, г, д:

а) очень хотел бы увидеть; б) скорее «да», чем «нет»; в) трудно сказать; г) скорее «нет», чем «да»; д) не хотел бы увидеть.

Социометрические вопросы были разбиты на три группы: первая группа (условное название «информационная»): сюда входят 1, 3, 5, 8 вопросы; вторая группа («неформальная»): 2, 4, 6а, 6б, 7 вопросы; третья группа («трудовая»): 9–13 вопросы.

Социометрический опрос позволяет выявить неформальных лидеров в коллективе. Именно с ними менеджеру следует работать в первую очередь, проводя то или иное изменение (см. главу 1.4 в [19]). С точки зрения социометрии неформальный лидер — это человек, которого выбирает большинство респондентов.

На основании полученных для конкретной студенческой группы результатов были сделаны следующие выводы:

1. В исследуемой группе есть два крупных четко выраженных кластера (подгруппы) и ряд мелких.

2. Первая бригада построена по принципу взаимной заинтересованности в совместной работе. Вторая бригада построена по принципу антипатии к руководителю первой бригады и внутри группы — кто с кем хотел бы работать.

3. В группе два социометрического лидера. Оба в первой бригаде. Одна из них — староста группы (т.е. формальный лидер).

4. Руководители бригад не являются социометрическими лидерами.

5. Если заменить формального лидера (условный номер 18 в списке респондентов) на одного из неформальных (условные номера 4 и 9), то получим следующие результаты. При замене 18 на 4 состав бригады не изменится, принцип объединения в группу остается прежним, количество прямых связей увеличивается (следовательно, такая замена была бы возможной с целью улучшения работы бригады). Если же заменить 18 на 9, то состав бригады изменится.

Дальнейшее развитие событий подтвердило результаты, полученные с помощью социометрического опроса. Через некоторое время коллектив первой бригады потребовал смены бригадира, новым стал социометрический лидер 4. Бригада продолжала работать в прежнем составе. При создании второй бригады в ней объединились студенты, стоявшие в оппозиции к основной части группы. Это диссидентство (инакомыслие) в дальнейшем породило трудности в управлении бригадой. Социометрический опрос принес вполне ощутимую пользу менеджерам Института высоких статистических технологий и эконометрики.

Методики социометрии представляют собой синтез выборочных исследований и экспертных оценок. Обратите внимание, что ответы членов малой группы являются нечисловыми данными — подмножествами конечного множества, ранжировками, градациями признака, измеренного в порядковой шкале.

13.5. Статистические методы в выборочных исследованиях научных организаций

Процессы, происходящие в отечественной науке, требуют новых подходов как к методам сбора данных о состоянии науки (инфраструктура, кадры, финансирование, материально-техническая база, результативность и т.д.), так и к методам анализа полученных данных. Разработка новых подходов проводится Институтом высоких статистических технологий и эконометрики совместно с Центром исследований и статистики науки (ЦИСН) Миннауки РФ и РАН [18]. За 1990-е гг. ЦИСН разработан ряд новых форм статистической отчетности, утвержденных государственными статистическими органами, проведено несколько специальных обследований (бюджетное финансирование науки, уровень жизни ученых, утечка умов — совместно с МВД РФ, деятельность аспирантуры и др.).

В качестве иллюстрации процессов, происходящих в последние годы в науке России, можно привести несколько чисел из статистического сборника [47], изданного ЦИСН: с 1990 по 1993 гг. численность работников основной деятельности научных организаций сократилась на 40 %, расходы на НИР (в ценах 1989 г.) — на 75 %, при этом, если в 1990 г. расходы на НИР составляли 2.03 % ВВП (валового внутреннего продукта), то в 1993 г. — только 0.81 %.

Сохранению потенциала российской науки уделяется большое внимание международной научной общественностью, международными организациями, в частности, Организацией экономического сотрудничества и развития (см. [48]), и научной общественностью отдельных стран. По некоторым оценкам зарубежная помощь российским ученым в 1992 г. составила около 8 % от объема государственных расходов на науку в России.

Кризис российской науки оживленно обсуждается и на страницах отечественных газет и журналов. Так, в № 1 за 1993 г. международной газеты «Наука и технология в России» была проведена дискуссия об эффективности научной деятельности и положении научных работников в России [49, 50], глубоко и по существу использующая результаты социологических исследований.

При обсуждении разнообразных задач изучения и управления наукой весьма важны исходные методологические принципы. Показательна упомянутая дискуссия, в которой один из авторов в качестве основного показателя использовал производительность труда научного работника [49], а другой — фондоемкость научной продукции [50], что явно затрудняло взаимопонимание. Характерно, что в обеих статьях широко использовались как статистические, так и экспертные данные. Поэтому бесспорным представляется вывод о том, что статистические данные о научном потенциале — база для теоретического и прикладного науковедения.

Основой для проведения анализа научного потенциала России является информационная база данных по российским организациям, выполняющим научно-исследовательские работы, сформированная в Центре исследований и статистики науки Миннауки РФ и РАН. Она основана на использовании совокупности независимых баз данных как справочного, так и информационного характера и включает около 250 показателей научно-технической деятельности предприятия (организации). Благодаря подробной справочной информации, все объекты могут быть классифицированы по регионам, типам организации, ведомственной принадлежности, областям наук, формам собственности, секторам деятельности. Единицей учета является отдельный документ «Отчет предприятия (организации) о выполнении научно-технических работ» — статистическая форма № 1-наука.

В числе показателей формы присутствуют как количественные, так и качественные. К качественным показателям, имеющим нечисловую природу, относятся:

- область науки;
- сектор науки;
- отрасль народного хозяйства;
- территория;
- министерская (ведомственная) принадлежность;
- организационно-правовая форма;
- форма собственности;
- тип организации (предприятия);
- сектор деятельности.

Среди количественных показателей можно выделить данные об объемах выполненных исследований (фундаментальных, прикладных разработок), затратах на НИР, объемах основных средств, количестве персонала различных категорий и др.

База для применения методов статистики нечисловых данных в наукометрии значительно расширилась с началом проведения регулярных статистических обследований по форме № 2-наука , которая, в отличие от формы № 1-наука, предусматривает возможность группировок данных по социально-экономическим целям исследований и разработок (в соответствии с международными стандартами), что расширяет возможности как для построения принципиально новых моделей, так и для проведения более качественных межстрановых сопоставлений. Была введена также краткая ежеквартальная форма статистической отчетности, позволяющая накапливать ряды данных и строить на их основе прогнозные модели. Впервые в стране с 1995 г. начался сбор данных по инновационной деятельности, что позволяет увязывать данные о развитии науки с разработкой и внедрением технологий и более точно оценивать результативность науки.

С 1970-х гг. в России активно развивается статистика объектов нечисловой природы, известная также как статистика нечисловых данных или нечисловая статистика. На основе подходов и результатов статистики объектов нечисловой природы разрабатывается методическое, математическое и программное обеспечение для социологических, маркетинговых, экспертных, прогнозных и других исследований. В разработке этого сравнительно нового направления статистических методов приоритет принадлежит отечественным ученым.

Показатели науки могут быть использованы для применения достижений статистики нечисловых данных на практике. Более того, представляется, что при анализе столь важной с практической точки зрения информации, как данные о научном потенциале, именно методы статистики нечисловых данных окажутся наиболее полезными, поскольку существенная часть данных носит нечисловой (в частности, качественный) характер.

Примерами объектов нечисловой природы являются: качественные признаки, например, пол человека или тип научной организации, вообще результат отнесения объекта в одну из заданных категорий (градаций); множество, например, совокупность научных организаций, занимающихся определенной тематикой; слово, предложение, текст, которые в памяти компьютера кодируются, как известно, с помощью цифр 0 и 1, но числами от этого не становятся; вектор, координаты которого — совокупность значений разнотипных признаков, например, результат составления статистического отчета о научно-технической деятельности (например, форма № 1-наука), в котором часть признаков носит качественный характер, а часть — количественный; ответы на вопросы социологической анкеты, часть из которых носит количественный характер (возможно, интервальный, т.е. респондент свое мнение выражает не числом, а интервалом), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты; упорядочение экспертом заявок на проведение научных работ при проведении конкурсов на выделение грантов; результаты контроля выполнения заданий по научно-техническим программам, вообще планов научных работ, по альтернативному признаку (выполнена или не выполнена конкретная позиция плана); разбиения объектов на группы сходных между собой (кластеры); ранжировки, например, упорядочения экспертами научных проектов по степени предпочтения (на одной из стадий процесса распределения грантов); толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходство тематики научных работ, оцениваемого экспертами с целью рационального формирования экспертных советов внутри определенной области науки; результаты парных сравнений и т.д. Интервальные данные тоже можно рассматривать как пример объектов нечисловой природы.

На развитие статистики нечисловых данных большое стимулирующее влияние оказывают запросы прикладных исследований, прежде всего в социально-экономической области. Так, многолетний (с 1970 г.) опыт проведения социологических и маркетинговых исследований, а в последнее время — анализа и прогнозирования потребительских цен и индекса инфляции, исследований рынка товаров народного потребления, образовательных услуг, программного обеспечения — привел к постановкам ряда нерешенных задач в области статистических методов анализа и прогнозирования на основе нечисловых данных.

Однако приходится с сожалением констатировать, что ряд теоретических разработок пока не доведен до уровня методик, математического и программного обеспечения. Этим мы занимались, в частности, в рамках работы по алгоритмизации и применению основных идей, подходов и методов рассматриваемого научного направления для анализа статистических данных о научных организациях России. На основе современных представлений о проведении социологических исследований и о методах анализа собранных данных разрабатывались соответствующие статистические методы, математическое и программное обеспечение. Поскольку нечисловые данные составляют до 90 % данных в социологии и большую часть — в экономике, то теоретические исследования в статистике нечисловых данных позволяют получать новые результаты в той центральной области эконометрики, в которой отечественные работы имеют приоритет на мировом уровне.

Для анализа данных о научных организациях России, об их научном потенциале весьма важны методы классификации, в том числе типологии. Проблемами теории и практики классификации в нашей стране занимались многие научные работники. С 1984 г. в рамках Союза научных и инженерных обществ действует «Комиссия по классификации», объединившая усилия нескольких десятков специалистов различных научных областей. Итоги десяти лет работы нашей комиссии показали, в частности, что наиболее естественно ставить и решать задачи классификации в рамках статистики объектов нечисловой природы. Это касается как распознавания образов с учителем (дискриминантного анализа), так и распознавания образов без учителя (кластерного анализа).

Примеры возможных применений современных статистических методов анализа выборочных данных в задачах изучения и управления научным потенциалом. В силу целого ряда причин (высокая динамика изменений, потребность в оперативной информации для принятия решений и т.д.) возникает острая необходимость регулярного проведения оперативных обследований научных организаций и социологических опросов ученых. Эта задача может быть успешно решена только путем проведения выборочных обследований [51]. Обсудим несколько примеров возможных применений современных статистических методов анализа выборочных данных в задачах изучения и управления научным потенциалом.

1. Изучение показателей научной деятельности, в частности, изучение предпочтений научных работников, руководителей научных организаций

и профессионалов-управленцев, экспертно-статистический подход к построению интегральных показателей (рейтингов).

Статистический и экспертный анализ показателей науки, выявление их рейтинга (какие показатели являются наиболее значимыми с учетом общепринятых международных стандартов в статистике науки [52]).

Как оценивать эффективность науки? В какой мере общепринятые показатели результативности науки (публикации, патенты, лицензии, индексы цитируемости и т.д.) отражают реальную эффективность науки?

2. Кластеризация научных организаций, выделение типов. При обсуждении вопросов финансирования, организационных преобразований, перспективного развития необходимо к различным типам научных организаций подходить дифференцированно. Насколько принятая типология научных организаций (см. [47, стр. 228]) соответствует тем или иным аналитическим задачам? Сколько типов существует реально? Какова естественная типология научных организаций? Если такая типология излишне сложна, не удастся ли ее упростить, рассматривая научные организации с определенных практических позиций — с точки зрения научного уровня, отдачи, финансирования, перспективности и выживания и т.д.

Представляет также интерес структура научных советов, комиссий и др. форм деятельности ученых, особенно в ситуации, когда через такие структуры идет финансирование.

3. Рейтинги проектов, научных организаций и др. В рамках достаточно однородной совокупности проектов НИР, заявок на гранты, научных организаций и т.д. с помощью методов многомерного статистического анализа можно выявить основные направления вариации. Однако главный фактор, вопреки распространенному способу интерпретации результатов факторного анализа (или метода главных компонент), не всегда соответствует оси «эффективность — неэффективность». Тем не менее, идея рейтинга (интегрального показателя качества) заслуживает проработки. Можно указать несколько подходов. Для решения этой задачи можно использовать три варианта экспертно-статистического метода интегральной оценки перспективности научного направления или научной организации:

- по интегральному показателю (линейная функция, параметры которой — показатели науки, а значения коэффициентов получаются от экспертов);
- по обучающим выборкам, полученным от квалифицированных экспертов;
- с помощью оценки параметров интегрального показателя по обучающим выборкам (собственно экспертно-статистический метод).

4. Расчет и краткосрочное прогнозирование показателей науки. Важнейшее значение для принятия управленческих решений в сфере науки имеет прогнозирование таких показателей как занятость, средняя зарплата в секторе «Наука и научное обслуживание» и др., особенно с учетом эффекта изменения макроэкономических показателей (в частности, индекса инфляции, курсов доллара, евро и других валют) и влияния тех или иных мер (экономических, законодательных, налоговых, таможенных и т.д.), предпринимаемых на государственном уровне.

5. Применение современных статистических методов анализа нечисловых и интервальных данных. Использование статистики объектов нечисловой природы в выборочных обследованиях даст возможность реализовать идеи, изложенные в упомянутом выше сборнике [17]. В частности, регрессионный анализ в пространствах разнотипных признаков (объектов нечисловой природы) даст возможность оценить эффективность финансирования, а статистика интервальных данных позволит учесть неизбежные неточности в имеющихся данных.

6. Анализ существующей системы приоритетов отечественной науки с точки зрения объемов базового финансирования отдельных институтов, структуры государственных научно-технических программ, системы бюджетных и внебюджетных фондов, проведения экспертных опросов ведущих ученых и специалистов. Выявление наиболее перспективных направлений развития российской науки, выделение приоритетных технологий, сопоставление научно-технического развития России и развитых стран.

7. Оценка основных тенденций развития научной сферы за последние годы (изменение структуры сети научных организаций по ведомственному признаку, численности, эффективности функционирования, направлениям исследований и т.д.). Выявление закономерностей между основными количественными и качественными показателями развития науки и базовыми макроэкономическими показателями с последующим моделированием процессов, происходящих в сфере науки, в их взаимосвязи с инновационными процессами в экономике и социальной сфере.

8. Выделение «групп риска». Частный случай обсуждаемых выше задач — выделение (по формальным — отчетным — признакам) научных организаций, само существование которых оказывается под вопросом в ближайшем будущем.

Прогноз «выживаемости» НИИ может быть построен с помощью обучающих выборок на основе непараметрических оценок плотности в пространстве разнотипных признаков, часть координат которых — количественные признаки, а часть — качественные.

Существует много других интересных проблем [53, 54]. Например, выявление цикличности развития научно-технического потенциала страны, изучение динамики реальной и формальной структуры науки [55], форм примитивизации научной деятельности и научной продукции в условиях резкого спада производства и сокращения финансирования научного труда, проблемы отражения в общественном сознании и в самосознании научного сообщества специфики научной деятельности (включая анализ распространенных догм) и др.

Применение современных статистических методов в выборочных исследованиях научных организаций позволяет получать результаты, интересные с теоретической точки зрения и полезные для практики управляющих воздействий на развитие российской науки.

13.6. Статистические методы в изучении способных к математике школьников

Автор много лет работал в Вечерней математической школе (ВМШ) при Московском математическом обществе, в 1969–1977 гг. был ее директором. Статистические методы использовались нами как для решения задач управления образовательным учреждением, так и для изучения способных к математике школьников с целью повышения эффективности процесса обучения.

ВМШ была создана в 1963 г. профессором Е.Б. Дынкиным как дальнейшее развитие системы математических кружков при МГУ им. М.В. Ломоносова. Ее цель — способствовать математическому развитию школьников, расширять их математический кругозор. Вначале основную массу составляли учащиеся 8–10 классов, а затем, с развитием системы дневных физико-математических школ, — учащиеся 5–8 классов. Численность учеников ВМШ колебалась в разные годы от 300 до 1 400 (1971 г.). Общая информация о ВМШ приведена в сборниках серии «Математическая школа. Лекции и задачи» (издательство МГУ им. М.В. Ломоносова, 1965–1968 гг., выпуски I–XIV, доступны в Интернете: http://ilib.mccme.ru/djvu/msch/index.html) и в брошюре [10]. ВМШ мы рассматривали как полигон для отработки нового содержания и новых форм внеклассной работы по математике в средних классах общеобразовательной школы. Необходимо было подготовить тексты занятий ВМШ и задач, предлагаемых школьникам, а также отработать письменные формулировки и стиль изложения. Для этого в журнале «Пионер» (тираж — 1,5 млн экземпляров) был организован отдел «Встречи с тремя Неизвестными» (имеются в виду постоянные персонажи отдела Икс, Игрек и Зет). В этом отделе только автором настоящего учебника с 1970 г. по 1977 г. было опубликовано 48 статей и заметок. Общее представление о работе отдела дает рецензия [56] (см. также [57]). Ряд материалов помещен в журнале «Квант». В отличие от «Пионера», этот журнал полностью имеется в Интернете (http://kvant.mccme.ru/).

На основе разработок ВМШ по заказу Академии педагогических наук СССР было составлено пособие для учителей «Внеклассная работа по математике в 6–8 классах» [58]. Оно было выпущено издательством «Просвещение» в 1977 и 1984 гг. общим тиражом 500 тыс. экземпляров. Книга переведена на молдавский, литовский, казахский, болгарский языки. Таким образом, опыт ВМШ широко используется, созданное на его основе пособие есть практически в каждой школе. И не только в России.

Следует подчеркнуть, что особое внимание уделялось внедрению в преподавание различных разделов прикладной математики. Так, впервые адаптирована для учащихся средней школы модель Вильсона теории управления запасами (см. [59], [58, тема 22] и раздел 8.3 данной книги). Изложение основ теории вероятностей и математической статистики начиналось с постановки задачи о проверке статистической гипотезы о том, что для дефектных изделий в партии составляет 23 % (см. [60; 15, разд. 2.2] и главу 10 данной книги).

Для повышения эффективности деятельности ВМШ было признано полезным ее осмысление с научных позиций. Для этого в составе ВМШ был создан научно-исследовательский сектор, деятельность которого развивалась по двум направлениям — анкетные обследования школьников и разработка математических моделей отдельных сторон обучения математике. Некоторые итоговые результаты работ в первом направлении рассматриваются ниже и освещены в [10, 61]. Во втором направлении разработаны две модели — оптимального распределения времени между активным (решение задач) и пассивным (слушание лекций, чтение литературы) усвоением материала и расчета оптимального числа преподавателей [62, 63].

Анкетные обследования мы рассматривали как способ методами социологии получить представление о социально-психологических характеристиках больших совокупностей способных к математике школьников. Первое анкетирование учеников ВМШ было проведено в 1970 г. Получено около 1 000 анкет. В их обработке участвовали А.Н. Аверкин и другие сотрудники ВМШ. Ряд результатов приведен в [10, 61].

В частности, изучались маркетинговые коммуникации образовательных услуг, оказываемых ВМШ. Распространение и информации о наборе в ВМШ было организовано так.

1. Учащиеся связанных с ВМШ средних школ (т.е. школ, в которые обычно переходили ученики ВМШ в конце учебного цикла) получали по два объявления о начале занятий ВМШ. Одно из объявлений предлагалось вывесить в прежней школе (где учился до перехода в математическую школу), а второе — отдать родителям для информирования их коллег по работе. Так распространялось до 1 000 объявлений.

2. Помещались заметки в печати (журналы «Пионер», «Квант», газеты).

3. Делались объявления на городских и районных собраниях учителей.

Все ученики ВМШ отвечали на анкету, где был и вопрос: «Откуда узнал о ВМШ?» Результаты обработки ответов на этот вопрос приведены в табл. 1.

Таблица 1

Ответы на вопрос: «Откуда узнал о ВМШ?»

Откуда узнал о ВМШ	**Из печати**			**От знакомых, одноклассников, друзей**			**От родителей и родственников**		
Класс	5–6	7	8	5–6	7	8	5–6	7	8
% ответов	30	29	27	28	33	45	28	15	7
Откуда узнал о ВМШ	**От учителя**			**Из объявления**			**Прочие ответы**		
Класс	5–6	7	8	5–6	7	8	5–6	7	8
% ответов	10	12	12	2	4	3	2	7	6

Обращает на себя внимание малое число узнавших о ВМШ из объявлений. Создается впечатление, что объявления как средство распространения информации не играют большой роли, от них можно отказаться. Однако более глубокий анализ показывает, что подобное заключение является неверным Дело в том, что объявления распространяют учащиеся дневных школ, связанных с ВМШ, тем самым они (учащиеся) являются носителями информации о начале работы ВМШ. Эту информацию они передают знакомым школьникам и учителям, а также родителям. Мы видим, что информация, дошедшая до многих будущих учеников ВМШ через знакомых, родителей, учителей, имела своим источником объявления, отказ от которых, таким образом, не оправдан.

Результаты анкетного обследования были использованы руководством ВМШ при планировании работы. Так, анкета показала, что учителя и почтовые объявления играют сравнительно малую роль как источники информации о ВМШ, в отличие от публикаций в печати и устной информации [10, 61]. В частности, рассылка объявлений по школам с помощью почты была признана неэффективной, поскольку, как показала проверка, при этом информация обычно не доходит до учеников. Поэтому мы практически отказались от обращений к учителям и рассылки объявлений по почте, а силы сосредоточили на публикациях в журналах и газетах, на «командировании учащихся физико-математических школ в «родные» им школы для распространения информации о ВМШ.

В 1972 г. 150 победителям третьей Олимпиады журнала «Пионер» были разосланы анкеты. Возвращенные 50 анкет были обработаны Д.А. Модель и нами. Результаты частично опубликованы, в частности, в [57]. Результаты этого обследования имели большое значение для организации работы отдела «Встречи с тремя Неизвестными» журнала «Пионер». Например, было выявлено явное предпочтение (со стороны читателей) задач последовательному изложению материала. А среди задач — алгоритмическим, а не «на доказательство». Подтверждено положительное отношение к занимательности формулировок и т.д. (см. [61]). Обнаруженные предпочтения школьников — потребителей образовательных услуг были учтены и при составлении упомянутого выше пособия по внеклассной работе [58]. Полученные в результате анкетного обследования результаты были, в частности, использованы в дискуссиях с членами редакционно-издательского совета министерства просвещения СССР.

В 1975 г. совместно с НИИ общей и педагогической психологии (НИИ ОиПП) Академии педагогических наук СССР было проведено углубленное анкетирование учеников ВМШ. В соответствии с планом исследования анкет всех видов было получено около 400. Со стороны НИИ ОиПП руководителями работы являлись Н.С. Лейтес и И.В. Дубровина. Рабочая группа со стороны ВМШ действовала под нашим руководством и включала в себя Г.А. Гусейнова, Д.А. Модель, С.Д. Ошанина, В.Б. Шехтмана, М.П. Шпильрейна. Анкеты и результаты обработки переданы в НИИ ОиПП. Ниже рассматриваем лишь один из результатов обработки данных, представляющий методический интерес с точки зрения статистических методов.

По решению Правления ВМШ от 19 октября 1975 г. учащиеся ВМШ заполнили анкеты из двух частей — «для заполнения на занятиях» (30 вопросов) и «для заполнения дома» (38 вопросов). Ряд вопросов состоял из нескольких подвопросов. Ответы на некоторые из вопросов давались в свободной форме и не подлежали количественной обработке. Руководители групп ВМШ заполнили карты на каждого ученика своих групп. В карте — 22 вопроса. В результате на каждого ученика было заполнено 3 документа общим объемом 10 машинописных страниц. Такой большой объем информации на каждого ученика был продиктован интересами специалистов по педагогической психологии.

Представлялось целесообразным использовать результаты опросов для построения типологии способных к математике школьников с учетом их социально-психологических свойств. Предполагалось, что для каждого типа будет построена своя оптимальная система обучения. Так, при анализе анкет победителей олимпиады журнала «Пионер» (см. выше) установлено, что одним больше нравятся те области математики, в которых «задачи легче получаются», другим — те, где «задачи самые трудные». Видимо, оптимальные системы обучения для этих двух типов учащихся различаются.

Естественно предположить, что требуемую типологию можно получить с помощью кластер-анализа ответов школьников на вопросы некоторой «анкеты» (кроме анкеты в обычном смысле, могут применяться описанные выше «карты» и специальные тесты). Построение типологий с помощью традиционных методов рассмотрено ВА. Крутецким [64], Н.С. Лейтесом [65] и др. Прообразом подобной анкеты является та, что применялась в описываемом исследовании.

При формировании групп для обучения по разным оптимальным программам отнесение школьника к определенному типу будет происходить по ответам на анкету. Поэтому желательно иметь анкету с возможно меньшим числом вопросов (считая, что вопросы примерно равноценны по трудоемкости, по затратам на получение ответов). Добиться этого можно, оставив по одному из каждой группы вопросов, ответы на которые близки, т.е. из каждого кластера, выделенного при решении задач кластер-анализа вопросов.

Для «пилотажного» исследования в рамках намеченной программы были выбраны 10 вопросов. Ответы на каждый из них — результаты измерений в шкале наименований (в терминологии теории измерений — см., например, [15, разд. 1.2]). Перейдем к общей терминологии: вопрос — это признак, ответ на вопрос — это значение признака. Некоторым вопросам изначально соответствовали количественные признаки. Переход к номинальным признакам делался путем введения градаций (см. ниже) с целью повышения устойчивости ответов, интерпретируемости, а также возможности применения соответствующего математического аппарата. Всего было выделено 80 градаций.

Анализировались ответы на следующие вопросы.

1. Единственный ребенок или нет.

2. Среднее время, затрачиваемое в день на выполнение домашних заданий (градации: до 0,5 часа; 0,5–1 час; 1–1,5 часа; и т.д.).

3. Любимый предмет в школе.

4. Интересующая область знания.

5. Чем занимается в свободное время (градации: искусством; спортом; наукой и техникой; спортом и искусством; прогулками; и т.п.).

6. В каких кружках, кроме ВМШ, занимается (градации: ни в каких; в спортивном; в техническом; в математическом; в музыкальном; в гуманитарном; в техническом и математическом; в спортивном, техническом и музыкальном; и т.д.).

7. Какой областью искусства интересуется (градации: литературой; кино и театром; рисованием; и т. д.).

8. Занимается спортом или нет.

9. Какой общественной работой занимается (градации: ответственной — председатель совета отряда, звеньевой и т.п.; стенгазетой; и т.д.).

10. Какие качества человека ему нравятся (градации: интеллектуальные (ум); моральные (честность, верность); веселость, остроумие; деловые качества; всевозможные сочетания перечисленных).

Обрабатывались анкеты 46 пятиклассников.

Для применения алгоритмов кластер-анализа необходимо получить матрицу попарных расстояний между признаками (вопросами). Поставим k-му признаку в соответствие матрицу $r^k = \|r_{ij}^k\|$, где $r_{ij}^k = 1$, если i-й и j-й ученики одинаково ответили на k-й вопрос, и $r_{ij}^k = 0$ в противном случае. Таким образом, матрица $r^k = \|r_{ij}^k\|$ задает разбиение множества всех опрошенных на группы одинаково ответивших на k-й вопрос. В качестве расстояния между p-м и q-м признаками (вопросами) будем использовать расстояние Кемени (см. [15, разд. 1.6]) между разбиениями r^p и r^q учеников на группы одинаково ответивших на p-й и q-й вопросы:

$$d_{pq} = d(r^p, r^q) = \sum_{i,j} \left| r_{ij}^p - r_{ij}^q \right|.$$

Рассчитанные по анкетам расстояния приведены в табл. 2 (вычисления проведены Г.А. Гусейновым). Данные табл. 2 — исходные для работы трех алгоритмов кластер-анализа: ближайшего соседа, дальнего соседа (см. раздел 6.4) и алгоритма максимизации суммы существенных внутренних связей, предложенного в работе [66].

Таблица 2

Расстояния между социально-психологическими признаками способных к математике школьников [61]

p, q	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	–	–	–	–	–	–	–	–	–
2	1028	0	–	–	–	–	–	–	–	–
3	1028	608	0	–	–	–	–	–	–	–
4	1050	688	610	0	–	–	–	–	–	–
5	1012	686	636	634	0	–	–	–	–	–
6	1006	566	538	616	562	0	–	–	–	–
7	1012	1026	748	692	774	732	0	–	–	–
8	960	1088	1144	1122	1120	1130	1110	0	–	–
9	1026	878	874	830	836	802	904	1040	0	–
10	990	744	674	744	718	580	814	1090	830	0

Примечание. Матрица $\left\| d_{pq} \right\|$ симметрична относительно главной диагонали, поэтому в табл. 2 приведены только элементы, стоящие на главной диагонали и расположенные ниже нее.

В работе [66] предлагается находить разбиение из условия

$$f(R,a) = \sum_{1 \le s \le m} \sum_{p,q \in R_s} (a_{pq} - a) \to \max_R ,$$

где $R = \{R_1, R_2, \ldots, R_m\}$ — искомое разбиение, a_{pq} — показатель связи между p-м и q-м признаками, a — параметр алгоритма. Для рассматриваемых данных положили $a_{pq} = n^2 - d_{pq}$, где $n = 10$ — число классифицируемых признаков.

Только что сформулированная оптимизационная задача в вычислительном плане достаточно сложна. В [66] предложен эвристический алгоритм ее решения, который независимо от его связи с оптимизационной остановкой представляет собой некоторый алгоритм кластер-анализа. В 1976 г. мы приняли решение его использовать, т.к. в то время он представлялся наиболее соответствующим задаче, обоснованным и современным. На программирование и расчеты у Г.А. Гусейнова ушло полгода. Результаты представлены в табл. 3. В столбце «значение параметра» приведено значение a, при котором впервые появляется разбиение, указанное в следующем столбце. Оно сохранялось при увеличении a вплоть до значения, данного в следующей строке.

Таблица 3

Классификация признаков по методу [66]
при различных значениях параметра *a*

Шаг	Параметр *a*	Кластеры
1	538	{3, 6}, {1}, {2}, {4}, {5}, {7}, {8}, {9}, {10}
2	580	{2, 3, 6}, {1}, {4}, {5}, {7}, {8}, {9}, {10}
3	608	{2, 3, 5, 6}, {1}, {4}, {7}, {8}, {9}, {10}
4	634	{2, 3, 4, 5, 6}, {1}, {7}, {8}, {9}, {10}
5	686	{2, 3, 4, 5, 6, 10}, {1}, {7}, {8}, {9}
6	744	{2, 3, 4, 5, 6, 7, 10}, {1}, {8}, {9}
7	874	{2, 3, 4, 5, 6, 7, 9, 10}, {1}, {8}
8	990	{2, 3, 4, 5, 6, 7, 9, 10}, {1, 8}
9	1040	{1, 2, 3, 4, 5, 6, 7, 9, 10}, {8}
10	1090	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

В соответствии с общими положениями теории классификации, рассмотренными в разделе 6.4, все алгоритмы кластер-анализа должны давать близкие с содержательной точки зрения результаты, если задача разбиения на кластеры имеет решение. Поэтому мы применили к данным табл. 2 два иерархических агломеративных алгоритма типа «Дендрограмма» — ближайшего соседа и дальнего соседа. По каждому из алгоритмов ручной счет занял 2–3 часа. Результаты представлены в табл. 4 и 5 соответственно. В столбцах «Параметр *a*» приведены расстояния между кластерами, при которых происходят очередные объединения.

Таблица 4

Классификация признаков по методу ближайшего соседа
при различных значениях параметра *a*

Шаг	Параметр *a*	Кластеры
1	538	{3, 6}, {1}, {2}, {4}, {5}, {7}, {8}, {9}, {10}
2	562	{3, 5, 6}, {1}, {2}, {4}, {7}, {8}, {9}, {10}
3	566	{2, 3, 5, 6}, {1}, {4}, {7}, {8}, {9}, {10}
4	580	{2, 3, 5, 6, 10}, {1}, {4}, {7}, {8}, {9}
5	610	{2, 3, 4, 5, 6, 10}, {1}, {7}, {8}, {9}
6	692	{2, 3, 4, 5, 6, 7, 10}, {1}, {8}, {9}
7	802	{2, 3, 4, 5, 6, 7, 9, 10}, {1}, {8}
8	960	{2, 3, 4, 5, 6, 7, 9, 10}, {1, 8}
9	990	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

Таблица 5

Классификация признаков по методу дальнего соседа при различных значениях параметра *a*

Шаг	Параметр a	Кластеры
1	538	{3, 6}, {1}, {2}, {4}, {5}, {7}, {8}, {9}, {10}
2	608	{2, 3, 6}, {1}, {4}, {5}, {7}, {8}, {9}, {10}
3	634	{2, 3, 6}, {4, 5}, {1}, {7}, {8}, {9}, {10}
4	688	{2, 3, 4, 5, 6}, {1}, {7}, {8}, {9}, {10}
5	744	{2, 3, 4, 5, 6, 10}, {1}, {7}, {8}, {9}
6	878	{2, 3, 4, 5, 6, 9, 10}, {1}, {7}, {8}
7	960	{2, 3, 4, 5, 6, 7, 9, 10}, {1, 8}, {7}
8	1026	{2, 3, 4, 5, 6, 7, 9, 10}, {1, 8}
9	1144	{1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

Содержательные выводы, приведенные в [61], сохраняются и при анализе табл. 4 и 5. (Выводы относятся к педагогической психологии и методике математики, поэтому здесь их не приводим.) Результаты обработки реальных данных табл. 2 тремя алгоритмами подтверждают вывод общей теории классификации (см. раздел 6.4) о целесообразности использования в задачах кластер-анализа наиболее простых в расчетном смысле алгоритмов.

Обсудим одну встречающуюся в литературе рекомендацию. Все три использованных нами алгоритма являются однопараметрическими. Для однопараметрических алгоритмов в работах [67, 68] предлагается выделять разбиения, которым соответствуют наибольшие интервалы устойчивости по параметру, т.е. наибольшие разности значений параметров в соседних строках таблиц 3–5. Предполагается, что эти разбиения являются «наиболее объективными». В табл. 6 приведены наиболее устойчивые по параметру разбиения, построенные по данным табл. 2 (подп. а)), а также «вторые по устойчивости» (подп. б)). Различие одноименных строк в табл. 6, соответствующих трем алгоритмам, говорит о необоснованности рекомендации, предложенной в работах [67, 68]. Содержательные выводы мы делаем по всей «дендрограмме», а не по наиболее устойчивому разбиению [61].

Таблица 6

Наиболее устойчивые по параметру разбиения на кластеры (согласно [67, 68])

Название метода кластер-анализа	Наиболее устойчивые по параметру разбиения на кластеры
Метод максимизации суммы существенных внутренних связей	а) {2, 3, 4, 5, 6, 7, 10}, {1}, {8}, {9}; б) {2, 3, 4, 5, 6, 7, 9, 10}, {1}, {8}
Метод ближайшего соседа	а) {2, 3, 4, 5, 6, 7, 9, 10}, {1}, {8}; б) {2, 3, 4, 5, 6, 7, 10}, {1}, {8}, {9}
Метод дальнего соседа	а) {2, 3, 4, 5, 6, 10}, {1}, {7}, {8}, {9}; б) {2, 3, 4, 5, 6, 7, 9, 10}, {1, 8}

Необоснованность предложения авторов [67, 68] можно продемонстрировать не только на численных примерах, но и с помощью методологического анализа. Рассмотрим алгоритм ближайшего соседа, использующий меру близости $d(x, y)$. Нетрудно видеть, что дендрограмма, полученная с его помощью, является инвариантной в порядковой шкале измерения меры близости $d(x, y)$, т.е. не меняется при любом строго возрастающем преобразовании меры близости. Параметром в смысле, принятом в работах [67, 68], является численное значение меры близости между кластерами. Важны те ее значения, при которых происходит объединение кластеров. Покажем, что «наиболее устойчивое по параметру разбиение» не является инвариантным в порядковой шкале. Более того, при соответствующем подборе преобразования меры близости в качестве «наиболее устойчивого по параметру» может оказаться любое разбиение, даваемое «обрезанием» дендрограммы на том или ином уровне. Построение элементарно: выбрав произвольный интервал между соседними объединениями кластеров, рассмотрим любое строго возрастающее преобразование, оставляющее этот интервал на месте и переводящее верхний и нижний лучи оси изменения параметра в малые интервалы, примыкающие к концам неподвижного интервала.

Можно получить опровержение метода работ [67, 68], ограничившись более узкими семействами преобразований. Рассмотрим степенные преобразования: $d \to d^{\alpha}$, $\alpha > 0$. Пусть при $d = d_1$ объединяются последние два кластера, а предыдущее объединение произошло при $d = d_2$ — после него остались эти два кластера Поскольку

$$\lim_{\alpha \to +\infty} \frac{d_1^{\alpha} - d_2^{\alpha}}{d_2^{\alpha}} = +\infty ,$$

то при достаточно большом α «наиболее объективном» по [67, 68] будет разбиение на два кластера!

В чем причина несостоятельности предложения [67, 68]? Устойчивость рассматривается односторонне — относительно изменения параметра в данном алгоритме, но не относительно перехода к другому алгоритму. Стоит сделать такой переход, как «устойчивое» разбиение становится крайне неустойчивым.

Подводя итоги настоящей главы, следует отметить, что в социологии с успехом используются самые разные статистические методы, позволяющие получать в рамках данной предметной области полезные теоретические и практические выводы и рекомендации. Статистические методы социологии активно развиваются. В частности, на недавних конференциях широко обсуждалась теория люсианов [70, 71].

Литература

1. *Осипов, Г.В.* Российская социология в XXI веке / Г.В. Осипов // Социологические исследования. — 2004. — № 3 (239). — С. 3–15.

2. *Толстова, Ю.Н.* Анализ социологических данных: методология, дескриптивная статистика, изучение связей между номинальными признаками / Ю.Н. Толстова. — Москва : Научный мир, 2000. 352 с.

3. *Толстова, Ю.Н.* Измерение в социологии / Ю.Н. Толстова. — Москва : Инфра-М, 1998. — 224 с.

4. *Татарова, Г.Г.* Методология анализа данных в социологии (введение) : учебник для вузов / Г.Г. Татарова. — Москва : NOTA BENE, 1999. — 224 с.

5. *Орлов, А.И.* Эконометрика / А.И. Орлов. — Изд. 3-е, испр. и доп. — Москва : Экзамен, 2004. — 576 с.

6. Методы современной математики и логики в социологических исследованиях. / под редакцией Э.П. Андреева. — Москва : Институт социологических исследований АН СССР, 1977. — 171 с.

7. Математические методы и модели в социологии / под редакцией В.Н. Варыгина. — Москва : Изд-во Института социологических исследований АН СССР, 1977. — 192 с.

8. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

9. *Шляпентох, В.Э.* Социология для всех: некоторые проблемы, результаты, методы / В.Э. Шляпентох. — Москва : Советская Россия, 1970. — 250 с.

10. *Орлов, А.И.* О теоретических основах внеклассной работы по математике и опыте Вечерней математической школы при Московском математическом обществе / А.И. Орлов // Бюллетень № 2 Всесоюзного центра статистических методов и информатики. — Москва : ВЦСМИ, 1991. — 48 с.

11. Тезисы докладов и выступлений на II Всероссийском социологическом конгрессе «Российское общество и социология в XXI веке: социальные вызовы и альтернативы» : в 3 т. — Москва : Альфа-М, 2003.

12. Актуальные проблемы социологической науки и социальной практики: Научная конференция «Сорокинские чтения — 2002». Москва, МГУ им. М.В. Ломоносова, 17–18 декабря 2002 г. : сборник научных докладов : в 3 томах. Т. 3: Математическое моделирование социальных процессов. Выпуск 5 / под общей редакцией А.А. Самарского, В.И. Добренькова, А.П. Михайлова. — Москва : МАКС Пресс, 2003. — 252 с.

13. *Орлов, А.И.* Статистические методы в российской социологии (тридцать лет спустя) / А.И. Орлов // Социология: методология, методы, математические модели. — 2005. — № 20. — С. 32–53.

14. *Толстова, Ю.Н.* Математические методы в социологии / Ю.Н. Толстова // Социология в России / под редакцией В.А. Ядова. — 2-е изд., перераб. и дополн. — Москва : Изд-во Института социологии РАН, 1998. — С. 83–89, 98–103.

15. *Орлов, А.И.* Прикладная статистика : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 672 с.

16. *Орлов, А.И.* Статистика объектов нечисловой природы и обработка социологических данных / А.И. Орлов // Математические методы в социологическом исследовании. — Москва : Наука, 1981. — С. 67–75.

17. Анализ нечисловой информации в социологических исследованиях. / под редакцией В.Г. Андреенкова, А.И. Орлова, Ю.Н. Толстовой. — Москва : Наука, 1985.

18. *Орлов, А.И.* Статистика объектов нечисловой природы и анализ данных о научном потенциале / А.И. Орлов, Е.Г. Нечаева, А.В. Соколов // Социология: методология, методы, математические модели. — 1995. — № 5–6. — С. 118–136.

19. *Орлов, А.И.* Принятие решений. Теория и методы разработки управленческих решений / А.И. Орлов. — Москва : МарТ, 2005. — С. 496.

20. *Орлов, А.И.* Случайные множества с независимыми элементами (люсианы) и их применения / А.И. Орлов // Алгоритмическое и программное обес-

печение прикладного статистического анализа. Ученые записки по статистике. — Т. 36. — Москва : Наука, 1980. — С. 287–308.

21. *Рыданова, Г.В.* Некоторые вопросы статистического анализа случайных бинарных векторов : специальность 01.01.05: «Теория вероятностей и математическая статистика» : диссертация на соискание ученой степени кандидата физико-математических наук / Рыданова Галина всеволодовна ; Московский государственный университет им. М.В. Ломоносова. — Москва, 1987. — 139 с.

22. *Шошин, П.Б.* Размытые числа как средство описания субъективных величин / П.Б. Шошин // Статистические методы анализа экспертных оценок. — Москва : Наука, 1977. — С. 234–250.

23. *Орлов, А.И.* Асимптотика квантования и выбор числа градаций в социологических анкетах / А.И. Орлов // Математические методы и модели в социологии. — Москва : Изд-во ИСИ АН СССР, 1977. — С. 42–55.

24. *Шубкин, В.Н.* Социологические опыты / В.Н. Шубкин. — Москва: Мысль, 1970. — 256 с.

25. *Саганенко, Г.И.* Надежность результатов социологического исследования / Г.И. Саганенко. — Ленинград : Наука, 1983. — 180 с.

26. *Шляпентох, В.Э.* Проблемы достоверности статистической информации в социологических исследованиях / В.Э. Шляпентох. — Москва : Статистика, 1973. — 144 с.

27. *Колмогоров, А.Н.* Элементы теории функций и функционального анализа : учебник / А.Н. Колмогоров, С.В. Фомин. — Москва : Наука, 1972. — 496 с.

28. *Гнеденко, Б.В.* Курс теории вероятностей : учебник / Б.В. Гнеденко. — 7-е изд., испр. — Москва : Эдиториал УРСС, 2001. — 320 с.

29. *Лоэв, М.* Теория вероятностей / М. Лоэв. — Москва : ИЛ, 1962. — 720 с.

30. *Моргенштерн, О.* О точности экономико-статистических наблюдений / О. Моргенштерн. — Москва : Статистика, 1968. — 396 с.

31. *Титма, М.Х.* Математические методы в арсенале социолога / М.Х. Титма, Л.М. Тоодинг / Социологические исследования. — 1986. —№ 4. — С. 123–128.

32. *Глотов, В.А.* Векторная стратификация / В.А. Глотов, В.В. Павельева. — Москва : Наука, 1984. — 94 с.

33. *Ильясов, Ф.Н.* Экспериментальное обоснование количества делений шкалы / Ф.Н. Ильясов / Социологические исследования. — 1984. — № 4. — С. 113–116.

34. *Литвак, Б.Г.* О выборе делений шкалы / Б.Г. Литвак. — Статистические методы анализа экспертных оценок. — Москва : Наука, 1977. — С. 228–234.

35. *Лавренченко, А.С.* Решение задач оптимального управления запасами при нестационарном пуассоновском потоке требований / А.С. Лавренченко // Сборник научных трудов Московского авиационного института. — 1974. — № 306. — С. 62–72.

36. *Орлов, А.И.* Отказ от пуассоновости спроса в одной модели управления запасами / А.И. Орлов // Исследования по вероятностно-статистическому моделированию реальных систем. — Москва : Изд-во ЦЭМИ АН СССР, 1977. — С. 93–96.

37. *Ротарь, Г.В.* Об одной задаче управления резервами / Г.В. Ротарь / Экономика и математические методы. — 1976. — Т. 12. — № 4. — С. 733–739.

38. *Ротарь, Г.В.* Одна задача об управлении резервами / Г.В. Ротарь // Теория вероятностей и ее применения. — 1972. — Т. XVII. — № 3. — С. 597–599.

39. *Саульев, В.К.* Математическая теория оптимального управления запасами / В.К. Саульев, А.С. Лавренченко. — Москва : МАИ, 1974. — 188 с.

40. *Хедли, Дж.* Анализ систем управления запасами / Дж. Хедли, Т. Уайтин. — Москва : Наука, 1969. — 512 с.

41. *Феллер, В.* Введение в теорию вероятностей и ее приложения : т. 2 / В. Феллер. — Москва : Мир, 1984. — 738 с.

42. *Воскресенский, А.В.* Принцип инвариантности в одной модели управления запасами : дипломная работа. — Москва : Механико-математический факультет МГУ им. М.В. Ломоносова, 1977.

43. *Орлов, А.И.* Принцип инвариантности в одной модели управления запасами / А.И. Орлов, А.В. Воскресенский // Прикладной многомерный статистический анализ. Ученые записки по статистике. — Т. 33. — Москва : Наука, 1978. — С. 307–311.

44. *Орлов, А.И.* Горизонтная устойчивость двухуровневой модели управления запасами //А.И. Орлов // Многомерный статистический анализ (математическое обеспечение). — Москва : Изд-во ЦЭМИ АН СССР, 1979. — С. 187–199.

45. *Морено, Дж.* Социометрия / Дж. Морено. — Москва : ИЛ, 1958. — 289 с.

46. *Максименко, В.С.* Зачем социологу математика / В.С. Максименко, В.И. Паниотто. — Киев : Радяньска школа, 1988. — 223 с.

47. Наука России — 1993 : статистический сборник. — Москва : ЦИСН, 1994. — 240 с.

48. Научно-техническая и инновационная политика. Российская федерация. Том 1. Оценочный доклад. — Москва : Организация экономического сотрудничества и развития, 1994. — 124 с.

49. *Орлов, А.И.* Социологический прогноз развития Российской науки на 1993–1995 годы / А.И. Орлов // Наука и технология в России. — 1993. — № 1. — С. 29.

50. *Страхов, В.Н.* Нужны ли подобные прогнозы? / В.Н. Страхов — Наука и технология в России. — 1993. — № 1. — С. 30–31.

51. *Орлов, А.И.* Организационные методы управления наукой и статистика объектов нечисловой природы / А.И. Орлов // Тезисы докладов Всесоюзного симпозиума «Медицинское науковедение и автоматизация информационных процессов». — Москва : ВНИИ медицинской и медико-технической информации Министерства здравоохранения СССР, 1984. — С. 215–216.

52. Frascati Manual: 1993. The Measurement of Scientific and Technological Activities. — Paris : OECD, 1994 — 261 с.

53. Развитие науки в России : статистический сборник. — Москва : ЦИСН, 1993. — 468 с.

54. *Налимов, В.В.* Наукометрия / В.В. Налимов, А.Б. Мульченко. — Москва : Наука, 1969. — 192 с.

55. *Орлов, А.И.* Прикладная статистика — «Золушка» научно-технической революции // А.И. Орлов // Наука и технология в России. — 1994. — № 1 (3). — С. 13–14.

56. *Ломакин, Ю.В.* «Встречи с тремя Неизвестными» в журнале «Пионер» / Ю.В. Ломакин, С.А. Козырева, Г.А. Соколова // Математика в школе. — 1976. — № 4. — С. 85–87.

57. *Орлов, А.И.* О математическом разделе «Встречи с тремя Неизвестными» журнала «Пионер» / А.И. Орлов // Заочное обучение математике школьников 8–10 классов. — Москва : Изд-во НИИ содержания и методов обучения АПН СССР, 1976. — С. 78–81.

58. *Гусев, В.А.* Внеклассная работа по математике в 6–8 классах / В.А. Гусев, А.И. Орлов, А.Л. Розенталь. — Москва : Просвещение, 1977. — 288 с. (2-е изд., испр. и доп. — Москва : Просвещение, 1984; перевод на молдавский язык. — Кишинев : Лумина, 1980; перевод на литовский язык. — Каунас : Швеса, 1982; перевод на казахский язык. — Алма-Ата : Мектеп, 1986.)

59. *Орлов, А.И.* Про управление запасами / А.И. Орлов // Подготовка студентов педагогических институтов к внеурочной работе по математике. — Вологда : Изд-во ВГПИ, 1975. — С. 10–20.

60. *Орлов, А.И.* Вероятностное пространство, неравенство Чебышева и закон больших чисел — основа курса теории вероятностей для школьников / А.И. Орлов // Подготовка студентов педагогических институтов к внеурочной работе по математике. — Вологда : Изд-во ВГПИ, 1976. — С. 13–29.

62. *Орлов А.И.* Проблемы устойчивости в некоторых моделях управления запасами и ресурсами / А.И. Орлов // Алгоритмы многомерного статистического анализа и их применения. — Москва : Изд-во ЦЭМИ АН СССР, 1975. — С. 94–105.

63. *Орлов, А.И.* Математические модели отдельных сторон обучения математике/ А.И. Орлов // Проблемы преподавания математики в вузах : сборник научно-методических статей по математике. — Вып. 7. — Москва : Высшая школа, 1978. — С. 28–34.

64. *Крутецкий, В.А.* Психология математических способностей школьников / В.А. Крутецкий. — Москва : Просвещение, 1968. — 432 с.

65. *Лейтес, Н.С.* Умственные способности и возраст / Н.С. Лейтес. — Москва : Педагогика, 1971. — 280 с.

66. *Куперштох, В.Л.* Сумма внутренних связей как показатель качества классификации / В.Л. Куперштох, Б.Г. Миркин, В.А. Трофимов // Автоматика и телемеханика. — 1976. — № 3. — С. 91–98.

67. *Плоткин, А.А.* Устойчивость разбиения как критерий оптимальности построенной классификации / А.А. Плоткин // Статистические методы анализа экспертных оценок. — Москва : Наука, 1977. — С. 111–123.

68. Типология потребления / С.А. Айвазян, Н.М. Римашевская, З.И. Бежаева [и др.]. — Москва : Наука, 1978. — 308 с.

69. *Новиков, Д.А.* Статистические методы в педагогических исследованиях (типовые случаи) / Д.А. Новиков. — Москва : МЗ-Пресс, 2004. — 67 с.

70. *Орлов, А.И.* Роль люсианов в теории экспертных оценок / А.И. Орлов // Теория активных систем : труды международной научно-технической конференции (16–18 ноября 2005 г., Москва, Россия) / под общей редакцией В.Н. Бурков, Д.А. Новиков. — Москва : ИПУ РАН, 2005. — С. 64–65.

71. *Орлов, А.И.* Перспективы применения люсианов в социологии / А.И. орлов // Тезисы II Всероссийской научной конференции «Сорокинские чтения: Будущее России: стратегии развития». — Москва : Альфа-М, 2005. — С. 213–216.

72. *Орлов, А.И.* Статистика нечисловых данных за сорок лет (обзор) / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2019. — Т. 85. — № 11. — С. 69–84.

73. *Орлов, А.И.* Статистика нечисловых данных — центральная часть современной прикладной статистики / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 156. — С. 111–142.

74. *Орлов, А.И.* Теория люсианов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 101. — С. 275–304.

75. *Орлов, А.И.* Асимптотика квантования, выбор числа градаций в социологических анкетах и двухуровневая модель управления запасами / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 123. — С. 660–687.

76. *Лойко, В. И.* Современные подходы в наукометрии: монография / В.И. Лойко, Е.В. Луценко, А.И. Орлов ; под научной редакцией профессора С.Г. Фалько. — Краснодар : Изд-во КубГАУ, 2017. — 532 с.

77. *Луценко, Е.В.* Количественная оценка степени манипулирования индексом Хирша и его модификация, устойчивая к манипулированию / Е.В. Луценко, А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 121. — С. 202–234.

78. *Орлов, А.И.* Статистические и экспертные методы в задачах экономики и управления наукой / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 166. — С. 1–35.

79. *Орлов, А.И.* Теория измерений как часть методов анализа данных: размышления над переводом статьи П.Ф. Веллемана и Л. Уилкинсона / А.И. Орлов // Социология: методология, методы, математическое моделирование. — 2012. — № 35. — С. 155–174.

80. *Орлов, А.И.* Статистические методы в российской социологии (тридцать лет спустя) / А.И. Орлов // Социология: методология, методы, математические модели. — 2005. — № 20. — С. 32–53.

81. *Орлов, А.И.* Математические методы в социологии за сорок пять лет / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 117. — С. 91–119.

Контрольные вопросы

1. Расскажите о развитии статистических методов в отечественной социологии.

2. Дайте определение понятию «люсиан».

3. Для проверки каких статистических гипотез используют теорию люсианов?

4. Как рекомендуют выбирать число градаций в социологических анкетах на основе изучения асимптотики квантования?

5. Как применяют асимптотику квантования в логистике?

6. Чем неформальный лидер отличается от формального?

7. Почему менеджеру полезно социометрическое исследование?

8. Для решения каких практических задач применяют выборочные исследования научных организаций?

9. Как анкетные исследования помогают повышать эффективность маркетинговых коммуникаций в области образовательных услуг?

10. Расскажите об использовании опросов в сочетании с методами кластер-анализа для построения типологии учащихся.

Темы докладов, рефератов, исследовательских работ

1. Статистические данные в социологии и методы их анализа.

2. Построение непараметрической теории парных сравнений на основе теории люсианов.

3. Несмещенные оценки в теории люсианов.

4. Принцип уравнивания погрешностей и выбор числа градаций в социологических анкетах.

5. Тождество Вальда для математического ожидания суммы случайного числа случайных слагаемых.

6. Асимптотическая теория суммирования случайного числа случайных слагаемых.

7. Современные методы социометрии.

8. Динамика статистических показателей, характеризующих состояние науки в России.

9. Статистические методы в науковедении и управлении наукой.

10. Наукометрия как область научных и прикладных исследований с интенсивным использованием статистических методов.

11. Статистические методы в педагогических исследованиях (на основе [69]).

ЧАСТЬ 4. ИНСТРУМЕНТЫ, ИСТОРИЯ И ПЕРСПЕКТИВЫ СТАТИСТИЧЕСКИХ МЕТОДОВ

ГЛАВА 14. ТЕОРЕТИЧЕСКИЕ ИНСТРУМЕНТЫ СТАТИСТИЧЕСКИХ МЕТОДОВ

В настоящей главе собраны основные математические утверждения, постоянно используемые при обосновании статистических методов. Эти утверждения отнюдь не всегда легко найти в литературе по теории вероятностей и математической статистике. Например, такие рассматриваемые далее теоремы и методы, как многомерная центральная предельная теорема, теоремы о наследовании сходимости и метод линеаризации, даже не включены в энциклопедию «Вероятность и математическая статистика» [1] — наиболее полный, по мнению составителей, свод знаний по заявленной тематике. Последний факт наглядно демонстрирует разрыв между математической дисциплиной «теория вероятностей и математическая статистика» и потребностями прикладной статистики и других статистических методов.

14.1. Законы больших чисел

Законы больших чисел позволяют описать поведение сумм случайных величин. Примером является следующий результат, доказанный русским математиком П.Л. Чебышевым (1821–1894) в 1867 г. Пусть сначала вероятностное пространство состоит из конечного числа элементов.

Теорема Чебышева. Пусть случайные величины $X_1, X_2, \ldots, X_k$ попарно независимы и существует число C такое, что $D(X_i)\leq C$ при всех $i = 1, 2, \ldots, k$. Тогда для любого положительного ε выполнено неравенство

$$P\left\{\left|\frac{X_1+X_2+\ldots+X_k}{k}-\frac{M(X_1)+M(X_2)+\ldots+M(X_k)}{k}\right|\geq\varepsilon\right\}\leq\frac{C}{k\varepsilon^2}. \qquad (1)$$

Частным случаем теоремы Чебышева является теорема Бернулли — первый в истории вариант закона больших чисел. Известный математики Якоб Бернулли (1654–1705), живший в городе Базель в Швейцарии, в самом конце XVII века доказал это утверждение в рамках математической модели (опубликовано доказательство было лишь после его смерти, в 1713 году). Современная формулировка теоремы Бернулли такова.

Теорема Бернулли. Пусть m — число наступлений события A в k независимых (попарно) испытаниях, и p есть вероятность наступления события A в каждом из испытаний. Тогда при любом $\varepsilon > 0$ справедливо неравенство

$$P\left\{\left|\frac{m}{k}-p\right| \geq \varepsilon\right\} \leq \frac{p(1-p)}{k\varepsilon^2}. \qquad (2)$$

Ясно, что при росте k выражения в правых частях формул (1) и (2) стремятся к 0. Таким образом, среднее арифметическое попарно независимых случайных величин сближается со средним арифметическим их математических ожиданий.

Напомним, что выше шла речь лишь о пространствах элементарных событий из конечного числа элементов. Однако приведенные теоремы верны и в общем случае, для произвольных пространств элементарных событий. Однако в список условий закона больших чисел необходимо добавить требование существования дисперсий. Легко видеть, что если существуют дисперсии, то существуют и математические ожидания. Закон больших чисел в форме Чебышева приобретает следующий вид.

Теорема Чебышева [2, с.147]. Если X_1, X_2, …, X_k, … — последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной,

$$D(X_1)\leq C,\ D(X_2)\leq C, \ldots\ D(X_i)\leq C, \ldots$$

то, каково бы ни было постоянное $\varepsilon > 0$,

$$\lim_{k\to\infty} P\left\{\left|\frac{1}{n}\sum_{j=1}^{k} X_j - \frac{1}{n}\sum_{j=1}^{k} MX_j\right| < \varepsilon\right\} = 1. \qquad (3)$$

С точки зрения прикладных статистических исследований ограниченность дисперсий вполне естественна. Она вытекает, например, из ограниченности диапазона изменения практически всех величин, используемых при реальных расчетах.

В 1923 г. А.Я. Хинчин показал, что если случайные величины не только независимы, но и одинаково распределены, то существование у них математического ожидания является необходимым и достаточным условием для применимости закона больших чисел [2, с. 150]. Найдены и более экзотические варианты закона больших чисел. Например, такой.

Теорема [2, с. 150–151]. Для того чтобы для последовательности X_1, X_2, ..., X_k, ... (как угодно зависимых) случайных величин при любом положительном ε выполнялось соотношение (3), необходимо и достаточно, чтобы при $n \to \infty$

$$M\frac{\left(\sum_{j=1}^{k}(X_j - MX_j)\right)^2}{n^2 + \left(\sum_{j=1}^{k}(X_j - MX_j)\right)^2} \to 0.$$

Законы больших чисел для случайных величин служат основой для аналогичных утверждений для случайных элементов в пространствах более сложной природы. В частности, в пространствах произвольной природы. Однако здесь мы ограничимся классическими формулировками, служащими основой для современных статистических методов.

Смысл классических законов больших чисел состоит в том, что выборочное среднее арифметическое независимых одинаково распределенных случайных величин приближается (сходится) к математическому ожиданию этих величин. Другими словами, *выборочные средние сходятся к теоретическому среднему.*

Это утверждение справедливо и для других видов средних. Например, выборочная медиана сходится к теоретической медиане. Это утверждение — тоже закон больших чисел, но не классический.

Существенным продвижением в теории вероятностей во второй половине XX в. явилось введение средних величин в пространствах произвольной природы и получение для них законов больших чисел, т.е. утверждений, состоящих в том, что эмпирические (т.е. выборочные) средние сходятся к теоретическим средним.

14.2. Центральные предельные теоремы

Простейший вариант Центральной предельной теоремы (ЦПТ) теории вероятностей таков.

Центральная предельная теорема (для одинаково распределенных слагаемых). Пусть X_1, X_2, ..., X_n, ... – независимые одинаково распределенные случайные величины с математическими ожиданиями $M(X_i) = m$ и дисперсиями

$D(X_i) = \sigma^2$, $i = 1, 2, \ldots, n, \ldots$ Тогда для любого действительного числа x существует предел

$$\lim_{n\to\infty} P\left(\frac{X_1 + X_2 + \ldots + X_n - nm}{\sigma\sqrt{n}} < x\right) = \Phi(x),$$

где $\Phi(x)$ — функция стандартного нормального распределения.

Эту теорему иногда называют теоремой Линдеберга — Леви [3, с. 122].

В ряде прикладных задач не выполнено условие одинаковой распределенности. В таких случаях центральная предельная теорема обычно остается справедливой, однако на последовательность случайных величин приходится накладывать те или иные условия. Суть этих условий состоит в том, что ни одно слагаемое не должно быть доминирующим, вклад каждого слагаемого в среднее арифметическое должен быть пренебрежимо мал по сравнению с итоговой суммой. Наиболее часто используется теорема Ляпунова.

Центральная предельная теорема (для разнораспределенных слагаемых) — *теорема Ляпунова*. Пусть $X_1, X_2, \ldots, X_n, \ldots$ — независимые случайные величины с математическими ожиданиями $M(X_i) = m_i$ и дисперсиями $D(X_i) = \sigma_i^2 \neq 0$, $i = 1, 2, \ldots, n, \ldots$ Пусть при некотором $\delta > 0$ у всех рассматриваемых случайных величин существуют центральные моменты порядка $2 + \delta$ и безгранично убывает «дробь Ляпунова»:

$$\lim_{k\to\infty} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} M \mid X_k - m_k \mid^{2+\delta} = 0,$$

где

$$B_k^2 = \sum_{i=1}^{k} \sigma_i^2 = D\left(\sum_{i=1}^{k} X_i\right).$$

Тогда для любого действительного числа x существует предел

$$\lim_{n\to\infty} P\left(\frac{X_1 + X_2 + \ldots + X_n - m_1 - m_2 - \ldots - m_n}{B_n} < x\right) = \Phi(x), \qquad (1)$$

где $\Phi(x)$ — функция стандартного нормального распределения.

В случае одинаково распределенных случайных слагаемых

$$m_1 = m_2 = \ldots = m_n = m, \quad B_n = D(X_1 + X_2 + \ldots + X_n) = \sigma\sqrt{n},$$

и теорема Ляпунова переходит в теорему Линдеберга — Леви.

История получения центральных предельных теорем для случайных величин растянулась на два века — от первых работ Муавра в 30-х гг. XVIII в. для необходимых и достаточных условий, полученных Линдебергом и Феллером в 30-х гг. XX в.

Теорема Линдеберга — Феллера. Пусть $X_1, X_2, \ldots, X_n, \ldots$ — независимые случайные величины с математическими ожиданиями $M(X_i) = m_i$ и дисперсиями $D(X_i) = \sigma_i^2 \neq 0$, $i = 1, 2, \ldots, n, \ldots$ Предельное соотношение (1), т.е. центральная предельная теорема, выполнено тогда и только тогда, когда при любом $\tau>0$

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-m_k|>\tau B_n}(x-m_k)^2\,dF_k(x0=0,$$

где $F_k(x)$ обозначает функцию распределения случайной величины X_k.

Доказательства перечисленных в настоящем разделе центральных предельных теорем для случайных величин можно найти в классическом курсе теории вероятностей [2].

Для обоснования многих статистических методов большое значение имеет многомерная центральная предельная теорема. В ней речь идет не о сумме случайных величин, а о сумме случайных векторов.

Необходимое и достаточное условие многомерной сходимости [3, с. 124]. Пусть F_n обозначает совместную функцию распределения k-мерного случайного вектора $(X_n^{(1)},\ldots,X_n^{(k)})$, $n = 1, 2, \ldots$, и $F_{\lambda n}$ — функция распределения линейной комбинации $\lambda_1 X_n^{(1)}+\ldots+\lambda_n X_n^{(k)}$. Необходимое и достаточное условие для сходимости F_n к некоторой k-мерной функции распределения F состоит в том, что $F_{\lambda n}$ имеет предел для любого вектора λ.

Приведенная теорема ценна тем, что с ее помощью сходимость векторов сводится к сходимости линейных комбинаций их координат, т.е. к сходимости обычных случайных величин, рассмотренных ранее. Однако она не дает возможности непосредственно указать предельное распределение. Это можно сделать с помощью следующей теоремы.

Теорема о многомерной сходимости. Пусть F_n и $F_{\lambda n}$ — те же, что в предыдущей теореме. Пусть F — совместная функция распределения k-мерного случайного вектора $(X_1,\ldots,X_k)$. Если функция распределения $F_{\lambda n}$ сходится при росте объема выборки к функции распределения F_λ для любого вектора λ, где F_λ — функция распределения линейной комбинации $\lambda_1 X_1+\ldots+\lambda_k X_k$, то F_n сходится к F.

Здесь сходимость F_n к F означает, что для любого k-мерного вектора $(x_1,\ldots,x_k)$ такого, что функция распределения F непрерывна в $(x_1,\ldots,x_k)$, числовая последовательность $F_n(x_1,...,x_k)$ сходится при росте n к числу $F(x_1,...,x_k)$. Другими словами, сходимость функций распределения понимается точно так же, как при обсуждении предельных теорем для случайных величин выше. Приведем многомерный аналог этих теорем.

Многомерная центральная предельная теорема [3]. Рассмотрим независимые одинаково распределенные k-мерные случайные вектора

$$U_n' = (U_{1n},...,U_{kn}), \quad n = 1,2,...,$$

где штрих обозначает операцию транспонирования вектора. Предположим, что случайные вектора U_n имеют моменты первого и второго порядка, т.е.

$$M(U_n) = \mu,\ D(U_n) = \Sigma,$$

где μ — вектор математических ожиданий координат случайного вектора, Σ — его ковариационная матрица. Введем последовательность средних арифметических случайных векторов:

$$\overline{U}_n = (\overline{U}_{1n},...,\overline{U}_{kn}), \quad n = 1,2,..., \quad \overline{U}_{in} = \frac{1}{n}\sum_{j=1}^{n} U_{ij}.$$

Тогда случайный вектор $\sqrt{n}(\overline{U}_n - \mu)$ имеет асимптотическое k-мерное нормальное распределение $N_k(0,\Sigma)$, т.е. он асимптотически распределен так же, как k-мерная нормальная величина с нулевым математическим ожиданием, ковариационной Σ и плотностью

$$N_k(u\,|\,0,\Sigma) = (2\pi)^{-k/2}\,|\,\Sigma\,|^{-1/2}\exp\left\{-\frac{1}{2}u'\Sigma^{-1}u\right\}.$$

Здесь $|\Sigma|$ — определитель матрицы Σ. Другими словами, распределение случайного вектора $\sqrt{n}(\overline{U}_n - \mu)$ сходится к k-мерному нормальному распределению с нулевым математическим ожиданием и ковариационной матрицей Σ.

Напомним, что многомерным нормальным распределением с математическим ожиданием μ и ковариационной матрицей Σ называется распределение, имеющее плотность

$$N_k(u\,|\,\mu,\Sigma) = (2\pi)^{-k/2}\,|\,\Sigma\,|^{-1/2}\exp\left\{-\frac{1}{2}[(u-\mu)'\Sigma^{-1}(u-\mu)]\right\}.$$

Многомерная центральная предельная теорема показывает, что распределения сумм независимых одинаково распределенных случайных векторов при большом числе слагаемых хорошо приближаются с помощью нормальных распределений, имеющих такие же первые два момента (вектор математических ожиданий координат случайного вектора и его корреляционную матрицу), как и исходные вектора. От одинаковой распределенности можно отказаться, но это потребует некоторого усложнения символики. В целом из теоремы о многомерной сходимости вытекает, что многомерный случай ничем принципиально не отличается от одномерного.

Пример. Пусть $X_1, \ldots X_n, \ldots$ — независимые одинаково распределенные случайные величины. Рассмотрим k-мерные независимые одинаково распределенные случайные вектора

$$U_n' = (X_n, X_n^2, X_n^3, \ldots, X_n^k), \quad n = 1,2,\ldots$$

Их математическое ожидание — вектор теоретических начальных моментов, а ковариационная матрица составлена из соответствующих центральных моментов. Тогда $\overline{U}_n$ — вектор выборочных центральных моментов. Многомерная центральная предельная теорема утверждает, что $\overline{U}_n$ имеет асимптотически нормальное распределение. Как вытекает из теорем о наследовании сходимости и о линеаризации (см. ниже), из распределения $\overline{U}_n$ можно вывести распределения различных функций от выборочных начальных моментов. А поскольку центральные моменты выражаются через начальные моменты, то аналогичное утверждение верно и для них.

14.3. Теоремы о наследовании сходимости

Суть проблемы наследования сходимости. Пусть распределения случайных величин X_n при $n \to \infty$ стремятся к распределению случайной величины X. При каких функциях f можно утверждать, что распределения случайных величин $f(X_n)$ сходятся к распределению $f(X)$, т.е. наследуется сходимость?

Хорошо известно, что для непрерывных функций f сходимость наследуется [3]. Однако в статистических методах используются различные обобщения этого утверждения. Необходимость обобщений связана с тремя обстоятельствами.

1. Статистические данные могут моделироваться не только случайными величинами, но и случайными векторами, случайными множествами, случай-

ными элементами произвольной природы (т.е. функциями на вероятностном пространстве со значениями в произвольном множестве).

2. Переход к пределу должен рассматриваться не только для случая безграничного возрастания объема выборки, но и в более общих случаях. Например, если в постановке статистической задачи участвуют несколько выборок объемов $n(1)$, $n(2)$, … , $n(k)$, то вполне обычным является предположение о безграничном росте всех этих объемов (что можно описать и как $\min\{n(1), n(2), \ldots, n(k)\} \to \infty$).

3. Функция f не обязательно является непрерывной. Она может иметь разрывы. Кроме того, она может зависеть от параметров, по которым происходит переход к пределу. Например, может зависеть от объемов выборок. Например, в главе 5 понадобилось рассмотреть функцию $f = f(n(1), n(2), \ldots , n(k))$.

Расстояние Прохорова и сходимость по направленному множеству. Введем необходимые для дальнейшего изложения понятия.

Расстояние (метрика) Прохорова. Пусть C — некоторое пространство, A — его подмножество, d — метрика в C. Введем понятие ε-окрестности множества A в метрике d:

$$S(A,\varepsilon) = \{x\in C: d(A,x) < \varepsilon\}.$$

Таким образом, ε-окрестность множества A — это совокупность всех точек пространства C, отстоящих от A не более чем на положительное число ε. При этом расстояние от точки x до множества A — это точная нижняя грань расстояний от x до точек множества A, т.е.

$$d(A,x) = \inf\{d(x,y): y\in A\}.$$

Пусть P_1 и P_2 — две вероятностные меры на C (т.е. распределения двух случайных элементов со значениями в C). Пусть D_{12} — множество чисел $\varepsilon > 0$ таких, что

$$P_1(A) \le P_2(S(A,\varepsilon)+\varepsilon$$

для любого замкнутого подмножества A пространства C. Пусть D_{21} — множество чисел $\varepsilon > 0$ таких, что

$$P_2(A) \le P_1(S(A,\varepsilon)+\varepsilon$$

для любого замкнутого подмножества A пространства C. Расстояние Прохорова $L(P_1,P_2)$ между вероятностными мерами (его можно рассматривать и как расстояние между случайными элементами с распределениями P_1 и P_2 соответственно) вводится формулой

$$L(P_1,P_2) = \max(\inf D_{12}, \inf D_{21}).$$

С помощью метрики Прохорова формализуется понятие сходимости распределений случайных элементов в произвольном пространстве.

Расстояние $L(P_1,P_2)$ введено академиком РАН Юрием Васильевичем Прохоровым в середине XX в. и широко используется в современной теории вероятностей.

Сходимость по направленному множеству [4, с. 95–96]. Бинарное отношение $\geq$ (упорядочение), заданное на множестве B, называется направлением на нем, если B не пусто и:

(а) если m, n и p — такие элементы множества B, что $m \geq n$ и $n \geq p$, то $m \geq p$;

(б) $m \geq m$ для любого m из B;

(в) если m и n принадлежат B, то найдется элемент p из B такой, что $p \geq m$ и $p \geq n$.

Направленное множество — это пара $(B, \geq)$, где $\geq$ — направление на множестве B. Направленностью (или «последовательностью по направленному множеству») называется пара $(f, \geq)$, где f — функция, $>$ — направление на ее области определения. Пусть f: $B \to Y$, где Y — топологическое пространство. Направленность $(f, \geq)$ сходится в топологическом пространстве Y к точке y_0, если для любой окрестности U точки y_0 найдется p из B такое, что $f(q) \in U$ при любом $q \geq p$. В таком случае говорят также о сходимости по направленному множеству.

Пусть $B = \{(n(1), n(2), \dots , n(k))\}$ — совокупность векторов, каждый из которых составлен из объемов k выборок. Пусть

$$(n(1), n(2), \dots , n(k)) \geq (n_1(1), n_1(2), \dots , n_1(k))$$

тогда и только тогда, когда $n(i) \geq n_1(i)$ при всех $i = 1, 2, \dots, k$. Тогда $(B, \geq)$ — направленное множество, сходимость по которому эквивалентна сходимости при $\min \{n(1), n(2), \dots , n(k)\} \to \infty$.

Чтобы охватить различные частные случаи, целесообразно предельные теоремы формулировать в терминах сходимости по направленному множеству. Будем писать $B = \{\alpha\}$. Пусть запись $\alpha\to\infty$ обозначает переход к пределу по направленному множеству.

Формулировка проблемы наследования сходимости. Пусть случайные элементы X_α со значениями в пространстве C сходятся при $\alpha\to\infty$ к случайному элементу X, где через $\alpha\to\infty$ обозначен переход к пределу по направленному множеству. Сходимость случайных элементов означает, что $L(X_\alpha, X) \to 0$ при $\alpha\to\infty$, где L — метрика Прохорова в пространстве C.

Пусть f_α: $C \to Y$ — некоторые функции. Какие условия надо на них наложить, чтобы из $L(X_\alpha, X) \to 0$ вытекало, что $L_1(f_\alpha(X_\alpha), f_\alpha(\mathrm{X})) \to 0$ при $\alpha\to\infty$, где L_1 — метрика Прохорова в пространстве Y? Другими словами, какие условия на функции f_α: $C \to Y$ гарантируют наследование сходимости?

В работах [5, 6] найдены необходимые и достаточные условия на функции f_α: $C \to Y$, гарантирующие наследование сходимости. Описанию этих условий посвящена оставшаяся часть подраздела.

Приведем для полноты изложения строгие формулировки математических предположений (в дальнейшем никому, кроме профессиональных математиков, не понадобятся).

Математические предположения. Пусть C и $У$ — полные сепарабельные метрические пространства, Пусть выполнены обычные предположения измеримости: X_α и X — случайные элементы C, $f_\alpha(X_\alpha)$ и $f_\alpha(X)$ — случайные элементы в $У$, рассматриваемые ниже подмножества пространств C и $У$ лежат в соответствующих σ–алгебрах измеримых подмножеств, и т.д.

Понадобятся некоторые *определения*. Разбиение $T_n = \{C_{1\mathrm{n}}, C_{2n}, \dots , C_{nn}\}$ пространства C — это такой набор подмножеств C_j, $j = 1, 2, \dots , n$, этого пространства, что пересечение любых двух из них пусто, а объединение совпадает с C. Диаметром $diam(A)$ подмножества A множества C называется точная верхняя грань расстояний между элементами A, т.е.

$$diam(A) = \sup \{d(x,y), \mathrm{x}\in A, y\in A\},$$

где $d(x,y)$ — метрика в пространстве C. Обозначим ∂A границу множества A, т.е. совокупность точек x таких, что любая их окрестность $U(x)$ имеет непустое пересечение как с A, так и с $C\backslash A$. Колебанием $\delta(f, B)$ функции f на множестве B называется $\delta(f, B) = \sup \{|f(x) — f(y)|, x\in B, y\in B\}$.

Достаточное условие для наследования сходимости. Пусть $L(X_\alpha, X) \to 0$ при $\alpha \to \infty$. Пусть существует последовательность T_n разбиений пространства C такая, что $P(X \in \partial A) = 0$ для любого A из T_n и, основное условие, для любого $\varepsilon > 0$

$$m_\varepsilon(\alpha, n) = \sum P(X \in A) \to 0 \qquad (1)$$

при $n \to \infty$ и $\alpha \to \infty$, где сумма берется по всем тем A из T_n, для которых колебание функции f_α на A больше ε, т.е. $\delta(f_\alpha, \text{A}) > \varepsilon$. Тогда $L_1(f_\alpha(X_\alpha), f_\alpha(X)) \to 0$ при $\alpha \to \infty$.

Необходимое условие для наследования сходимости. Пусть $У$ — конечномерное линейное пространство, $У = R^k$. Пусть случайные элементы $f_\alpha(X)$ асимптотически ограничены по вероятности при $\alpha \to \infty$, т.е. для любого $\varepsilon > 0$ существуют число $S(\varepsilon)$ и элемент направленного множества $\alpha(\varepsilon)$ такие, что $P(\|f_\alpha(X)\| > S(\varepsilon)) < \varepsilon$ при $\alpha \geq \alpha(\varepsilon)$, где $\|f_\alpha(X)\|$ — норма (длина) вектора $f_\alpha(X)$. Пусть существует последовательность T_n разбиений пространства C такая, что

$$\lim_{n \to \infty} \max\{diam(C_{jn}),\ C_{jn} \in T_n\} = 0,$$

т.е. последовательность T_n является безгранично измельчающейся. Самое существенное — пусть условие (1) не выполнено для последовательности T_n. Тогда существует последовательность случайных элементов X_α такая, что $L(X_\alpha, X) \to 0$ при $\alpha \to \infty$, но $L_1(f_\alpha(X_\alpha), f_\alpha(X))$ не сходится к 0 при $\alpha \to \infty$.

Несколько огрубляя, можно сказать, что *условие (1) является необходимым и достаточным для наследования сходимости*.

Пример 1. Пусть C и $У$ — конечномерные линейные пространства, функции f_α не зависят от α, т.е. $f_\alpha \equiv f$, причем функция f ограничена. Тогда условие (1) эквивалентно требованию интегрируемости по Риману-Стилтьесу функции f по мере $G(A) = P(X \in A)$. В частности, условие (1) выполнено для непрерывной функции f.

В конечномерных пространствах C вместо сходимости $L(X_\alpha, X) \to 0$ при $\alpha \to \infty$ можно говорить о слабой сходимости функций распределения случайных векторов X_α к функции распределения случайного вектора X. Речь идет о «сходимости по распределению», т.е. о сходимости во всех точках непрерывности функции распределения случайного вектора X. В этом случае разбиения могут состоять из многомерных параллелепипедов [5, глава 2].

Пример 2. Полученные выше результаты дают обоснование для рассуждений типа следующего (ср., например, утверждения в главе 5 выше). Пусть по двум независимым выборкам объемов m и n соответственно построены статистики X_m и Y_n. Пусть известно, что распределения этих статистик сходятся при безграничном росте объемов выборок к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Пусть $a(m, n)$ и $b(m, n)$ — некоторые коэффициенты. Тогда согласно результатам примера 1 распределение случайной величины $Z(m, n) = a(m, n)X_m + b(m, n)Y_n$ сближается с распределением нормально распределенной случайной величины с математическим ожиданием 0 и дисперсией $a^2(m, n) + b^2(m, n)$. Если же $a^2(m, n) + b^2(m, n) =$ $= 1$, например,

$$a(m,n) = \sqrt{\frac{m}{m+n}}, \quad b(m,n) = \sqrt{\frac{n}{m+n}},$$

то распределение $Z(m, n)$ сходится при безграничном росте объемов выборок к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1.

14.4. Метод линеаризации

При разработке статистических методов часто возникает следующая задача [3, с. 338]. Имеется последовательность k-мерных случайных векторов $X_n = (X_{1n}, X_{2n}, \dots, X_{kn})$, $n = 1, 2, \dots$, такая, что $X_n \to a = (a_1, a_2, \dots, a_k)$ при $n \to \infty$, и последовательность функций f_n: $R^k \to R^1$. Требуется найти распределение случайной величины $f_n(X_n)$.

Основная идея — рассмотреть главный линейный член функции f_n в окрестности точки a. Из математического анализа известно, что

$$f_n(X_n) - f_n(a) = \sum_{j=1}^{k} \frac{\partial f_n(a)}{\partial x_j}(X_{jn} - a_j) + O_n(\| X_n - a \|^2),$$

где остаточный член является бесконечно малой величиной более высокого порядка малости, чем линейный член. Таким образом, произвольная функция может быть заменена на линейную функцию от координат случайного вектора. Эта замена проводится с точностью до бесконечно малых более высокого порядка. Конечно, должны быть выполнены некоторые математические условия регулярности. Например, функции f_n должны быть дважды непрерывно дифференцируемы в окрестности точки a.

Если вектор X_n является асимптотически нормальным с математическим ожиданием a и ковариационной матрицей $\sum/n$, где $\sum = \|\sigma_{ij}\|$, причем $\sigma_{ij} = nM(X_i - - a_i)(X_j - a_j)$, то линейная функция от его координат также асимптотически нормальна. Следовательно, при очевидных условиях регулярности $f_n(X_n)$ — асимптотически нормальная случайная величина с математическим ожиданием $f_n(a)$ и дисперсией

$$\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{\partial f_n(a)}{\partial x_i}\frac{\partial f_n(a)}{\partial x_j}\sigma_{ij}.$$

Для практического использования асимптотической нормальности $f_n(X_n)$ остается заменить неизвестные моменты a и $\sum$ на их оценки. Например, если X_n — это среднее арифметическое независимых одинаково распределенных случайных векторов, то a можно заменить на X_n, а $\sum$ — на выборочную ковариационную матрицу.

Пример. Пусть Y_1, Y_2, … , Y_n — независимые одинаково распределенные случайные величины с математическим ожиданием a и дисперсией σ^2. В качестве X_n ($k = 1$) рассмотрим выборочное среднее арифметическое

$$\overline{Y} = \frac{Y_1 + Y_2 + \ldots + Y_n}{n}.$$

Как известно, в силу закона больших чисел $\overline{Y} \to a = M(У)$. Следовательно, для получения распределений функций от выборочного среднего арифметического можно использовать метод линеаризации. В качестве примера рассмотрим $f_n(y) = f(y) = y^2$. Тогда

$$\left(\overline{Y}\right)^2 - a^2 = \frac{df(a)}{dy}\left(\overline{Y} - a\right) + O\left(\left(\overline{Y} - a\right)^2\right) = 2a\left(\overline{Y} - a\right) + O\left(\left(\overline{Y} - a\right)^2\right).$$

Из этого соотношения следует, что с точностью до бесконечно малых более высокого порядка

$$\left(\overline{Y}\right)^2 = a^2 + 2a\left(\overline{Y} - a\right).$$

Поскольку в соответствии с Центральной Предельной Теоремой выборочное среднее арифметическое является асимптотически нормальной случайной величиной с математическим ожиданием a и дисперсией σ^2/n, то квадрат этой статистики является асимптотически нормальной случайной величиной с математическим ожиданием a^2 и дисперсией $4a^2\sigma^2/n$. Для практического использования может оказаться полезной замена параметров (асимптотического нормального распределения) на их оценки, а именно, математического ожидания — на $(\overline{Y})^2$, а дисперсии — на $4(\overline{Y})^2 s^2/n$, где s^2 — выборочная дисперсия.

Большое внимание (целая глава!) уделено методу линеаризации в классическом учебнике Е.С. Вентцель [7].

14.5. Принцип инвариантности

Пусть $Y_1, Y_2, \dots, Y_n$ — независимые одинаково распределенные случайные величины с непрерывной функцией распределения $F(x)$. Многие используемые в статистических методах функции от результатов наблюдений выражаются через эмпирическую функцию распределения $F_n(x)$. К ним относятся статистики Колмогорова, Смирнова, омега-квадрат. Отметим, что и другие статистики выражаются через эмпирическую функцию распределения, например:

$$\overline{Y} = \int_{-\infty}^{+\infty} x dF_n(x).$$

Полезным является преобразование Н.В. Смирнова $t = F(x)$. Тогда независимые случайные величины $Z_j = F(Y_j)$, $j = 1, 2, \dots, n$, имеют равномерное распределение на отрезке [0; 1]. Рассмотрим построенную по ним эмпирическую функцию распределения $F_n(t)$, $0 \le t \le 1$. *Эмпирическим процессом* называется случайный процесс

$$\xi_n(t) = \sqrt{n}(F_n(t) - t).$$

Рассмотрим критерии проверки согласия функции распределения выборки с фиксированной функцией распределения $F(x)$. Статистика критерия Колмогорова записывается в виде

$$K_n = \sup_{0 \le t \le 1} |\xi_n(t)|,$$

статистика критерия Смирнова — это

$$S_n = \sup_{0 \le t \le 1} \xi_n(t),$$

а статистика критерия омега-квадрат (Крамера — Мизеса — Смирнова) имеет вид

$$\omega_n^2 = \int_0^1 \xi_n^2(t)dt .$$

Случайный процесс $\xi_n(t)$ имеет нулевое математическое ожидание и ковариационную функцию $M\xi_n(s)\xi_n(t) = \min(s,t) - st$. Рассмотрим гауссовский случайный процесс $\xi(t)$ с такими же математическим ожиданием и ковариационной функцией. Он называется броуновским мостом. (Напомним, что гауссовским процесс именуется потому, что вектор $(\xi(t_1), \xi(t_2), \ldots, \xi(t_k))$ имеет многомерное нормальное распределение при любых наборах моментов времени $t_1, t_2, \ldots, t_k$.).

Пусть f — функционал, определенный на множестве возможных траекторий случайных процессов. *Принцип инвариантности* [1] состоит в том, что последовательность распределений случайных величин $f(\xi_n)$ сходится при $n \to \infty$ к распределению случайной величины $f(\xi)$. Сходимость по распределению обозначим символом =>. Тогда принцип инвариантности кратко записывается так: $f(\xi_n)$ => $f(\xi)$. В частности, согласно принципу инвариантности, статистика Колмогорова и статистика омега-квадрат сходятся по распределению к распределениям соответствующих функционалов от случайного процесса ξ:

$$K_n = \sup_{0 \le t \le 1} |\xi_n(t)| \Rightarrow \sup_{0 \le t \le 1} |\xi(t)|, \ \omega_n^2 = \int_0^1 \xi_n^2(t)dt \Rightarrow \int_0^1 \xi^2(t)dt .$$

Таким образом, от проблем прикладной статистики сделан переход к теории случайных процессов. Методами этой теории найдены распределения случайных величин

$$\sup_{0 \le t \le 1} |\xi(t)|, \ \int_0^1 \xi^2(t)dt .$$

Принцип инвариантности — инструмент получения предельных распределений функций от результатов наблюдений, используемых в прикладной статистике.

Обоснование принципу инвариантности может быть дано на основе теории сходимости вероятностных мер в функциональных пространствах [8]. Более простой подход, позволяющий к тому же получать необходимые и достаточные условия в предельной теории статистик интегрального типа (принцип инвариантности к ним нельзя применить), рассмотрен в главе 4.3.

Почему «принцип инвариантности» так назван? Обратим внимание, что предельные распределения рассматриваемых статистик не зависят от их функции распределения $F(x)$. Другими словами, предельное распределение инвариантно относительно выбора $F(x)$.

В более широком смысле термин «принцип инвариантности» применяют тогда, когда предельное распределение не зависит от тех или иных характеристик исходных распределений [1]. В этом смысле наиболее известный «принцип инвариантности» — это Центральная Предельная Теорема, поскольку предельное стандартное нормальное распределение — одно и то же для всех возможных распределений независимых одинаково распределенных слагаемых (лишь бы слагаемые имели конечные математическое ожидание и дисперсию).

14.6. Устойчивость выводов и принцип уравнивания погрешностей

Устойчивость математических моделей. Проблемам познания, в том числе в технических исследованиях, естественно-научных и социально-экономических областях, посвящено огромное количество работ. Однако это не значит, что обо всем в этой области уже все сказано. А о некоторых положениях целесообразно говорить еще и еще раз, пока они ни станут общеизвестными.

В идеале каждую модель порождения и анализа данных следовало бы рассматривать как аксиоматическую теорию. В этом идеальном случае создание и использование модели происходит в соответствии с известной триадой «практика — теория — практика». А именно, сначала вводятся некоторые математические объекты, соответствующие интересующим исследователя реальным объектам, и на основе представлений о свойствах реальных объектов формулируются необходимые для успешного моделирования свойства математических объектов, которые и принимаются в качестве аксиом. Затем аксиоматическая теория развивается как часть математики, вне связи с представлениями о реальных объектах. На заключительном этапе полученные в математической теории результаты интерпретируются содержательно. Получаются утверждения о реальных объектах, являющиеся следствиями тех и только тех их свойств, которые ранее были аксиоматизированы.

После построения математической модели реального явления или процесса встает вопрос об ее адекватности. Иногда ответ на этот вопрос может дать эксперимент. Рассогласование модельных и экспериментальных данных следует интерпретировать как признак неадекватности некоторых из принятых аксиом. Однако для проверки адекватности социально-экономических моделей зачастую невозможно поставить решающий эксперимент в отличие, скажем, от физических моделей. С другой стороны, для одного и того же явления или процесса, как правило, можно составить много возможных моделей, если угодно, много разновидностей одной базовой модели. Поэтому необходимы какие-то дополнительные условия, которые позволяли бы их множества возможных моделей и эконометрических методов анализа данных выбрать наиболее подходящие. В качестве одного из подобных условий выдвигается требование *устойчивости* модели и метода анализа данных относительно допустимых отклонений исходных данных и предпосылок модели или условий применимости метода.

Отметим, что в большинстве случаев исследователей и практических работников интересуют не столько сами модели и методы, сколько решения, которые с их помощью принимаются. Ведь модели и методы для того и разрабатываются, чтобы подготавливать решения. Вместе с тем очевидно, что решения, как правило, принимаются в условиях неполноты информации. Так, любые числовые параметры известны лишь с некоторой точностью. Введение в рассмотрение возможных неопределенностей исходных данных требует каких-то заключений относительно устойчивости принимаемых решений по отношению к этим допустимым неопределенностям.

Введем основные понятия согласно монографии [5]. Будем считать, что имеются *исходные данные*, на основе которых принимаются *решения*. Способ переработки (отображения) исходных данных в решение назовем *моделью*. Таким образом, с общей точки зрения модель — это функция, переводящая исходные данные в решение, т.е. способ перехода значения не имеет. Очевидно, любая рекомендуемая для практического использования модель должна быть исследована *на устойчивость* относительно допустимых отклонений исходных данных. Укажем некоторые возможные применения результатов подобного исследования:

- заказчик научно-исследовательской работы получает представление о точности предлагаемого решения;
- удается выбрать из многих моделей наиболее адекватную;
- по известной точности определения отдельных параметров модели удается указать необходимую точность нахождения остальных параметров;

- переход к случаю «общего положения» позволяет получать более сильные с математической точки зрения результаты.

Примеры. По каждому из четырех перечисленных возможных применений в [5, 9] приведены различные примеры. В статистических методах точность предлагаемого решения связана с разбросом исходных данных и с объемом выборки. Выбору наиболее адекватной модели посвящены многие рассмотрения в главах 5 и 6, связанные с обсуждением моделей однородности и регрессии. Использование рационального объема выборки в статистике интервальных данных [9] исходит из принципа уравнивания погрешностей. Этот принцип основан на том, что по известной точности определения отдельных параметров модели удается указать необходимую точность нахождения остальных параметров. Другим примером применения принципа уравнивания погрешностей является нахождение необходимой точности оценивания параметров в моделях логистики, рассмотренных в главе 5 монографии [5]. Наконец, переходом к случаю «общего положения» в прикладной статистике является, в частности, переход к непараметрическим методам, необходимый из-за невозможности обосновать принадлежность результатов наблюдений к тем или иным параметрическим семействам.

Специалисты по математическому моделированию и теории управления считают устойчивость одной из важных характеристик технических, социально-экономических, медицинских и иных моделей. Достаточно глубокие исследования ведутся по ряду направлений.

Первоначальное изучение влияния малого изменения одного параметра обычно называют *анализом чувствительности*. Оно описывается значением частной производной. Если модель задается дифференцируемой функцией, то итог анализа чувствительности — вектор значений частных производных в анализируемой точке.

Теория устойчивости решений дифференциальных уравнений развивается по крайней мере с XIX в. [10]. Выработаны соответствующие понятия — устойчивость по Ляпунову, корректность, доказаны глубокие теоремы. Для решения некорректных задач академиком АН СССР А.Н. Тихоновым в начале 1960-х годов был предложен метод регуляризации. Модели явлений и процессов, выражаемые с помощью дифференциальных уравнений, могут быть исследованы на устойчивость путем применения хорошо разработанного математического аппарата.

Вопросы устойчивости изучались практически во всех направлениях прикладных математических методов — и в математическом программировании, и в теории массового обслуживания (теории очередей), и в эколого-экономических моделях, и в различных областях эконометрики.

Общая схема устойчивости. Прежде чем переходить к конкретным постановкам, обсудим «общую схему устойчивости», дающую понятийную базу для обсуждения проблем устойчивости в различных предметных областях.

Определение 1. Общей схемой устойчивости называется математический объект $\{\mathrm{A}, \mathrm{B}, d, f, \mathrm{E}\}$.

Здесь A — множество, интерпретируемое как пространство исходных данных; B — множество, называемое пространством решений. Однозначное отображение $f:\mathrm{A}\to\mathrm{B}$ называется моделью. Об этих трех составляющих общей схемы устойчивости уже шла речь выше.

Оставшиеся два понятия нужны для уточнения понятий близости в пространстве исходных данных и пространстве решений. Подобные уточнения могут быть сделаны разными способами. Самое «слабое» уточнение — на языке топологических пространств. Тогда возможны качественные выводы (сходится — не сходится), но не количественные расчеты. Самое «сильное» уточнение — на языке метрических пространств. Промежуточный вариант — используются показатели различия (отличаются от метрик тем, что не обязательно выполняются неравенства треугольника) или вводимые ниже понятия.

Пусть d-показатель устойчивости, т.е. неотрицательная функция, определенная на подмножествах $У$ множества B и такая, что из $Y_1 \subseteq Y_2$ вытекает $d(Y_1) \le d(Y_2)$. Часто показатель устойчивости $d(Y)$ определяется с помощью метрики, псевдометрики или показателя различия (меры близости) ρ как диаметр множества $У$, т.е.

$$d(Y) = \sup\{\rho(y_1, y_2), y_1 \in Y, y_2 \in Y\}.$$

Таким образом, говоря попросту, в пространстве решений с помощью показателя устойчивости вокруг образа исходных данных может быть сформирована система окрестностей. Но сначала надо такую систему сформировать в пространстве исходных данных.

Пусть $\mathrm{E} = \{E(x,\alpha), x \in \mathrm{A}, \alpha \in \Theta\}$ — совокупность допустимых отклонений. Система подмножеств множества A такая, что каждому элементу множества исходных данных $x \in \mathrm{A}$ и каждому значению параметра α из некоторого множества параметров Θ соответствует подмножество $E(x,\alpha)$ множества исходных данных. Оно называется множеством допустимых отклонений в точке x при значении параметра, равном α. Наглядно можно представить себе, что вокруг точки x взята окрестность радиуса α.

Определение 2. Показателем устойчивости в точке x при значении параметра, равном α, называется число

$$\beta(x, E(x,\alpha)) = d(f(E(x,\alpha))).$$

Другими словами, это — диаметр образа множества допустимых колебаний при рассматриваемом в качестве модели отображении. Очевидно, что этот показатель устойчивости зависит как от исходных данных, так и от диаметра множества возможных отклонений в исходном пространстве. Для непрерывных функций показатель устойчивости обычно называется модулем непрерывности.

Естественно посмотреть, насколько сузится образ окрестности возможных отклонений при максимально возможном сужении этой окрестности.

Определение 3. Абсолютным показателем устойчивости в точке x называется число

$$\beta(x, \mathrm{E}) = \inf\{\beta(x, E(x,\alpha)), \alpha \in \Theta\}.$$

Если функция f непрерывна, а окрестности — именно те, о которых идет речь в математическом анализе, то максимальное сужение означает сужение к точке и абсолютный показатель устойчивости равен 0. Но в теории измерений и статистике интервальных данных мы сталкиваемся с совсем иными ситуациями. В теории измерений окрестностью исходных данных являются все те вектора, что получаются из исходного путем преобразования координат с помощью допустимого преобразования шкалы, а допустимое преобразование шкалы берется из соответствующей группы допустимых преобразований. В статистике интервальных данных под окрестностью исходных данных естественно понимать — при описании выборки — куб с ребрами 2Δ и центром в исходном векторе. И в том, и в другом случае максимальное сужение не означает сужение к точке.

Естественным является желание ввести характеристики устойчивости на всем пространстве. Не вдаваясь в математические тонкости (см. о них монографию [5]), рассмотрим меру μ на пространстве A такую, что мера всего пространства равна 1 (т.е. $\mu(\mathrm{A}) = 1$).

Определение 4. Абсолютным показателем устойчивости на пространстве исходных данных A по мере μ называется число

$$\gamma(\mu) = \int_{\mathrm{A}} \beta(x, \mathrm{E})\, d\mu.$$

Здесь имеется в виду интеграл Лебега. Интегрирование проводится по (абстрактному) пространству исходных данных A по мере μ. Естественно, должны быть выполнены некоторые внутриматематические условия. Читателю, незнакомому с интегрированием по Лебегу, достаточно мысленно заменить в предыдущей формуле интеграл на сумму (а пространство A считать конечным, хотя и состоящим из большого числа элементов).

Определение 5. Максимальным абсолютным показателем устойчивости называется

$$\gamma = \sup\{\beta(x, \mathrm{E}), x \in \mathrm{A}\}.$$

Легко видеть, что $\gamma = \sup \gamma(\mu)$, где супремум берется по всем описанным выше мерам.

Итак, построена иерархия показателей устойчивости математических моделей реальных явлений и процессов. Она с успехом использовалась в различных исследованиях, подробно развивалась, в частности, в монографии [5]. Приведем еще одно полезное определение.

Определение 6. Модель f называется абсолютно ε-устойчивой, если $\gamma \le \varepsilon$, где γ — максимальный абсолютный показатель устойчивости.

Пример. Если показатель устойчивости формируется с помощью метрики ρ, совокупность допустимых отклонений E — это совокупность всех окрестностей всех точек пространства исходных данных A, то 0-устойчивость модели *f* эквивалентна непрерывности модели *f* на множестве A.

Основная проблема в общей схеме устойчивости — проверка ε-устойчивости данной модели f относительно данной системы допустимых отклонений E.

Часто оказываются полезными следующие два обобщения основной проблемы.

Проблема А (характеризации устойчивых моделей). Даны пространство исходных данных A, пространство решений B, показатель устойчивости *d*, совокупность допустимых отклонений E и неотрицательное число ε. Описать достаточно широкий класс ε-устойчивых моделей *f*. Или: найти все ε-устойчивые модели среди моделей, обладающих данными свойствами, т.е. входящих в данное множество моделей.

Проблема Б (характеризации систем допустимых отклонений). Даны пространство исходных данных A, пространство решений B, показатель устойчивости *d*, модель *f* и неотрицательное число ε. Описать достаточно широкий класс систем допустимых отклонений E, относительно которых модель f *является* ε-устойчивой. Или: найти все такие системы допустимых отклонений E среди совокупностей допустимых отклонений, обладающих данными свойствами, т.е. входящих в данное множество совокупностей допустимых отклонений.

Ясно, что проблемы А и Б можно рассматривать не только для показателя устойчивости γ, но и для других только что введенных показателей устойчивости, а именно, $\gamma(\mu), \beta(x, \mathrm{E}), \beta(x, E(x,\alpha))$.

Язык общей схемы устойчивости позволяет описывать конкретные задачи специализированных теорий устойчивости в различных областях исследований, выделять в основные элементы в них, ставить проблемы типа А и Б. В частности, на этом языке легко формулируются задачи теории устойчивости решений дифференциальных уравнений, теории робастности статистических процедур (см. главу 3.4), проблемы адекватности теории измерений, достигаемой точности расчетов в статистике интервальных данных и в логистике (см. монографию [5]), и т.д.

Для примера рассмотрим определение устойчивости по Ляпунову решения $\varphi(t,x)$ нормальной автономной системы дифференциальных уравнений $\dot{y} = g(y)$ с начальными условиями $\varphi(0,x) = x$. Здесь пространство исходных данных A — конечномерное евклидово пространство, множество допустимых отклонений $E(x,\alpha)$ – окрестность радиуса α точки $x \in \mathrm{A}$, пространство решений B — множество функций на луче $[0;+\infty)$ с метрикой

$$\rho(y_1, y_2) = \sup_{t \geq 0} | y_1(t) - y_2(t) | .$$

Модель f — отображение, переводящее начальные условия x в решение системы дифференциальных уравнений с этими начальными условиями $\varphi(t,x)$.

В терминах общей схемы устойчивости положение равновесия a называется *устойчивым по Ляпунову*, если $\beta(a, \mathrm{E}) = 0$. Для формулировки определения асимптотической устойчивости по Ляпунову надо ввести в пространстве решений B псевдометрику

$$\rho_1(y_1, y_2) = \overline{\lim_{t \to \infty}} \, | y_1(t) - y_2(t) | .$$

Положение равновесия a называется асимптотически устойчивым, если $\beta_1(a, E(a,\varepsilon)) = 0$ для некоторого $\varepsilon > 0$, где показатель устойчивости β_1 рассчитан с использованием псевдометрики ρ_1.

Таким образом, общая схема устойчивости естественным образом включает в себя классические понятия теории устойчивости по Ляпунову. Вместе с тем стоит отметить, что эта схема дает общий подход к различным проблемам устойчивости. Она дает систему понятий, которые в каждом конкретном случае должны приспосабливаться к решаемой задаче.

До настоящего момента для определенности речь шла о допустимых отклонениях в пространстве исходных данных. Часто оказывается необходимым говорить и об отклонениях от предпосылок модели. С чисто формальной точки зрения для этого достаточно расширить понятие «исходные данные» до пары (x, f), т.е. включив «прежнюю» модель в качестве второго элемента пары. Все остальные определения остаются без изменения. Теперь отклонения в пространстве решений вызываются не только отклонениями в исходных данных x, но и отклонениями от предпосылок модели, т.е. отклонениями f. Это соображение нам понадобилось в разделе 3.4, посвященном робастности статистических процедур.

Устойчивость по отношению к объему выборки. Различные асимптотические постановки в прикладной статистике также естественно рассматривать как задачи устойчивости. Если при безграничном возрастании объема выборки некоторая величина стремится к пределу, то в терминах общей схемы устойчивости это означает, что она 0-устойчива в соответствующей псевдометрике (см. выше обсуждение асимптотической устойчивости по Ляпунову). С содержательной точки зрения употребление термина «устойчивость» в такой ситуации представляется вполне оправданным, поскольку рассматриваемая величина мало меняется при изменении объема выборки.

Рассмотрим проблему и методы оценки близости предельных распределений статистик и распределений, соответствующих конечным объемам выборок. При каких объемах выборок уже можно пользоваться предельными распределениями? Каков точный смысл термина «можно» в предыдущей фразе? Основное внимание уделяется переходу от точных формул допредельных распределений к пределу и применению метода статистических испытаний (Монте-Карло).

Начнем с обсуждения взаимоотношений асимптотической математической статистики и практики анализа статистических данных. Как обычно подходят к обработке реальных данных в конкретной задаче? Первым делом строят статистическую модель. Если хотят перенести выводы с совокупности результатов наблюдений на более широкую совокупность, например, предсказать что-либо, то рассматривают, как правило, вероятностно-статистическую модель.

Например, традиционную модель выборки, в которой результаты наблюдений — реализации независимых (в совокупности) одинаково распределенных случайных величин. Очевидно, *любая модель лишь приближенно соответствует реальности.* В частности, естественно ожидать, что распределения результатов наблюдений несколько отличаются друг от друга, а сами результаты связаны между собой, хотя и слабо.

Итак, первый этап — переход от реальной ситуации к математической модели. Далее — неожиданность: на настоящем этапе своего развития математическая теория статистики зачастую не позволяет провести необходимые исследования для имеющихся объемов выборок. Более того, отдельные математики пытаются оправдать свой отрыв от практики соображениями о структуре этой теории, на первый взгляд убедительными. Неосторожная давняя фраза Б.В. Гнеденко и А.Н. Колмогорова: «Познавательная ценность теории вероятностей раскрывается только предельными теоремами» (см. классическую монографию [11], одну из наиболее ценных математических книг XX в.) взята на вооружение и более близкими к нам по времени авторами. Так, И.А. Ибрагимов и Р.З. Хасьминский пишут: «Решение неасимптотических задач оценивания, хотя и весьма важное само по себе, как правило, не может являться объектом достаточно общей математической теории. Более того, соответствующее решение часто зависит от конкретного типа распределения, объема выборки и т.д. Так, теория малых выборок из нормального закона будет отличаться от теории малых выборок из закона Пуассона» (см. напичканную формулами монографию [12, с. 7]).

Согласно цитированным авторам и их единомышленникам, основное содержание математической теории статистики — предельные теоремы, полученные в предположении, что объемы рассматриваемых выборок стремятся к бесконечности. Эти теоремы опираются на предельные соотношения теории вероятностей, типа Закона Больших Чисел и Центральной Предельной Теоремы. Ясно, что сами по себе подобные утверждения относятся к математике, т.е. к сфере чистой абстракции, и не могут быть непосредственно применены для анализа реальных данных. Их практическое использование, о котором «чистые» математики предпочитают не думать, опирается на важное предположение: *«При данном объеме выборки достаточно точными являются асимптотические формулы».*

Конечно, в качестве первого приближения представляется естественным воспользоваться асимптотическими формулами, не тратя сил на анализ их точности. Но это — лишь начало долгой цепи исследований. Как же обычно преодолевают разрыв между результатами асимптотической математической статистики и потребностями практики статистического анализа данных? Какие «подводные камни» подстерегают на этом пути?

Точные формулы и асимптотика. Начнем с наиболее продвинутой в математическом плане ситуации, когда для статистики известны как предельное распределение, так и распределения при конечных объемах выборки.

Примером является двухвыборочная односторонняя статистика Н.В. Смирнова. Рассмотрим две независимые выборки объемов m и n из непрерывных функций распределения $F(x)$ и $G(x)$ соответственно. Для проверки гипотезы однородности двух выборок (ср. главу 5)

$$H_0\colon F(x) = G(x) \text{ для всех действительных чисел } x$$

в 1939 г. Н.В. Смирнов в статье [13] предложил использовать статистику

$$D^+(m,n) = sup\,(F_m(x) - G_n(x))\,,$$

где $F_m(x)$ — эмпирическая функция распределения, построенная по первой выборке, $G_n(x)$ — эмпирическая функция распределения, построенная по второй выборке, супремум берется по всем действительным числам x. Для обсуждения проблемы соотношения точных и предельных результатов ограничимся случаем равных объемов выборок, т.е. $m = n$. Положим

$$H(n,t) = P(D^+(n,n) \geq \frac{t}{\sqrt{n}}).$$

В цитированной статье [13] Н.В. Смирнов установил, что при безграничном возрастании объема выборки n вероятность $H(n, t)$ стремится к $exp(-t^2)$.

В работе [14] 1951 г. Б.В. Гнеденко и В.С. Королюк показали, что при целом $c = t\sqrt{n}$ (именно при таких t вероятность $H(n, t)$ как функция t имеет скачки, поскольку статистика Смирнова $D^+(n,n)$ кратна $1/n$) рассматриваемая вероятность $H(n, t)$ выражается через биномиальные коэффициенты, а именно,

$$H(n,t) = \frac{\binom{2n}{n-c}}{\binom{2n}{n}}. \qquad (1)$$

К сожалению, непосредственные расчеты по формуле (1) возможны лишь при сравнительно небольших объемах выборок, поскольку величина n! (n-факториал) уже при n=100 имеет более 200 цифр и не может быть без преобразований использована в вычислениях. Следовательно, наличие точной формулы для интересующей нас вероятности не снимает необходимости использования предельного распределения и изучения точности приближения с его помощью.

Широко известная формула Стирлинга для гамма-функции и, в частности, для факториалов позволяет преобразовать последнее выражение в асимптотическое разложение. Т.е. построить бесконечный степенной ряд (по степеням n) такой, что каждая следующая частичная сумма дает все более точное приближение для интересующей нас вероятности $H(x, t)$. Это и было сделано в работе А.А. Боровкова 1962 г. Большое количество подобных разложений для различных статистических задач приведено в работах В.М. Калинина и О.В. Шалаевского конца 1960-х — начала 1970-х гг. (Интересно отметить, что асимптотические разложения в ряде случаев расходятся, т.е. остаточные члены имеют нетривиальную природу.)

Затем в работах конца семидесятых годов была сделана попытка теоретически оценить остаточный член второго порядка. Итоги подведены в монографии [5, параграф 2.2, с. 37–45]. Справедливо равенство

$$H(n, t) = exp\,(-t^2).(1 + f(t)/n + g(n,t)/n^2),$$

где

$$f(t) = t^2\,(1/2 - t^2/6).$$

Целью последних из названных работ было получение равномерных по n, t оценок остаточного члена второго порядка $g(n,t)$ сверху и снизу в области, задаваемой условиями

$$0 < \frac{t}{\sqrt{n}} < A, \quad 0 < t < t_{\max}, \quad , n \geq n_0. \tag{2}$$

где A, $t_{\max}$, n_0 — некоторые параметры. С помощью длинных цепочек оценок остаточных членов в формулах, получаемых при преобразовании формулы (1) к предельному виду, сформулированная выше цель была достигнута. Для различных наборов параметров A, $t_{\max}$, n_0 получены равномерные по n, t оценки (сверху и снизу) остаточного члена второго порядка $g(n,t)$ в области (2). Так, например, при A = 0,5, t_{max} = 1,73, n_0 = 8 нижняя граница равна (– 0,71), а верхняя есть 2,65.

Основными недостатками такого подхода являются, во первых, зависимость оценок от параметров A, $t_{\max}$, n_0, задающих границы областей, во-вторых, завышение оценок, иногда в сотни раз, обусловленное желанием получить равномерные оценки по области (оценкой реальной погрешности в конкретной точке является значение следующего члена асимптотического разложения).

Поэтому при составлении рассчитанной на практическое использование методики [15] проверки однородности двух выборок с помощью статистики Смирнова было решено перейти на несколько другую методологию (назовем ее «методологией заданной точности»), которую кратко можно описать следующим образом:

1) выбирается достаточно малое положительное число p, например $p = 0{,}05$ или $p = 0{,}20$;

2) приводятся точные значения $H(n, t)$ для всех значений n таких, что

$$|H(n, t) - exp(-t^2)| > p\ exp(-t^2);$$

3) если же последнее неравенство не выполнено, то используется вместо $H(n, t)$ предельное значение $exp(-t^2)$.

Таким образом, принятая в методике [15] методология предполагает интенсивное использование вычислительной техники. Результатами расчетов являются *граничные значения* объемов выборок $n(p,t)$ такие, что при меньших значениях объемов выборок рекомендуется пользоваться точными значениями функции распределения статистики Смирнова, а при больших — предельными. Описывается этот результат таблицей, а не формулой. Отметим, что при построении реальных таблиц не обойтись без выбора того или иного конкретного значения p, задающего объемы таблиц.

Оценки скорости сходимости. Теоретические оценки скорости сходимости в различных задачах прикладной математической статистики иногда формулируются в весьма абстрактном виде. Так, в 1960–1970-х гг. была популярна задача оценки скорости сходимости распределения классической статистики омега-квадрат (Крамера — Мизеса — Смирнова). Для максимума модуля разности допредельной и предельной функций распределения этой статистики различные авторы доказывали, что для любого $e>0$ существует константа $C(e)$

такая, что он не превосходит $C(e)n^{-w+e}$. Прогресс состоял в увеличении константы w. Сформулированный выше результат был доказан последовательно для w = 1/10, 1/6, 1/5, 1/4, 1/3, 1/2 и 1 (подробнее история этих исследований рассказана в §2.3 монографии [5]).

Конечно, все эти исследования не могли дать конкретных практических рекомендаций. Однако необходимой исходной точкой является само существование предельного распределения. Представим себе, что некто, не зная, что у распределения Коши нет математического ожидания, моделирует выборочные средние арифметические результатов наблюдений из этого распределения. Ясно, что его попытки оценить скорость сходимости выборочных средних к пределу обречены на провал.

Последовательное улучшение теоретических оценок скорости сходимости дает надежду на быструю реальную сходимость. Действительно, численные расчеты показали, что предельным распределением для статистики омега-квадрат (Крамера — Мизеса — Смирнова) можно пользоваться уже при объеме выборки, равном 4.

Использование датчиков псевдослучайных чисел. Если же предельное распределение известно, то возникает возможность изучить скорость сходимости численно методом статистических испытаний (Монте-Карло). Однако при этом обычно возникают две проблемы.

Во-первых, откуда известно, что скорость сходимости монотонна? Если при данном объеме выборки различие мало, то будет ли оно мало и при дальнейших объемах? Иногда отклонения допредельного распределения от предельного объясняются довольно сложными причинами. Так, для распределения хи-квадрат они связаны с рядом до сих пор не решенных теоретико-числовых проблем о числе целых точек в эллипсоиде растущего диаметра.

Во-вторых, с помощью датчиков псевдослучайных чисел получаем допредельные распределения с погрешностью, которая может преуменьшать различие. Поясним мысль аналогией. Растущий сигнал измеряется с погрешностями. Когда можно гарантировать, что его величина наверняка превзошла заданную границу?

Напомним, что проблема качества датчиков псевдослучайных чисел продолжает оставаться открытой (см. главу 11 в [9]). Для моделирования в пространствах фиксированной размерности датчики псевдослучайных чисел решают поставленные задачи. Но для рассматриваемых здесь задач размерность не фиксирована — мы не знаем, при каком конкретно объеме выборки можно переходить к предельному распределению согласно «методологии заданной точности».

Нужны дальнейшие работы по изучению качества датчиков псевдослучайных чисел в задачах неопределенной размерности. Поскольку критиков датчиков обычно обвиняют в том, что они сами их не используют, отметим, что мы применяли этот инструментарий при изучении помех, создаваемых электровозами (см. монографию [5]), при изучении статистических критериев проверки однородности двух выборок (см. работу [16]).

А нужна ли вообще асимптотика? В настоящее время развивается актуальное направление статистических методов, связанное с интенсивным использованием вычислительной техники для изучения свойств статистических процедур. Как уже отмечалось, математические методы в статистике обычно позволяют получать лишь асимптотические результаты, и для переноса выводов на конечные объемы выборок приходится применять вычислительные методы. Разработан и успешно применяется компьютерно-статистический подход, основанный на интенсивном использовании современной вычислительной техники. Основная идея такова: в качестве альтернативы асимптотическим методам математической статистики используется анализ результатов статистического моделирования (порядка 2 000–100 000 испытаний) выборок конкретных объемов (10, 50, 100, 200, 500, 1 000). При этом анализ предельных распределений заменяется на анализ распределений соответствующих статистик при указанных объемах выборок.

К достоинствам подхода относится возможность замены теоретических исследований расчетами. Достаточно простая программная система дает (в принципе) возможность численно изучить свойства любого статистического алгоритма для любого конкретного распределения результатов наблюдений и любого конкретного объема выборки. К недостаткам рассматриваемого подхода относится зависимость от свойств датчиков псевдослучайных чисел, а также — что более важно — неизвестность предельного распределения (и даже самого факта его существования), а потому невозможность обоснованного переноса полученных выводов на объемы выборок, отличные от исследованных. Более существенный недостаток — зависимость выводов от того, какое именно распределение моделируется — нормальное, логистическое, Вейбулла — Гнеденко или какое-либо еще. Ведь выводы могут сильно измениться при переходе к другому распределению, а все распределения — их бесконечное количество — перебрать невозможно. Поэтому с точки зрения теории математической статистики полученные рассматриваемым способом результаты следует рассматривать как правдоподобные (а не доказательные, как в классической математической статистике).

Кроме того, они принципиально неточные. Даже в наиболее благоприятных условиях отклонение (в метрике «супремум разности», другое название — метрике Колмогорова) смоделированного распределения, построенного по 2 000 испытаниям, от теоретического предельного распределения может достигать $1{,}358 \times (1/2\ 000)^{1/2} = 0{,}030$. Это означает, в частности, что процентные точки, соответствующие уровням значимости 0,05 и особенно 0,01, могут сильно отличаться от соответствующих процентных точек предельных распределений. Конечно, точность можно увеличить, взяв существенно больше испытаний — не 2 000, а 100 000. Очевидно, следующий этап работ — изучение точности полученных в рассматриваемом подходе выводов, прежде всего приближений и процентных точек, в зависимости от числа испытаний в методах Монте-Карло.

Однако сразу все не сделаешь. Поэтому необходимо развивать новые компьютерные подходы к давним задачам прикладной статистики. Так, весьма интересными и полезными и являются результаты, касающиеся сравнения реальных и номинальных уровней значимости в задачах проверки статистических гипотез [16].

Однако стоит сделать два замечания. В работе [17] сравниваются два плана контроля надежности технических изделий. Оказывается, что при объемах выборки, меньших 150, лучше первый план, а при объемах, больших 150 — второй. Значит, если бы методом статистического моделирования сравнивались эти планы, причем был взят достаточно большой объем выборки n=100, то лучшим был бы признан первый план, что неверно — наступит момент (объем выборки), когда лучшим станет второй план.

Другая относящаяся к делу ассоциация — из весьма содержательной монографии о прикладной математике [18]. Будем суммировать бесконечный ряд с членами $z_n = 1/n$. Поскольку члены его убывают, то обычно используемые алгоритмы остановят вычисления на каком-то шагу. А сумма-то — бесконечна!

Кажется, что компьютер дал универсальную отмычку ко всем проблемам вообще и в области прикладной статистики в частности. Но это только кажется.

Принцип уравнивания погрешностей состоит в том, что погрешности различной природы должны вносит примерно одинаковый вклад в общую погрешность математической модели. Так, определение рационального объема выборки в статистике интервальных данных основано на уравнивании влияния метрологической и статистической погрешностей. Согласно подходу [5] выбор числа градаций в социологических анкетах целесообразно проводить на основе уравнивания погрешностей квантования и неопределенности в ответах респондентов. В классической модели управления запасами целесообразно уравнять влияние неточностей в определении параметров на отклонение целевой функции от оптимума. Из принципа уравнивания погрешностей следует, что относительные погрешности определения различных параметров модели должны совпадать. При использовании классической модели управления запасами погрешность, порожденная отклонением спроса от линейного, оценивается по данным об отпуске товаров. Это дает возможность оценить допустимые отклонения для других параметров. В частности, установить, что расхождения между методиками не являются существенными [5].

В терминах общей схемы устойчивости рассмотрим для простоты записи случай двух параметров. Пусть $A = [0, \infty)\times[0,\infty)$ и $E(x, \alpha) = E(x, (\varepsilon, \delta))$, где $\varepsilon > 0$ и $\delta > 0$ задают точность определения соответствующих параметров, так что $E(x,(\varepsilon_1,\delta_1)) \subseteq E(x,(\varepsilon_2,\delta_2))$ при $\varepsilon_1 \leq \varepsilon_2$, $\delta_1 \leq \delta_2$. Пусть ε задано, а δ исследователь может выбрать, причем известно, что уменьшение δ связано с увеличением расходов. Как выбирать δ? Представляется естественным «уравнять» отклонения, порожденные различными параметрами, т.е. определить δ из условия

$$\beta(x, E(x, (\varepsilon, \delta)) - \beta(x, E(x, (\varepsilon, 0)) \approx \beta(x, E(x, (0, \delta)). \qquad (3)$$

Если затраты и полезный эффект точно известны, то δ можно определить путем решения соответствующей оптимизационной задачи. В противном случае соотношение (3) предлагается использовать в качестве эвристического правила.

Литература

1. Вероятность и математическая статистика : энциклопедия / главный редактор Ю.В. Прохоров. — Москва : Большая Российская энциклопедия, 1999. — 910 с.

2. *Гнеденко, Б.В.* Курс теории вероятностей : учебник / Б.В. Гнеденко. — 7-е изд., испр. — Москва : Эдиториал УРСС, 2001. — 320 с.

3. *Рао, С.Р.* Линейные статистические методы и их применения / С.Р. Рао. — Москва : Наука, 1968. — 548 с.

4. *Келли, Дж.* Общая топология / Дж. Келли. — Москва : Наука, 1968. — 384 с.

5. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

6. *Орлов, А.И.* Асимптотическое поведение статистик интегрального типа / А.И. Орлов // Вероятностные процессы и их приложения : межвузовский сборник. — Москва : МИЭМ, 1989. — С. 118–123.

7. *Вентцель, Е.С.* Теория вероятностей / Е.С. Вентцель. — Москва : Наука, 1964. — 576 с.

8. *Биллингсли, П.* Сходимость вероятностных мер / П. Биллингсли. — Москва : Наука, 1977. — 352 с.

9. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 2-е изд., испр. и доп. — Москва : Экзамен, 2003. — 576 с.

10. *Поляк, Б.Т.* Робастная устойчивость и управление / Б.Т. Поляк, П.С. Щербаков. — Москва : Наука, 2002. — 303 с.

11. *Гнеденко, Б.В.* Предельные распределения для сумм независимых случайных величин / Б.В. Гнеденко, А.Н. Колмогоров. — Москва : Ленинград : ГИТТЛ, 1949. — 264 с.

12. *Ибрагимов, И.А.* Асимптотическая теория оценивания / И.А. Ибрагимов, Р.З. Хасьминский. — Москва : Наука, 1979. — 528 с.

13. *Смирнов, Н.В.* Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках / Н.В. Смирнов // Бюллетень МГУ им. М.В. Ломоносова. Сер. А. — 1939. — Т. 2. — № 2. — С. 3–14.

14. *Гнеденко, Б.В.* О максимальном расхождении двух эмпирических распределений / Б.В. Гнеденко, В.С. Королюк // Доклады АН СССР. — 1951. — Т. 80. — № 4. — С. 525–528.

15. Методика. Проверка однородности двух выборок параметров продукции при оценке ее технического уровня и качества. — Москва : Всесоюзный научно-исследовательский институт стандартизации Госстандарта СССР, 1987. — 116 с.

16. *Камень, Ю.Э.* Реальные и номинальные уровни значимости в задачах проверки статистических гипотез / Ю.Э. Камень, Я.Э. Камень, А.И. Орлов // Заводская лаборатория. — 1986. — Т. 52. — № 12. — С. 55–57.

17. *Левин, Б.Р.* Использование непараметрических методов при обработке результатов испытаний на надежность / Б.Р. Левин, Н.О. Демидович // Надежность средств связи. — Киев : Техніка, 1976. — С. 59–72.

18. *Блехман, И.И.* Механика и прикладная математика: Логика и особенности приложений математики / И.И. Блехман, А.Д. Мышкис, Я.Г. Пановко. — Москва : Наука, 1983. — 328 с.

19. *Орлов, А.И.* Вероятность и прикладная статистика: основные факты : справочник / А.И. Орлов. — Москва : КноРус, 2015. — 190 с.

20. *Орлов, А.И.* Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат и ошибки при их применении / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 97. — С. 32–45.

21. *Орлов, А.И.* Система моделей и методов проверки однородности двух независимых выборок / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 157. — С. 145–169.

22. *Орлов, А.И.* Взаимосвязь предельных теорем и метода Монте-Карло / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 114. — С. 27–41.

23. *Орлов, А.И.* Применение метода Монте-Карло при изучении свойств статистических критериев однородности двух независимых выборок / А.И. Орлов // Научный журнал КубГАУ. — 2019. — № 154. — С. 55–83.

24. *Орлов, А.И.* Устойчивые экономико-математические методы и модели. Разработка и развитие устойчивых экономико-математических методов и моделей для модернизации управления предприятиями / А.И. Орлов. — Saarbrücken : Lambert Academic Publishing, 2011. — 436 с.

25. *Орлов, А.И.* Новый подход к изучению устойчивости выводов в математических моделях / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 100. — С. 146–176.

26. *Орлов, А.И.* Свойства общей схемы устойчивости / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 161. — С. 121–149.

27. *Орлов, А.И.* Теоретические инструменты статистических методов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 101. — С. 253–274.

Контрольные вопросы и задачи

1. Какова роль законов больших чисел в обосновании статистических методов?

2. Приведите примеры использования центральных предельных теорем при построении статистических процедур доверительного оценивания и проверки гипотез.

3. Почему при обосновании статистических методов необходимо использовать теоремы о наследовании сходимости?

4. Примените метод линеаризации для изучения распределения выборочной дисперсии (исходя из асимптотической нормальности при $n \to \infty$ среднего арифметического двумерных векторов $(X_k, (X_k)^2)$, $k = 1, 2, \ldots, n$).

5. Как применяется в прикладной статистике принцип инвариантности?

6. Почему необходимо изучать устойчивость выводов, полученных на основе математических моделей, по отношению к допустимым отклонениям исходных данных и предпосылок моделей?

7. В чем состоит основная идея принципа уравнивания погрешностей?

Темы докладов, рефератов, исследовательских работ

1. Законы больших чисел и различные варианты Центральной предельной теоремы — основные результаты классической теории вероятностей.

2. Место теорем о наследовании сходимости и метода линеаризации в асимптотической прикладной статистике.

3. Необходимые и достаточные условия наследования сходимости.

4. Принцип инвариантности для классических непараметрических статистик.

5. Проблема устойчивости в математическом моделировании.

6. Теория измерений и концепция устойчивости.

7. Принцип уравнивания погрешностей в статистике интервальных данных.

8. Принцип уравнивания погрешностей в логистике (управлении запасами).

ГЛАВА 15. О РАЗВИТИИ СТАТИСТИЧЕСКИХ МЕТОДОВ

15.1. Основные этапы становления статистических методов

Самая ранняя из известных нам статистических работ входит в Библию. В Ветхий Завет включена Четвертая книга Моисеева под названием «Числа». Глава 1 этой книги начинается так (цитируем по синодальному изданию):

«1. И сказал Господь Моисею в пустыне Синайской, в скинии собрания, в первый день второго месяца, во второй год по выходе их из земли Египетской, говоря:

2. Исчислите все общество сынов Израилевых по родам их, по семействам их, по числу имен, всех мужеского пола поголовно,

3. От двадцати лет и выше, всех годных для войны у Израиля, по ополчениям их исчислите их — ты и Аарон.

4. С вами должны быть из каждого колена по одному человеку, который в роде своем есть главный.

***

21. Исчислено в колене Рувимовом сорок шесть тысяч пятьсот.

***

23. Исчислено в колене Симеоновом пятьдесят девять тысяч триста.

***

46. И было всех вошедших в исчисление шестьсот три тысячи пятьсот пятьдесят».

Практическая направленность этого статистического исследования вполне очевидна. Обратите внимание, что оно предпринято по решению руководства страны (в библейских терминах — «общества сынов Израилевых»), причем к работам привлечены региональные начальники (главные по коленам, на которые делилось государство). Четко указана совокупность, подлежащая переписи — мужчины от 20 лет и старше, годные для войны (военнообязанные).

Древность исследования проявляется только в том, что стандартные описания результатов учета военнообязанных по коленам выражены словами. Сейчас мы представили бы результаты в виде таблицы (табл. 1). Таблицы такого типа постоянно составляют органы государственной статистики и в настоящее время (см. портал http://www.gks.ru/wps/portal Федеральной службы государственной статистики РФ (краткое название — Росстат).

Итак, при сравнении с деятельностью Росстата описанное в Библии исследование, выполненное под руководством Моисея, является вполне современным по своим задачам и методам.

Таблица 1

Число всех годных для войны у Израиля

№ п/п	Родоначальник колена	Число военнообязанных
1	Рувим	46 500
2	Симеон	59 300
3	Гад	45 650
4	Иуда	74 600
5	Иссахар	54 400
6	Завулон	57 400
7	Ефрем	40 500
8	Манассия	32 200
9	Вениамин	35 400
10	Дан	62 700
11	Асир	41 500
12	Неффалим	53 400
	Всего:	603 550

Развитие представлений о статистике. В Библии не было терминов «статистика» или «статистик». Впервые термин «статистик» мы находим в художественной литературе — в «Гамлете» Шекспира (1602 г., акт 5, сцена 2). Смысл этого слова у Шекспира — знать, придворные. По-видимому, оно происходит от латинского слова *status*, что в оригинале означает «состояние» или «политическое состояние».

В течении следующих 400 с небольшим лет термин «статистика» понимали и понимают по-разному. В работе [1] сотрудниками Межфакультетской лаборатории статистических методов МГУ им. М.В. Ломоносова собрано более 200 определений этого термина, некоторые из которых приводятся ниже.

Вначале под статистикой понимали описание экономического и политического состояния государства или его части. Например, к 1792 г. относится определение: «Статистика описывает состояние государства в настоящее время или в некоторый известный момент в прошлом». И в настоящее время деятельность государственных статистических служб достаточно хорошо соответствует этому определению.

Однако постепенно термин «статистика» стал использоваться более широко. По Наполеону Бонапарту: «Статистика — это бюджет вещей». Тем самым статистические методы были признаны полезными не только для административного управления, но и для управления на уровне отдельного предприятия. Согласно формулировке 1833 г. «цель статистики заключается в представлении фактов в наиболее сжатой форме». Здесь статистика уже не связывается ни с государствоведением, ни с социально-экономическими проблемами вообще.

Приведем еще два высказывания. «Статистика состоит в наблюдении явлений, которые могут быть подсчитаны или выражены посредством чисел» (1895). «Статистика — это численное представление фактов из любой области исследования в их взаимосвязи» (1909).

В XX в. статистику часто рассматривают прежде всего как самостоятельную научную дисциплина. «Статистика есть совокупность методов и принципов, согласно которым проводится сбор, анализ, сравнение, представление и интерпретация числовых данных» (1925). В 1954 г. академик АН УССР Б.В. Гнеденко дал следующее определение: «Статистика состоит из трех разделов:

1) сбор статистических сведений, т.е. сведений, характеризующих отдельные единицы каких-либо массовых совокупностей;

2) статистическое исследование полученных данных, заключающееся в выяснении тех закономерностей, которые могут быть установлены на основе данных массового наблюдения;

3) разработка приемов статистического наблюдения и анализа статистических данных. Последний раздел, собственно, и составляет содержание математической статистики».

Термин «статистика» употребляют еще в двух смыслах. Во-первых, в обиходе под «статистикой» часто понимают набор количественных данных о каком-либо явлении или процессе. Во-вторых, специалисты в области статистических методов называют «статистикой» функцию от результатов наблюдений, используемую для оценивания характеристик и параметров распределений и проверки гипотез.

Чтобы подойти к современному состоянию, в частности, разъяснить термин «прикладная статистика», кратко рассмотрим историю реальных статистических работ.

Краткая история статистических методов. Как уже отмечалось, типовые примеры раннего этапа применения статистических методов описаны в Ветхом Завете. Там, в частности, приводится число воинов в различных племенах («коленах»). С математической точки зрения дело сводилось к подсчету числа попаданий значений наблюдаемых признаков в определенные градации.

В дальнейшем результаты обработки статистических данных стали представлять в виде таблиц и диаграмм, как это и сейчас делает Росстат. Надо признать, что по сравнению с Ветхим Заветом есть прогресс — в Библии не было таблиц и диаграмм. Однако нет продвижения по сравнению с работами российских статистиков конца девятнадцатого — начала двадцатого века (типовой монографией тех времен можно считать книгу [2], которая в настоящее время еще легко доступна).

Сразу после возникновения теории вероятностей (Паскаль, Ферма, 17 век) вероятностные модели стали использоваться при обработке статистических данных. Например, изучалась частота рождения мальчиков и девочек, было установлено отличие вероятности рождения мальчика от 1/2, анализировались причины того, что в парижских приютах эта вероятность не та, что в самом Париже, и т.д. Имеется достаточно много публикаций по истории теории вероятностей с описанием раннего этапа развития статистических методов исследований, к лучшим из них относится очерк [3]. Отметим, что основатель современного бухгалтерского учета Лука Пачолли (1445–1517) хорошо известен и историкам теории вероятностей. Это символично, поскольку вопросы учета и статистики тесно переплетаются в деятельности современного инженера и менеджера.

В 1794 г. (по другим данным — в 1795 г.) К. Гаусс разработал метод наименьших квадратов, один из наиболее популярных ныне статистических методов, и применил его при расчете орбиты астероида Церера — для борьбы с ошибками астрономических наблюдений [4]. В X1X в. заметный вклад в развитие практической статистики внес бельгиец Кетле, на основе анализа большого числа реальных данных показавший устойчивость относительных статистических показателей, таких, как доля самоубийств среди всех смертей [5]. Интересно, что основные идеи статистического приемочного контроля и сертификации продукции обсуждались академиком Петербургской АН М.В. Остроградским (1801–1862) и применялись в российской армии еще в середине X1X в. [3]. Статистические методы управления качеством и сертификации продукции сейчас весьма актуальны (см. главу 10).

Параметрическая статистика. Современный этап развития статистических методов можно отсчитывать с 1900 г., когда англичанин К. Пирсон основан журнал «*Biometrika*». Первая треть XX в. прошла под знаком параметрической статистики. Изучались методы, основанные на анализе данных из параметрических семейств распределений, описываемых кривыми семейства Пирсона. Наиболее популярным было нормальное (гауссово) распределение. Для проверки гипотез использовались критерии Пирсона, Стьюдента, Фишера, основанные на вероятностно-статистических моделях, в которых результаты измерений (наблюдений, испытаний, опытов, анализов) имели нормальное распределение. В те годы были предложены метод максимального правдоподобия, дисперсионный анализ, сформулированы основные идеи планирования эксперимента.

Разработанную в первой трети XX в. теорию анализа данных называем параметрической статистикой, поскольку ее основной объект изучения — это выборки из распределений, описываемых одним параметром или небольшим числом параметров (2–4). Наиболее общим является семейство кривых Пирсона, задаваемых четырьмя параметрами.

С математической точки зрения параметрическая статистика дает интересные теоретические схемы, на основе которых позволяет построить развитую теорию. Для профессионалов укажем на теорию достаточных статистик, неравенство Рао — Крамера, теорию оптимального оценивания и другие внутриматематические достижения.

Критика параметрической статистики вытекает из ее оторванности от практики статистической работы. Как правило, нельзя указать каких-либо веских причин, по которым распределение результатов конкретных наблюдений непременно должно входить в то или иное параметрическое семейство. Исключения есть, и они хорошо известны: если вероятностная модель предусматривает суммирование независимых случайных величин, то сумму естественно описывать нормальным распределением; если же в модели рассматривается произведение таких величин, то итог, видимо, приближается логарифмически нормальным распределением, и т.д. Однако подобных моделей нет в подавляющем большинстве реальных ситуаций, и приближение реального распределения с помощью кривых из семейства Пирсона или его подсемейств — чисто формальная операция.

Именно из таких соображений критиковал параметрическую статистику академик АН СССР С.Н. Бернштейн в 1927 г. в своем докладе на Всероссийском съезде математиков [6]. Однако эта теория, к сожалению, до сих пор остается основой преподавания статистических методов и продолжает использоваться основной массой прикладников, использующих статистические методы в различных отраслях народного хозяйства и областях науки, но далеких от новых достижений в статистической науке. Почему так происходит? Чтобы попытаться ответить на этот вопрос, обратимся к наукометрии, т.е. к статистическим методам в науковедении.

Наукометрия статистических исследований. В рамках движения за создание Всесоюзной статистической ассоциации (учреждена в 1990 г.) был проведен назад анализ статистики как области научно-практической деятельности. Он показал, в частности, что актуальными для специалистов в настоящее время являются не менее чем 100 тысяч публикаций (подробнее см. статьи [7,8]). Реально же каждый из нас знаком с существенно меньшим количеством книг и статей. Так, в известном трехтомнике М Кендалла и А. Стьюарта [9–11] — наиболее полном на русском языке издании по статистическим методам — всего около 2 тысяч литературных ссылок. При всей очевидности соображений о многократном дублировании в публикациях ценных идей приходится признать, что каждый специалист по статистическим методам владеет лишь небольшой частью накопленных в этой области знаний. Поэтому нет ничего удивительного в том, что приходится постоянно сталкиваться с игнорированием или повторением ранее полученных результатов, с уходом в тупиковые (с точки зрения практики) направления исследований, с беспомощностью при обращении к реальным данным, и т.д. Все это — одно из проявлений адапционного механизма торможения развития науки, вызванного ее быстрым ростом, о котором еще 30 лет назад писали В.В. Налимов и другие науковеды (см., например, [12]).

Традиционный предрассудок состоит в том, что каждый новый результат, полученный исследователем — это кирпич в непрерывно растущее здание науки, который непременно будет проанализирован и использован научным сообществом, а затем и при решении практических задач. Реальная ситуация — совсем иная. Основа профессиональных знаний исследователя, инженера, экономиста, менеджера, социолога, историка, геолога, медика закладывается в период обучения. Затем знания пополняются в том узком направлении, в котором работает специалист. Следующий этап — тиражирование знаний при обучении нового поколения. В результате вузовские учебники отстоят от современного развития на десятки лет. Так, учебники по математической статистике, согласно мнению экспертов, по научному уровню в основном соответствуют 40–60-м гг. XX в. А потому середине XX в. соответствует большинство вновь публикуемых исследований и тем более — прикладных работ. Одновременно приходится признать, что результаты, не вошедшие в учебники, независимо от их ценности почти все забываются. Достаточно взглянуть на длинные ряды библиотечных полок с номерами научных журналов за последние сто лет. Сколько из них были хотя бы раз открыты в текущем веке? Кроме того, сейчас все популярнее поиск информации в Интернете — вплоть до того, что кое-кто из молодых даже забывает о существовании библиотек. А ведь в Интернете можно найти лишь небольшую часть опубликованных в XX в. научных работ.

Активно продолжается развитие тупиковых направлений. Психологически это понятно. Приведу пример из своего опыта. В свое время по заказу Госстандарта я разработал методы оценки параметров гамма-распределения [13]. Поэтому мне близки и интересны работы по оцениванию параметров по выборкам из распределений, принадлежащих тем или иным параметрическим семействам, понятия функции максимального правдоподобия, эффективности оценок, использование неравенства Рао — Крамера и т.д. К сожалению, я знаю, что это — тупиковая ветвь теории статистики, поскольку реальные данные не подчиняются каким-либо параметрическим семействам, надо применять иные статистические методы, о которых речь пойдет ниже. Понятно, что специалистам по параметрической статистике, потратившим многие годы на совершенствование в своей области, психологически трудно согласиться с этим утверждением. В том числе и мне. Но необходимо идти вперед. Поэтому настоящий учебник во многом очищен от тупиковых подходов. В том числе и от неравенства Рао-Крамера. Однако я включил разделы 3.1 и 3.2, посвященные оцениванию параметров распределений, поскольку эта тематика часто обсуждается в литературе, причем с устаревших позиций. Например, вместо уходящих в прошлое оценок максимального правдоподобия в настоящее время рекомендуют использовать одношаговые оценки (раздел 3.2).

Непараметрическая статистика. Статистические методы, которые не основаны на нереалистическом предположении о том, что рассматриваемые выборки взяты из распределений, описываемых одним параметром или небольшим числом параметров (2–4), называют *непараметрическими*. При математическом обосновании непараметрических статистических методов обычно вводят те или иные условия регулярности, например, требуют непрерывности функции распределения результатов наблюдений или существования математического ожидания и дисперсии. Как правило, подобные условия регулярности носят внутриматематический характер и не ограничивают прикладные возможности непараметрических методов.

Примерами являются критерии Колмогорова, Смирнова, Реньи, Вилкоксона, омега-квадрат (Крамера — Мизеса — Смирнова) [14], предназначенные для проверки гипотез согласия и однородности и разработанные в 30–40-х гг. XX в. История непараметрических коэффициентов корреляции Спирмена и Кендалла уходит корнями в работы начала XX в. В 50-х гг. с известной работы Н.В. Смирнова [15] началась разработка методов непараметрического оценивания плотности. Непараметрическая статистика активно развивается и в XXI в. (см. главу 16).

Во второй половине XX в. появились новые области статистических методов — робастная статистика, компьютерное статистическое моделирование (методы Монте-Карло, бутстреп-методы), статистика нечисловых и интервальных данных. Эти области активно развиваются и в настоящее время. О них пойдет речь в главе 16.

Иные причины привели к появлению и распространению прикладной статистики. Что означает этот термин? Вполне естественно, что математическая статистика выступает как метатеория по отношению к статистическим методам в той или иной области применения — к эконометрике, т.е. статистическим методам в экономике [16], к наукометрии [12], к биометрике и другим «метрикам». По цитированному выше определению Б.В. Гнеденко: «разработка приемов статистического наблюдения и анализа статистических данных составляет содержание математической статистики». Почему понадобилась новая научная область — прикладная статистика — между математической статистикой и статистическими методами в конкретных областях применений? Для ответа на этот вопрос необходимо обсудить внутреннюю логику развития статистических методов как научно-прикладной дисциплины.

Появление прикладной статистики. В нашей стране термин «прикладная статистика» вошел в широкое употребление в 1981 г. после выхода массовым тиражом (33940 экз.) сборника «Современные проблемы кибернетики (прикладная статистика)». В этом сборнике обосновывалась трехкомпонентная структура прикладной статистики [17]. Во-первых, в нее входят ориентированные на прикладную деятельность статистические методы анализа данных (эту область можно назвать прикладной математической статистикой и включать также и в прикладную математику). Однако прикладную статистику нельзя целиком относить к математике. Она включает в себя две внематематические области. Во-первых, методологию организации статистического исследования: как планировать исследование, как собирать данные, как подготавливать данные к обработке, какие вероятностно-статистические модели использовать, какие статистические методы выбирать для обработки данных, как представлять результаты. Во-вторых, организацию компьютерной обработки данных, в том числе разработку и использование баз данных и электронных таблиц, статистических программных продуктов, например, диалоговых систем анализа данных. В нашей стране термин «прикладная статистика» использовался и ранее 1981 г., но лишь внутри сравнительно небольших и замкнутых групп специалистов [17].

Прикладная статистика и математическая статистика — это две разные научные дисциплины. Различие четко проявляется не только в исследованиях, но и при преподавании. Курс математической статистики состоит в основном из доказательств теорем, как и соответствующие учебники и учебные пособия.

В курсах прикладной статистики основное — методология анализа данных и алгоритмы расчетов, а теоремы приводятся как обоснования этих алгоритмов, доказательства же, как правило, опускаются (их можно найти в научной литературе).

К настоящему времени беспристрастному наблюдателю очевидно четко выраженное размежевание этих двух научных дисциплин. Математическая статистика исходит из сформулированных в 1930–1950 гг. постановок математических задач, происхождение которых связано с рассматриваемыми в те времена проблемами анализа статистических данных. Начиная с 70-х гг. XX в. исследования по математической статистике посвящены обобщению и дальнейшему математическому изучению этих старых задач. Поток новых математических результатов (теорем) не ослабевает, но новые практические рекомендации по обработке статистических данных при этом не появляются. Можно сказать, что математическая статистика как научное направление замкнулась внутри себя.

Сам термин «прикладная статистика» возник как реакция на описанную выше тенденцию. Прикладная статистика нацелена на решение реальных задач. Поэтому в ней возникают новые постановки математических задач анализа статистических данных, развиваются и обосновываются новые методы. Обоснование часто проводится математическими методами, т.е. путем доказательства теорем. Большую роль играет методологическая составляющая — как именно ставить задачи, какие предположения принять с целью дальнейшего математического изучения. Велика роль современных информационных технологий, в частности, компьютерного эксперимента.

Рассматриваемое соотношение математической и прикладной статистик отнюдь не являются исключением. Как правило, математические дисциплины проходят в своем развитии ряд этапов. Вначале в какой-либо прикладной области возникает необходимость в применении математических методов и накапливаются соответствующие эмпирические приемы (для геометрии это — «измерение земли», т.е. землемерие, в Древнем Египте). Затем возникает математическая дисциплина со своей аксиоматикой (для геометрии это — время Евклида). Затем идет внутриматематическое развитие и преподавание (известно, что большинство результатов элементарной геометрии получено учителями гимназий в XIX в.). При этом на запросы исходной прикладной области перестают обращать внимание, и та для решения своих задач порождает новые научные дисциплины (сейчас «измерением земли» занимается не геометрия, а геодезия и картография). Затем научный интерес к исходной дисциплине иссякает, но преподавание по традиции продолжается (элементарная геометрия «ушла» из вузов, но до сих пор изучается в средней школе, хотя трудно понять, в каких практических задачах может понадобиться, например, теорема о том, что высоты треугольника пересекаются в одной точке). Следующий этап — окончательное вытеснение дисциплины из реальной жизни в историю науки (объем преподавания элементарной геометрии в настоящее время постепенно сокращается, в частности, ей все меньше уделяется внимания на вступительных экзаменах в вузах). К интеллектуальным дисциплинам, закончившим свой жизненный путь, относится средневековая схоластика. Как справедливо отмечает профессор МГУ им. М.В. Ломоносова В.Н. Тутубалин [18], теория вероятностей и математическая статистика успешно двигаются по ее пути — вслед за элементарной геометрией.

Резюмируем сказанное. Хотя статистические данные собираются и анализируются с незапамятных времен (см., например, Книгу Чисел в Ветхом Завете), современная математическая статистика как наука была создана, по общему мнению специалистов, сравнительно недавно — в первой половине XX в. Именно тогда были разработаны основные идеи и получены результаты, излагаемые ныне в учебных курсах математической статистики. После чего специалисты по математической статистике занялись внутриматематическими проблемами, а для теоретического обслуживания проблем практического анализа статистических данных стала формироваться новая дисциплина — прикладная статистика.

В настоящее время статистическая обработка данных проводится, как правило, с помощью соответствующих программных продуктов. Однако между математической и прикладной статистикой существует разрыв — большинство методов, включенных в статистические пакеты программ (*Statgraphics*, *SPSS*, *Statistica*), не отражены в учебниках по математической статистике. В итоге специалист по математической статистике зачастую не может обработать реальные данные, а пакеты программ применяют и разрабатывают лица, не имеющие необходимой теоретической подготовки. Естественно, что они допускают разнообразные ошибки, в том числе в таких ответственных документах, как государственные стандарты по статистическим методам (о грубых ошибках в ГОСТах рассказано в статье [19] и главе 10).

Что дает прикладная статистика народному хозяйству? Так называлась статья [20], в которой приводились многочисленные примеры успешного использования прикладной статистики и других статистических методов при решении практических задач. Перечень примеров можно продолжать практически безгранично (см., например, сводку [21]).

Методы прикладной статистики используются в зарубежных и отечественных экономических и технических исследованиях, работах по управлению (менеджменту), в медицине, социологии, психологии, истории, геологии и других областях. Их применение дает заметный экономический эффект. Например, в США — не менее 20 миллиардов долларов ежегодно только в области статистического контроля качества. В 1988 г. затраты на статистический анализ данных в нашей стране оценивались в 2 миллиарда рублей ежегодно [22]. Согласно расчетам сравнительной стоимости валют на основе потребительских паритетов [16], эту величину можно сопоставить с 6 миллиардами долларов США. Следовательно, объем отечественного «рынка статистических услуг» был на порядок меньше, чем в США, что совпадает с оценками и по другим показателям, например, по числу специалистов.

Своеобразие исторического пути России привело к тому, что в нашей стране нет специализированного научного журнала по статистическим методам. Публикации по новым статистическим методам, по их применениям в технико-экономических исследованиях, в инженерном деле постоянно появляются, прежде всего, в журнале «Заводская лаборатория», в секции «Математические методы исследования». Надо назвать также журналы «Автоматика и телемеханика» (издается Институтом проблем управления Российской академии наук), «Экономика и математические методы» (издается Центральным экономико-математическим институтом РАН).

Однако необходимо констатировать, что для большинства менеджеров, экономистов и инженеров прикладная статистика и другие статистические методы являются пока экзотикой. Это объясняется тем, что в вузах современным статистическим методам почти не учат. Во всяком случае, по состоянию на 2021 г. каждый квалифицированный специалист в этой области — самоучка.

Этому выводу не мешает то, что в вузовских программах обычно есть два курса, связанных со статистическими методами. Один из них — «Теория вероятностей и математическая статистика». Этот небольшой курс обычно читают специалисты с математических кафедр. Они успевают дать лишь общее представление об основных понятиях математической статистики. Кроме того, внимание математиков обычно сосредоточено на внутриматематических проблемах, их больше интересует доказательства теорем, а не применение современных статистических методов в задачах экономики и менеджмента. Другой курс — «Статистика» или «Общая теория статистики», входящий в стандартный блок экономических дисциплин. Фактически он является введением в прикладную статистику и содержит первые начала эконометрических методов (по состоянию на 1900 г.).

Статистические методы как учебный предмет опираются на два названных вводных курса. Она призвана вооружить специалиста современным статистическим инструментарием. Специалист — это инженер, экономист, менеджер, геолог, медик, социолог, психолог, историк, химик, физик и т.д. Во многих странах мира — Японии и США, Франции и Швейцарии, Перу и Ботсване и др. — статистическим методам обучают в средней школе. ЮНЕСКО постоянно проводят конференции по вопросам такого обучения [23]. В СССР и СЭВ, а теперь — по плохой традиции — и в России игнорируют этот предмет в средней школе и лишь слегка затрагивают его в высшей. Результат на рынке труда очевиден — снижение конкурентоспособности специалистов.

Проблемы прикладной статистики и других статистических методов постоянно обсуждаются специалистами. Широкий интерес вызвала дискуссия в журнале «Вестник статистики», в рамках которой были, в частности, опубликованы статьи [8, 20]. На появление в нашей стране прикладной статистики отреагировали и в США [24].

В нашей стране получены многие фундаментальные результаты прикладной статистики. Огромное значение имеют работы академика РАН А.Н. Колмогорова [25]. Во многих случаях именно его работы дали первоначальный толчок дальнейшему развитию ряда направлений прикладной статистики. Зачастую еще 50–70 лет назад А.Н. Колмогоров рассматривал те проблемы, которые только сейчас начинают широко обсуждаться. Как правило, его работы не устарели и сейчас. Свою жизнь посвятили прикладной статистике члены-корреспонденты АН СССР Н.В. Смирнов и Л.Н. Большев. В настоящем учебнике постоянно встречаются ссылки на лучшую публикацию XX в. по статистическим методам — составленные ими подробно откомментированные «Таблицы …» [14].

Об отечественных исследованиях в области прикладной статистики и других статистических методов подробнее поговорим в следующем разделе.

15.2. Статистические методы в России

Специалисты по истории статистики установили [5], что в России, как и в других странах, статистические исследования проводились с момента возникновения государств. Цели этих исследований, как и описанных в Библии работ под руководством Моисея, вытекали из потребностей государственного управления, прежде всего налогообложения и обороны страны. С XII в. (в традиционной хронологии) на Руси проводились переписи населения [5]. Развитие статистической науки началось в России сразу же с выделением в начале XVIII века исследовательской деятельности как необходимой составляющей забот государства. Проще говоря, сразу же с организацией первого научного учреждения — Академии наук.

Первое статистико-экономическое обозрение России было составлено Иваном Кирилловичем Кириловым (1689–1737), обер-секретарем Сената, под названием «Цветущее состояние Всероссийского государства…». Первый в России научный труд по вопросам организации учета населения — «Разсуждение о ревизии поголовной и касаюсчемся до оной» — был написан в 1747 г. Василием Никитичем Татищевым (1686–1750), известным государственным деятелем той эпохи. Он, в частности, одним из первых применял анкеты для сбора статистических данных. Большой вклад в теорию и практику отечественной статистики внес Михаил Васильевич Ломоносов (1711–1765).

Подробное описание развития статистической науки и практики в России можно найти в трудах по истории социально-экономической ветви статистики (см., например, [5, 26]). К сожалению, в этих работах обычно не рассматривается развитие отечественной вероятностно-статистической научной школы (см., например, [3]). О причинах такой однобокости скажем ниже.

Реформы императора Александра Второго, прежде всего создание земств (органов местного самоуправления), дали мощный стимул развитию статистики. Связано это было прежде всего с тем, что штатное расписание губернских и уездных земств, как правило, включало должность статистика. Так, к концу 1894 г. за 15 лет активной статистической деятельности были собраны, разработаны и опубликованы земствами материалы крестьянских подворных переписей по 172 уездам, охватившим около 4 миллионов крестьянских дворов — примерно четвертую часть всего населения России [5, с. 109].

Проведение статистических исследований было делом чести для отечественной интеллигенции. Так, Антон Павлович Чехов по собственной инициативе в 1890 г. перепись на Сахалине, лично опросив несколько тысяч каторжников [27].

Расцвет статистики в конце XIX в. проявился в появлении большого числа оригинальных исследований, выполненных на высоком профессиональном уровне. Одна из них хорошо известна и в настоящее время, что объясняется личностью автора. Речь идет о книге В.И. Ульянова (Ленина) «Развитие капитализма в России. Процесс образования внутреннего рынка для крупной промышленности» [2]. Она была издана в 1899 г., когда автору было 29 лет. По современным критериям за эту монографию автору можно было бы присудить ученую степень доктора экономических наук. Это утверждение свидетельствует не только о высоком профессиональном уровне В.И. Ульянова как исследователя, но и об известной деградации социально-экономической статистики за последние сто лет.

Наибольшие достижения в XX в. были получены в России в математической статистике. Упомянем работы А.А. Чупрова (1874–1926) по теории корреляции. Несколько позже началась деятельность А.Н. Колмогорова.

Среди математиков XX столетия академик АН СССР А.Н. Колмогоров (1903–1987) должен быть назван первым. Именно его работы дали первоначальный толчок дальнейшему развитию ряда направлений, важных для современных статистических методов. Еще 50–70 лет назад А.Н. Колмогоров рассматривал те проблемы, которые только сейчас начинают широко обсуждаться.

Вероятностно-статистические методы исследования в работах А.Н. Колмогорова. С современной точки зрения [25] обсудим работы А.Н. Колмогорова по аксиоматическому подходу к теории вероятностей, критерию согласия эмпирического распределения с теоретическим, свойствам медианы как оценки центра распределения, эффекту «вздувания» коэффициента корреляции, теории средних величин, статистической теории кристаллизации металлов, методу наименьших квадратов, свойствам сумм случайного числа случайных слагаемых, статистическому контролю, несмещенным оценкам, аксиоматическому получению логарифмически нормального закона распределения при дроблении, методам обнаружения различий при экспериментах типа погодных.

Факты жизни и творчества А.Н. Колмогорова подробно рассмотрены в сборнике [28]. Его основные работы изданы в трех томах [29–31]. Андрей Николаевич считал, что хорошая математическая работа должна содержать простую идею (желательно геометрического характера), использовать «тонкую» аналитику, а хорошая и полезная прикладная работа должна опираться на фундаментальные теоретические основы.

Аксиоматический подход к теории вероятностей [32] позволил рассматривать теорию вероятностей и математическую статистику как часть математики, проводить рассуждения на математическом уровне строгости. В частности, было введено четкое различие между частотой и вероятностью, случайная величина стала рассматриваться как функция от элементарного исхода, и т.д. За основу методов статистического анализа данных стало возможным брать вероятностно-статистические модели, сформулированные в математических терминах. В результате удалось четко отделить строгие утверждения от обсуждения философских вопросов случайности, преодолеть подход на основе понятия равновозможности, имеющий ограниченное практическое значение. Наиболее существенно, что после работ А.Н. Колмогорова нет необходимости связывать вероятности тех или иных событий с пределами частот. Так называемые «субъективные вероятности» получили смысл экспертных оценок вероятностей.

После выхода (в 1933 г. на немецком языке и в 1936 г. — на русском) основополагающей монографии [32] аксиоматический подход к теории вероятностей стал общепринятым в научных исследованиях в этой области. Во многом перестроилось преподавание. Повысился научный уровень многих прикладных работ. Однако традиционный подход оказался живучим. С целью повышения строгости формулировок приходится помещать в наших учебниках ([16, 33] и др.) сводки терминов и определений в области вероятностно-статистических методов, опирающаяся на аксиоматику [32].

В послевоенные годы А.Н. Колмогоров формализовал понятие случайности на основе теории информации [31]. Грубо говоря, числовая последовательность является случайной, если ее нельзя заметно сжать (т.е. описать существенно короче) без потери информации. Однако этот подход не был предназначен для использования в прикладных работах и преподавании. Он представляет собой важное методологическое и теоретическое продвижение.

Критерии согласия. В работе 1933 г. «Об эмпирическом определении закона распределения» [30, с. 134–141] А.Н. Колмогоров предложил и изучил «критерий Колмогорова». Пусть элементы выборки (независимые случайные величины) объема n имеют непрерывную функцию распределения $F(x)$. Эмпирической функцией распределения $F_n(x)$ называется доля элементов выборки, не превосходящих x. Критерий Колмогорова предназначен для проверки гипотезы

$$H_0 : F(x) \equiv F_0(x),$$

где $F_0(x)$ — заданная функция распределения. Его статистика имеет вид

$$D_n = \sqrt{n} \sup_x | F_n(x) - F_0(x) | .$$

В [30, с. 134–141] показано, что функция распределения статистики D_n имеет предел,

$$\lim_{n\to\infty} P\{D_n \le \lambda\} = K(\lambda),$$

и рассчитана первая таблица функции распределения Колмогорова $K(\lambda)$.

Работа [30, с. 134–141] породила одно из основных направлений непараметрической статистики. И в настоящее время непараметрические критерии согласия (Колмогорова, Смирнова, омега-квадрат и др.) широко используются. Они были разработаны для проверки согласия с *полностью известным* теоретическим распределением. Основная идея критериев Колмогорова, омега-квадрат и аналогичных им состоит в измерении расстояния между функцией эмпирического распределения и функцией теоретического распределения. Различаются эти критерии видом расстояний в пространстве функций распределения. Расчетные формулы, таблицы распределений и критических значений широко распространены (см. [14] и главу 5 выше).

Часто возникает задача проверки гипотезы согласия эмпирического распределения с параметрическим семейством, например, с семейством нормальных, Вейбулла — Гнеденко или гамма-распределений. Представляется естественным оценить параметры распределения по выборке, а затем формально воспользоваться критериями согласия Колмогорова, Смирнова или омега-квадрат. При этом вместо фиксированной теоретической функции распределения подставляют функцию из параметрического семейства, в которой параметры заменены на их выборочные оценки. В отличие от классических критериев, при этом измеряются расстояния от эмпирической функции распределения до многообразий (в пространстве функций распределения), соответствующих параметрическим семействам. Развита [8] математическая техника проектирования в функциональных пространствах, которая позволяет строить методы проверки рассматриваемых гипотез.

Однако распределения таких критериев (как предельные, так и при конечных объемах выборок) *существенно отличаются* от распределений классических критериев согласия Колмогорова, Смирнова или омега-квадрат. Такие критерии в отличие от классических обычно называют «критериями согласия с параметрическим семейством типа Колмогорова — Смирнова и типа омега-квадрат». (Как показано в [35] на основе анализа исходных публикаций, корректно употреблять термины «критерий Колмогорова», «критерий Смирнова», «критерий типа Колмогорова — Смирнова», но нельзя говорить о несуществующем «критерии Колмогорова — Смирнова».) В [36] собраны основные факты о критериях согласия с параметрическими семействами типа Колмогорова-Смирнова и типа омега-квадрат и необходимые краткие таблицы. Современное положение дел в этой области отражено в [25]. Наиболее существенное продвижение в изучении критериев типа Колмогорова — Смирнова достигнуто Ю.Н. Тюриным [37].

«Вздувание» коэффициента корреляции — явление, обнаруженное А.Н. Колмогоровым в работе 1933 г. «К вопросу о пригодности найденных статистическим путем формул прогноза» [30, с. 161–167]. Предположим, что имеется много наборов предикторов (факторов, признаков). Для каждого из них строится наилучшее приближение отклика с помощью линейной функции от предикторов. Показателем качества приближения служит коэффициент корреляции между откликом и наилучшей линейной функцией от предикторов (в настоящее время чаще используют его квадрат, называемый коэффициентом детерминации). Эффект «вздувания» коэффициента корреляции состоит в том, что при увеличении числа проанализированных наборов предикторов заметно растет максимальный из соответствующих коэффициентов корреляции — показателей качества приближения. Создается впечатление, что тот набор предикторов, на котором достигается рассматриваемый максимум, дает хорошее приближение для отклика. Однако это приближение развеивается при попытке использовать соответствующую зависимость для прогноза — по новым данным коэффициент корреляции между откликом и ранее найденной линейной функцией от предикторов оказывается значительно меньшим.

В настоящее время весьма популярны методы поиска «наиболее информативного множества признаков» в регрессионном и дискриминантном анализе. Соответствующие алгоритмы, как правило, основаны на переборе большого числа наборов признаков. Поэтому, как показано в [38], актуальность работы А.Н. Колмогорова [30, с. 161–167] в настоящее время существенно повысилась. Эффект «вздувания» коэффициента корреляции является одним из проявлений неклассического поведения статистических характеристик в ситуации, когда одна и та же статистическая процедура осуществляется многократно, например, при множественных проверках статистических гипотез (см. раздел 4.5).

В течение полувека А.Н. Колмогоров интересовался статистическими постановками, в которых число неизвестных параметров растет вместе с объемом данных. К ним относится и работа [30, с. 161–167]. А в 1970-х годах он стимулировал исследования по т.н. «асимптотике Колмогорова»

$$p \to \infty,\ \ n \to \infty,\ \ p/n \to \lambda > 0,$$

где p — число параметров, n — объем выборки. Эта асимптотика весьма актуальна как для многомерного статистического анализа, так и для статистики нечисловых данных [39], а также для задач статистического приемочного контроля [16, параграф 13.5] и анализа социологических данных (см. главу 13).

Метод медианы в теории оценивания. Пусть $X_1, X_2, \ldots, X_n$ — независимые одинаково распределенные случайные величины с функцией распределения F и непрерывной плотностью f. Пусть μ и σ^2 — соответственно математическое ожидание и дисперсия, а m — медиана распределения F (т.е. $P\{X_1 \geq m\} \geq 1/2$ и $P\{X_1 \leq m\} \geq 1/2$). Медиана всегда существует, но не всегда определяется однозначно. Обычно в качестве оценки для μ используют (в случае нормального закона, прежде всего) выборочное среднее арифметическое

$$\overline{X}_n = n^{-1} \sum_{k=1}^{n} X_k \,,$$

обладающее при условии нормальности F оптимальными свойствами. Что делать, если распределение F отлично от нормального? В работе 1931 г. «Метод медианы в теории ошибок» [30, с. 111–114] А.Н. Колмогоров предлагает в этом случае оценивать по выборке другую среднюю характеристику распределения — медиану m (для симметричных распределений эти две характеристики совпадают). Пусть $X_n(k)$ — k-я порядковая статистика, построенная по рассматриваемой выборке. Если n четно, то в качестве оценки m_n медианы m возьмем $X_n(n/2)$; если же $n = 2k+1$, то в качестве оценки m возьмем $X_n(k)$. С целью сравнения оценок $\overline{X}_n$ и m_n рассмотрим преобразованные величины

$$\alpha_n = \sqrt{n}\left(\overline{X}_n - \mu\right), \ \beta_n = \sqrt{n}\left(m_n - m\right).$$

Согласно центральной предельной теореме предельное (при $n \to \infty$) распределение величины α_n является асимптотически нормальным с нулевым средним и дисперсией σ^2. Можно показать [30, с. 111–114], что распределение величины β_n является асимптотически нормальным с нулевым средним и дисперсией $\sigma_m = (1/2)/f(m)$, если $f(m)$ отлично от 0. Мерой сравнительной точности обоих методов является отношение $\lambda = \sigma_m/\sigma = (1/2)/[\sigma f(m)]$. В случае нормальной плотности f имеем $\lambda = (\pi/2)^{1/2} \approx 5/4$. Как показал А.Н. Колмогоров [30, с. 111–114], для унимодальных распределений отношение λ может принимать любое значение из интервала $(0; \sqrt{3})$, но не может превосходить $\sqrt{3}$.

Средние по Колмогорову. Естественная система аксиом приводит к так называемым ассоциативным средним. Их общий вид нашел в 1930 г. А.Н. Колмогоров [29, с. 136–138]. Теперь их называют «средними по Колмогорову». Для чисел $X_1, X_2, ..., X_n$ среднее по Колмогорову вычисляется как

$$G\{(F(X_1)+F(X_2)+...+F(X_n))/n\},$$

где F — строго монотонная функция (т.е. строго возрастающая или строго убывающая), G — функция, обратная к F. Среди средних по Колмогорову — много хорошо известных средних величин. Так, если $F(x) = x$, то среднее по Колмогорову — это среднее арифметическое. Для положительных $X_1, X_2, ..., X_n$: если $F(x) = \ln x$, то среднее по Колмогорову — это среднее геометрическое, если $F(x) = 1/x$, то среднее гармоническое, если $F(x) = x^2$, то среднее квадратическое, и т.д. Однако такие популярные средние, как медиана и мода, нельзя представить в виде средних по Колмогорову. В настоящем учебнике к средним по Колмогорову обращались в связи с рассмотрением выбора алгоритмов для анализа данных, измеренных в той или иной шкале (см. раздел 10.3). Так, для алгоритмов усреднения установлено, что в шкале интервалов из всех средних по Колмогорову допустимым является только среднее арифметическое, а в шкале отношений — только степенные средние с $F(x) = x^c$, (при c, отличном от 0) и среднее геометрическое. Таким образом, среднее геометрическое или среднее квадратическое температур (в шкале Цельсия) или расстояний не имеют смысла. В качестве среднего в шкале интервалов надо применять среднее арифметическое. А также можно использовать медиану или моду.

Статистической теории кристаллизации металлов посвящена работа 1937 г. [30, с. 178–182]. Построена модель возникновения центров кристаллизации и нарастания закристаллизованной массы. При широких допущениях найдена точная формула для вероятности $p(t)$, с которой наудачу выбранная точка P объема, заполненного подлежащим кристаллизации веществом, попадет в течение промежутка кристаллизации t внутрь уже закристаллизованной массы. С достаточным приближением можно считать, что доля вещества, закристаллизовавшегося за время t, также равно $p(t)$. Рассчитано число центров кристаллизации, образующихся в течение всего процесса кристаллизации. Полученные в работе [30, с. 178–182] результаты до сих пор представляют интерес для всех специалистов, связанных с изучением и использованием процессов кристаллизации металлов и иных веществ.

Метод наименьших квадратов. В двух работах А.Н. Колмогорова [30, с. 267–283, с. 283–288] 1946–1947 гг. построена геометрическая теория метода наименьших квадратов, выявляющая роль ортогонального проектирования на подпространства конечномерного евклидова пространства с целью получения оценок параметров. Эта идея затем широко использовалась как в научных исследованиях, так и при преподавании.

Вторая идея состоит в построении алгоритмов доверительного оценивания и проверки гипотез на основе предположения о нормальности распределения погрешностей измерения. К настоящему времени вторая идея изжила себя, поскольку установлено, что в подавляющем большинстве случаев распределение погрешностей заметно отличается от нормального (см. раздел 2.1). Поэтому современный подход (глава 6) к методу наименьших квадратов является непараметрическим, т.е. в определенном смысле наблюдается возврат к доколмогоровским взглядам.

Суммы случайного числа случайных слагаемых рассмотрены в работе 1949 г. [30, с. 308–313], выполненной совместно с Ю.В. Прохоровым, в дальнейшем академиком АН СССР. Эта статья стимулировала исследования по важному для приложений виду предельных теорем (см. [40, с. 300–312; 41 с. 223–228]). Речь идет прежде всего о статистическом последовательном анализе [42], в частности, об изучении времени наблюдения в задаче последовательного различения двух простых гипотез. Предельные теоремы [43, 44] о суммах случайного числа случайных слагаемых находят применения в задачах статистического контроля качества и надежности по Вальду, в моделях управления запасами в логистике (см. раздел 8.3 и монографию [45]) и др.

Статистический контроль. А.Н. Колмогоров — основоположник современной теории статистического приемочного контроля в нашей стране. Более 150 лет статистические методы применяются в России для проверки соответствия продукции установленным требованиям, т.е. для сертификации. Так, еще в 1846 г. действительный член Петербургской академии наук М.В. Остроградский рассматривал задачу статистического контроля партий мешков муки или штук сукна армейскими поставщиками. Однако современный этап начался в 1951 г. с брошюры А.Н. Колмогорова [46]. С тех пор в России в статистическом контроле качества было сделано многое, особенно в области теории [47–49]. Вопросы статистического контроля постоянно рассматриваются на страницах журнала «Заводская лаборатория» — основного места публикации отечественных работ по статистическим методам [7, 19].

Большое значение для развития статистических методов управления качеством имеют статья А.Н. Колмогорова 1933 г. [30, с. 134–141] о критерии согласия эмпирического распределения с теоретическим и статья 1950 г. о несмещенных оценках [30, с. 340–363]. Актуальность первой из них определяется недостатками в используемых до сих пор статистических методах управления качеством. Широко распространенные ошибки состоят в том, что для критериев согласия с параметрическими семействами используют критические значения классических критериев. При этом, например, гипотеза нормальности принимается гораздо чаще, чем следует. Поскольку в действующей нормативно-технической документации дальнейшие этапы анализа данных часто зависят от того, принимается нормальность или нет, то ошибки при такой проверке могут иметь далеко идущие последствия. Так, при анализе характеристик эластомерных материалов при ошибочном подходе из 30 выборок нормальность была отвергнута лишь для 2, а при правильном — для 26, т.е. в подавляющем большинстве случаев. Указанные ошибки встречаются в массе публикаций (хотя специалистам суть дела хорошо известна уже почти 50 лет [50]). Наиболее известным примером является полностью ошибочный ГОСТ 11.006-74 (СТ СЭВ 1190-78) «Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим» (в настоящее время отменен как нормативный документ, но, к сожалению, продолжает использоваться как научная и методическая публикация; ошибки вскрыты, и замена алгоритмов расчета обоснована в [106]).

Популярны и другие ошибки при применении рассматриваемых критериев согласия. Некоторые пытаются их использовать для сгруппированных данных, что приводит к излишне частому принятию гипотез [40]. Другие вместо эмпирической функции распределения рассматривают иные оценки теоретической функции распределения. Например, при использовании вероятностной бумаги удобно ординату точки, соответствующей i-ой порядковой статистике, установить равной $(i–0.5)/n$, а не i/n, как в классической эмпирической функции распределения. Возникает искушение построенную таким методом оценку использовать в критериях согласия вместо эмпирической функции распределения. Увы, распределение изменится (впрочем, в данном случае при росте объема выборки различие будет исчезать). Ряд ошибок рассмотрен в [14].

Несмещенные оценки. При оценивании по выборке параметров распределений (либо функций от них) рекомендуют использовать метод максимального правдоподобия, дающий при выполнении условий регулярности асимптотически оптимальные оценки. Однако часто возникают трудности с решением уравнений правдоподобия. Поэтому вместо оценок максимального правдоподобия применяют асимптотически им эквивалентные одношаговые оценки (см. раздел 3.2) или оценки иных видов. Среди последних популярными [52, гл. 2] являются несмещенные оценки. При конечном объеме выборки оценки максимального правдоподобия в ряде случаев хуже несмещенных оценок [53]. Основная идея использования несмещенных оценок состоит по Колмогорову [30, с. 340–363] в следующем. Во многих важных случаях эти оценки существуют. С другой стороны, чрезмерное разнообразие несмещенных оценок может быть значительно сокращено, если воспользоваться несмещенными оценками, которые выражаются через надлежащим образом выбранные достаточные статистики. Надо употреблять только несмещенные оценки, выражающиеся через достаточные статистики: оказывается, что при этом мы не суживаем круг задач, в котором несмещенные оценки существуют, и при переходе от произвольной (даже плохой) несмещенной оценки к осредненной оценке, выражающейся через достаточную статистику, мы можем только уменьшить дисперсию оценки. Имеет место [52, гл. 2] теорема Рао — Блекуэлла — Колмогорова: оптимальная оценка, если она существует, является функцией от достаточной статистики.

А.Н. Колмогоров первым ([30, с. 340–363], [46]) применил несмещенные оценки в задачах статистического контроля. Он впервые использовал несмещенные оценки для определения эффективности реально используемых планов выборочного контроля по альтернативному признаку. На основе идей А.Н. Колмогорова рядом авторов были построены несмещенные оценки для предъявленного и пропущенного брака, для априорного распределения числа дефектных изделий в контролируемых партиях, а также получены несмещенные оценки при контроле по альтернативному и количественному признакам (см. [49], а также комментарии Ю.К. Беляева и Я.П. Лумельского в [30, с. 522–523]). Несмещенные оценки основных показателей контроля включены в некоторые государственные стандарты (ГОСТ 24660-81, например).

Полученная А.Н. Колмогоровым несмещенная оценка плотности нормального распределения нашла широкое применение в задачах контроля по количественному признаку. В дальнейшем этот результат был перенесен на многомерное нормальное распределение, а также применен для задач статистической классификации. Метод проверки гипотез по совокупности малых выборок, разработанный нами в [45], также основан на использовании несмещенных оценок. Этот метод применяется при статистическом приемочном контроле по нескольким альтернативным признакам [16, раздел 13.5]. Отметим, что в этом случае оказывается нецелесообразным переход к осредненной оценке, выражающейся через достаточную статистику.

Введенные А.Н. Колмогоровым верхние и нижние оценки могут быть использованы и в тех случаях, когда несмещенные оценки не существуют. Именно так обстоит дело при оценивании пропущенного брака при биномиальном распределении и плане одноступенчатого контроля. Рядом авторов были получены верхние и нижние оценки функций неизвестных параметров, а также оценки с минимальным смещением.

О логнормальном законе распределения. В 1940 г. Н.К. Разумовский привел много случаев, в которых логарифмы размеров частиц (золотин в золотоносных россыпях, частиц горных пород при их дроблении и т.п.) приближенно подчиняются нормальному закону распределения. В 1941 г. А.Н. Колмогоров указал общую схему случайного процесса последовательного дробления частиц, при которой в пределе, при неограниченном продолжении дробления, нормальный закон для логарифмов размеров частиц может быть установлен теоретически [30, с. 264–266]. (Напомним, что положительная случайная величина X имеет логнормальный закон распределения, если логарифм величины X имеет нормальный закон распределения; условия, при которых вероятностная модель приводит к нормальному закону, хорошо известны.)

Обнаружение различий. В 70–80-х гг. XX в. под научным руководством А.Н. Колмогорова на механико-математическом факультете МГУ им. М.В. Ломоносова работала группа исследователей, занимавшаяся статистическим анализом эффективности экспериментальных методов управления погодой. Речь идет об изменении количеств выпавших осадков, борьбе с градом и рассеянии туманов. Среди прочих [54] вероятностных моделей использовалась и следующая.

Имеется n объектов $U_1, U_2, \ldots, U_n$ и с каждым объектом U_k связана пара чисел a_k и b_k, $k = 1, 2, \ldots, n$. Пусть $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ — последовательность независимых случайных величин, причем величина ε_k принимает значение 1 (считаем, что имеет место воздействие) с вероятностью p_k и значение 0 (воздействие отсутствует) с вероятностью $q_k = 1 - p_k$, $k = 1, 2, \ldots, n$. В результате наблюдений над объектами нам известны значения случайных пар (ε_k, X_k), $k = 1, 2, \ldots, n$, где $X_k = a_k$ при $\varepsilon_k = 0$ и $X_k = b_k$ при $\varepsilon_k = 1$. Задача состоит в сравнении двух последовательностей $a^{(n)} = (a_1, a_2, \ldots, a_n)$ и $b^{(n)} = (b_1, b_2, \ldots, b_n)$. Тем самым в этой модели (могущей быть использованной и в других случаях, когда необходимо установить наличие или отсутствие эффекта воздействия) предполагается, что числа a_k и b_k неслучайны и вся случайность связана с процессом рандомизации. С помощью оценок Горвица — Томпсона и их обобщений [55] можно построить [54] ряд статистических критериев для проверки гипотезы

$$H(n): A(n) = B(n),$$

где

$$A(n) = \sum\nolimits_{k=1}^{k=n} a_k, \quad B(n) = \sum\nolimits_{k=1}^{k=n} b_k .$$

А.Н. Колмогоров заметил, что дисперсии оценок в критериях могут быть заметно уменьшены, если имеются хорошие методы прогноза, позволяющие до начала наблюдений указывать оценки a_k* и b_k* для a_k и b_k соответственно. Полагая $a_k = a_k^* + \Delta a_k$ и $b_k = b_k^* + \Delta b_k$, мы можем упомянутые выше процедуры применить не к a_k и b_k, а к Δa_k и Δb_k. При этом получаются оценки, правильные независимо от качества прогноза, но они будут лучше оценок без обращения к прогнозам лишь в случае хороших прогнозов, когда величины $|\Delta a_k|$ и $|\Delta b_k|$ значительно меньше $|a_k|$ и $|b_k|$ соответственно.

Упомянем также работы А.Н. Колмогорова по теории стрельбы, выполненные в военные годы, по генетике и биологии, по лингвистике. Особенно велик вклад, сделанный А.Н. Колмогоровым и его учениками (М.Д. Миллионщиков, А.С. Монин, А.М. Обухов, А.М. Яглом и др.) в теорию турбулентности. Здесь прежде всего следует упомянуть знаменитый колмогоровский «закон двух третей» о распределении энергии в спектре турбулентности, полученный из простых соображений размерности (подробнее см. [28, с. 445, 475; 29]).

Бесспорно, что многие работы А.Н. Колмогорова [29–31] представляют несомненный интерес для всех, кто разрабатывает или применяет статистические методы. Его мысли еще долго будут приносить нам всем практическую пользу. Отечественная вероятностно-статистическая научная школа порождена идеями А.Н. Колмогорова. Это хорошо видно на примере работ его ученика Б.В. Гнеденко.

Статистические методы в работах Бориса Владимировича Гнеденко. При анализе актуальных для XXI в. работ академика АН УССР Б.В. Гнеденко (1912–1995) основное внимание уделим предельным теоремам теории вероятностей, математической статистике, теории надежности, статистическим методам управления качеством и теории массового обслуживания. Одна из основных научных заслуг Б.В. Гнеденко — обоснование необходимости развития статистических методов как самостоятельного научного направления, подробное рассмотрение ряда проблем, относящихся к этому направлению.

В XXI в. наиболее ценным для нас является удивительное умение Б.В. Гнеденко (далее — Б.В.) объединить в своем творчестве глубокие теоретические изыскания и практические разработки. В настоящее время все глубже становится разрыв между внутриматематическими изысканиями, от которых в обозримом будущем нельзя ждать практической пользы, и попытками решения прикладных задач методами, устаревшими на полвека. Уникальность Б.В. и состоит в том, что он своей личностью устранял этот пагубный разрыв. Он был одновременно великим теоретиком и великим прикладником. Чем больше проходит времени с того момента, как Б.В. завершил свои труды, тем яснее становится основополагающая роль его идей, его методологического подхода в нашей нынешней работе. Научный путь Б.В. заслуживает подробного осмысления.

Из теоретических исследований Б.В. больше всего известны работы по предельным теоремам теории вероятностей, в том числе классическая монография о суммах независимых случайных величин 1949 г., написанная совместно с А.Н. Колмогоровым, статьи по предельным распределениям крайних членов вариационного ряда. Основополагающие результаты получены им в математической статистике, например, в задаче проверки однородности двух выборок. Для прикладников Б.В. — лидер в области теории надежности, массового обслуживания, статистических методов управления качеством продукции. По его «Курсу теории вероятностей» учились многие поколения специалистов. Большое значение имеют работы по истории науки и по другим направлениям, среди которых особенно выделяется методология научных исследований.

От практики — к теории, от теории — к практике (четыре этапа научного пути). Научный путь Б.В. можно разбить на четыре этапа. Первый (1930–1934) прошел на кафедре математики текстильного института в г. Иваново, куда он был направлен в 1930 г. после окончания Саратовского университета. Именно там Б.В. пришел к глубокому убеждению, что полноценная творческая жизнь математика связана с широким использованием математических методов в решении задач практики и одновременном развитии самих математических методов, без чего невозможно глубокое изучение и удовлетворение потребностей практики. В ивановский период он увлекся теорией вероятностей.

Второй этап (1934–1945) — исследовательская работа в Москве. В 1934 г. Б.В. поступил в аспирантуру Московского университета. Его научными руководителями стали А.Я. Хинчин и А.Н. Колмогоров. Еженедельно собирался общегородской семинар по теории вероятностей, где с новыми результатами выступали известные ученые А.Н. Колмогоров, Е.Е. Слуцкий, Н.В. Смирнов, А.Я. Хинчин, а также аспиранты, молодые физики, биологи и инженеры. Б.В. увлекся предельными теоремами для сумм независимых случайных величин. В июне 1937 г. он защитил кандидатскую диссертацию «О некоторых результатах по теории безгранично-делимых распределений», а в начале июня

1941 г. — докторскую диссертацию, состоящую из двух частей: теории суммирования независимых случайных величин и теории распределения максимального члена вариационного ряда. В годы Великой Отечественной войны Б.В. Гнеденко принимал активное участие в решении многочисленных задач, связанных с обороной страны.

Третий этап научного пути Б.В. — украинский (1945–1960). В 1945 г. Академия наук Украинской ССР избрала Б.В. Гнеденко своим членом-корреспондентом и направила во Львов, где он восстанавливал Львовский университет и организовывал учреждения Академии наук УССР. Во Львове Б.В. Гнеденко читал разнообразные курсы: математический анализ, вариационное исчисление, теорию аналитических функций, теорию вероятностей, математическую статистику и др. Его научная работа в этот период также была весьма разнообразна. Ему удалось доказать в окончательной формулировке локальную предельную теорему для независимых, одинаково распределенных решетчатых слагаемых (1948 г.). Здесь начались исследования по непараметрическим методам статистики. Но, по нашему мнению, основное значение имела работа Б.В. Гнеденко над учебником «Курс теории вероятностей» [41] (первое издание — 1949 г.) и монографией «Предельные распределения для сумм независимых случайных величин» [56].

В 1950 г. Президиум АН УССР перевел Б.В. в Киев, где в Институте математики АН УССР был организован отдел теории вероятностей и математической статистики. Одновременно Б.В. заведовал кафедрой математического анализа в Киевском университете.

Естественно, что очень скоро вокруг него образовалась группа молодых ученых, увлекшаяся теорией вероятностей и задачами математической статистики. Первыми киевскими учениками Б.В. были В.С. Королюк и В.С. Михалевич, впоследствии известные ученые. Характерно для Б.В., что в Киеве он организовал городской семинар по истории математики при Институте математики АН УССР. Он объединил многих ученых, работающих в области истории науки.

В 1953–1954 гг. Б.В. работал в ГДР, а по возвращении Президиум АН УССР поручил ему возглавить работу по организации Вычислительного центра. Ядром группы ученых были сотрудники академика С.А. Лебедева, разработчика первой в Европе ЭВМ, получившей название МЭСМ (малая электронная счетная машина). Одновременно Б.В. возглавил работу по созданию курса программирования для ЭВМ, который начал читать студентам Киевского университета — будущим сотрудникам Вычислительного центра. Этот курс [57] — первая в СССР книга по программированию. Начались работы по проектированию универсальной машины «Киев» и специализированной машины для решения систем линейных алгебраических уравнений. В этот период Президиум АН УССР возложил на Б.В. Гнеденко обязанности директора Института математики АН УССР и председателя бюро физико-математического отделения.

Широкая организационная деятельность не ослабила научной и педагогической деятельности Б. В. Гнеденко. Именно к этому периоду относится начало разработки им двух новых направлений прикладных научных исследований — теории массового обслуживания и вопросов использования математических методов в современной медицине.

Четвертый этап научного пути (1960–1995) — снова Москва. В 1960 г. Б.В. переехал в Москву и возобновил работу в Московском университете. Сразу же Б.В. организовал московский семинар по математической теории надежности и теории массового обслуживания, привлекший многочисленных участников. Большое внимание Б.В. уделял разработке основ теории надежности, решению задач теории резервирования с восстановлением, оптимальной профилактики, управлению качеством промышленной продукции в процессе производства.

В 1965 г. А.Н. Колмогоров передает Б.В. руководство кафедрой теории вероятностей механико-математического факультета МГУ, которой Б.В. заведовал до своих последних дней.

Методологическими проблемами математики Б.В. систематически интересовался с конца 1950-х годов. Он — член научного совета при Президиуме АН СССР по философским проблемам естествознания. С первых дней Общества по распространению научных и политических знаний (общество «Знание») он принимает активное участие в его работе. Жизненному и научному пути Б.В. посвящена статья [58] и другие публикации.

Общее количество опубликованных научных трудов Б.В. — около тысячи. Рассмотрим подробнее основные направления его научной деятельности.

Суммирование независимых случайных величин. В 30-е годы внимание Б.В. привлекли задачи, связанные с суммированием независимых случайных величин (с.в.). Интерес к таким задачам появился в математике еще в XVII в. Невозможность прямых вычислений распределений сумм независимых с.в. приводит к необходимости получения и изучения асимптотических формул для них, т. е. таких формул, которые позволяют находить с нужной точностью требующиеся нам вероятности, связанные с суммами с.в. Эти формулы даются предельными теоремами теории вероятностей. Таким образом, аппроксимация многократных сверток распределений потребовала развития глубокой математической теории, которая называется теорией предельных теорем для сумм независимых с.в. или теорией суммирования.

Начало развития этой теории связано с работами Я. Бернулли и А. Муавра начала XVIII в., в которых были доказаны закон больших чисел (ЗБЧ) и центральная предельная теорема (ЦПТ) для независимых с.в., принимающих два значения. Эти исследования были продолжены в XIX веке П. Лапласом, С. Пуассоном, К. Гауссом и другими учеными, но вплоть до 1860-х гг. рассматривались лишь с.в., принимающие два значения. Лишь в 1867 г. П.Л. Чебышев получил ЗБЧ в общем виде, а достаточно общая форма ЦПТ была найдена лишь в работах А.М. Ляпунова и А.А. Маркова на рубеже XIX–XX вв. Наиболее бурное развитие теории суммирования пришлось на 20–40 гг. XX в. и связано с именами А.Н. Колмогорова, Б.В. Гнеденко, А.Я. Хинчина, П. Леви, В. Феллера и Дж. Линдеберга.

Класс возможных предельных распределений для сумм независимых случайных величин, как показали А.Я. Хинчин и Г.М. Бавли, совпадает с классом безгранично-делимых распределений. Оставалось выяснить условия существования предельных распределений и условия сходимости к каждому возможному предельному распределению. Заслуга постановки этих задач и их решения принадлежит Б.В. Он в 1937 г. предложил оригинальный метод, получивший название метода сопровождающих безгранично-делимых законов. Единым приемом удалось получить все ранее найденные в этой области результаты, а также и ряд новых.

В теории суммирования доказывались как интегральные предельные теоремы, то есть теоремы о сходимости ф.р., так и локальные теоремы, то есть теоремы о сходимости плотностей (для гладких распределений) и об асимптотическом поведении вероятностей отдельных значений для решетчатых распределений. В 20–40 гг. XX в. были получены исчерпывающие результаты о ЗБЧ в классической формулировке. Отметим, что законы больших чисел в пространствах нечисловой природы, найденные в последней четверти XX в., формулировались и доказывались исходя из совсем иных подходов — не на основе суммирования, а на основе решений оптимизационных задач (см., например, [16, 33]).

Во всех разделах теории суммирования Б.В. получил фундаментальные результаты, пролившие свет на существо дела. Итогом развития классической теории суммирования явилась публикация в 1949 г. монографии Б.В. Гнеденко и А.Н. Колмогорова [56], которую можно назвать монументом создателям этой теории. Методы и результаты теории суммирования применяются в различных разделах теории вероятностей, статистических методов и их применений, а книга [56] остается источником новых идей для многих исследователей. Эта книга — одно из наиболее замечательных достижений математики XX в.

Предельные теоремы для крайних порядковых и разделимых статистик. Работы по предельным теоремам для крайних порядковых статистик публикуются уже в течение почти сотни лет, начиная с двадцатых годов XX в. Среди авторов таких публикаций: Додж, фон Мизес, Фреше, Фишер и Типпет, Б. де Финетти, Гумбель. В.Б. Невзоров и другие. Здесь наиболее полные и глубокие результаты получены Б.В. [59].

Пусть x_1, ..., x_n — независимые одинаково распределенные с функцией распределения F случайные величины; тогда величины $x_n(1) = \min x_k$ и $x_n(n) = \max x_k$ называются крайними (или экстремальными) порядковыми статистиками, а также крайними членами вариационного ряда. Предположим, что для функции распределения F найдутся последовательности констант $\{a_n > 0\}, \{b_n\}$, для которых существуют невырожденные предельные (с ростом n) функции распределения G крайних членов преобразованной выборки $\{a_n^{-1}(x_k - b_n)\}$. Тогда, согласно общей теории, функция G имеет один из трех типов. Среди них широко используемое на практике распределение Вейбулла — Гнеденко [60]. Борисом Владимировичем получены необходимые и достаточные условия, относящиеся к F, чтобы получить тот или иной тип G.

Являясь выдающимся специалистом по теории суммирования независимых случайных величин, Б.В. решил результаты этой теории применить к суммированию зависимых случайных величин. Поэтому он проявил интерес [61] к таким случайным величинам $w_1,\dots,w_N$, совместное распределение которых совпадает с условным совместным распределением некоторых независимых случайных величин $\theta_1,\dots,\theta_N$, при условии фиксации суммы последних в некоторой точке. Отправляясь от величин $w_1,\dots,w_N$, можно построить [61] класс сумм зависимых случайных величин, называемых в отечественной литературе разделимыми статистиками. Распределения последних известным образом выражаются через распределения сумм соответствующих независимых случайных величин (векторов). Тем самым, для получения предельных (с ростом числа слагаемых) теорем для разделимых статистик надо воспользоваться результатами суммирования независимых величин или их многомерными аналогами — в случае векторов.

Теория массового обслуживания. Большим и весьма практически важным разделом современных статистических методов, в становление и развитие которого Б.В. внес неоценимый вклад, является теория массового обслуживания (ТМО). Первый цикл работ в этом направлении он выполнил в Иванове. В частности, он занимался изучением связи неровноты пряжи по номеру и весу, выяснением эффективности перехода от обслуживания одного станка к обслуживанию нескольких станков, оценкой длины среднего перехода между станками, который выполняет ткачиха в процессе обслуживания ткацких станков, выявлением особенностей метода станкообходов для нормирования рабочего времени станка и рабочего. Этой тематике посвящена первая книга Б.В. [61].

В опубликованной перед самой войной работе [62] Б.В. решает задачу определения среднего числа зарегистрированных счетчиком Гейгера-Мюллера частиц (известно, что в силу наличия «мертвой зоны» счетчик Гейгера-Мюллера регистрирует не все попадающие в него частицы). В терминах ТМО рассматриваемая модель может быть описана как однолинейная система массового обслуживания (СМО) с потерями, нестационарным пуассоновским входящим потоком и постоянным временем обслуживания. Заметим, что и к настоящему времени СМО с нестационарным входящим потоком исследованы крайне мало.

К задачам ТМО Б.В. возвращается в 1950-х гг, хотя, по собственному признанию, уже во время войны он не раз размышлял над ними. И теперь это направление, наряду с теорией суммирования и математической теорией надежности, становится одним из основных в его научной деятельности. Б.В. обобщает формулы Эрланга на системы с ненадежными восстанавливаемыми приборами, рассматривая как случай с потерей требования при отказе прибора, так и случай перехода недообслуженного требования на другой свободный прибор, и т.д.

В 1956 г. Б.В. прочитал первый в СССР спецкурс по ТМО. В 1958 г. цикл его лекций по теории массового обслуживания был опубликован, а затем послужил основой для широко известной монографии [63], выпущенной в 1966 г. Эта книга и до сих пор остается одной из основополагающих при подготовке специалистов по ТМО не только в нашей стране, но и за рубежом. Отметим еще две его монографии ([64, 65]), оказавших значительное влияние на развитие ТМО.

В последующие годы Б.В. опубликовал еще более 30 статей, относящихся к ТМО. В этих статьях, наряду с решением отдельных задач по ТМО, он дает детальные обзоры существующих методов исследования, формулирует новые проблемные направления. Важнейшей задачей Б.В. считал пропаганду на всех уровнях, начиная от школьников и кончая профессиональными математиками и управленцами высокого уровня, широчайшего внедрения методов ТМО в инженерную практику.

О работах Б.В. в области математической статистики, теории надежности и контроля качества. Статистические методы были в центре научных и педагогических интересов Б.В. на протяжении всей его творческой жизни. «Каждому специалисту нужно знать математическую статистику» — так называется одна из его статей [66]. Уже в первых его публикациях, посвященных математическому анализу проблем текстильного производства, проявился живой интерес и умение Б.В. работать с реальными данными.

Мировую известность Б.В. как статистику принес цикл работ, выполненный им вместе со своими учениками и сотрудниками в конце 1940-х — первой половине 1950-х гг. Он изучал проблему проверки гипотезы однородности двух независимых выборок с помощью статистики, равной максимуму разности соответствующих эмпирических функций распределения (т.н. двухвыборочная односторонняя статистика Н.В. Смирнова). Б.В. предложил метод вычисления точного распределения статистики критерия для конечных выборок равного объема, позволивший получить простое доказательство найденных ранее Н.В. Смирновым предельных теорем и достаточно точные асимптотические разложения. А.Н. Колмогоров высоко оценил исследования Б.В. по непараметрической статистике [67]. И сейчас, через 50 лет, эти результаты Б.В. попрежнему актуальны для применения математических методов исследования (см., например, раздел 14.6 выше).

По статистике Б.В. опубликовал более 50 работ. Среди них — посвященные проблемам статистического образования, а также приложениям статистических методов в технических исследованиях, теории надежности и контроле качества, экономике и социальных науках, биологии и медицине, во многих других областях.

Б.В. всегда был среди тех ученых, которые, с одной стороны, глубоко понимали необходимость развития вычислительной техники как основы и предпосылки внедрения результатов теоретических (и в том числе математико-статистических) исследований в практику; а с другой — предвидели широкие горизонты новых исследований, которые представляли высокопроизводительные компьютеры. Он не только руководил созданием Вычислительного центра АН УССР, но и был у истоков создания Института кибернетики АН УССР. Как уже отмечалось, Б.В. был написан первый в СССР учебник по программированию [57]. Начатые Б.В. в сотрудничестве с Н.М. Амосовым работы по машинной диагностике сердечных заболеваний во многих своих аспектах являются примером высококлассного прикладного статистического исследования, по своей тематике относящегося к проблемам классификации. К сожалению, Б.В. не дали завершить эти исследования. Являясь одним из виднейших математиков, работавших в то время на Украине, он был вынужден покинуть Киев и переехать в 1960 г. в Москву.

Вопросами теории надежности и проблемами управления (а значит, и контроля) качества Б.В. начал заниматься еще во второй половине 1950-х гг. По мере знакомства с уровнем качества продукции промышленных предприятий в нем крепла уверенность в необходимости использования математических методов для объективной оценки качества и прогноза надежности изделий. К разработке математической теории надежности он привлек своих учеников И.Н. Коваленко, В.С. Королюка, Т.П. Марьяновича. Сам Б.В. в это время выполнил ряд прикладных работ, связанных с анализом надежности и методикой расчета нагрузки электрических сетей промышленных предприятий.

В Москве, будучи одним из создателей и признанным лидером советской школы математической теории надежности, Б.В. приобрел огромное неформальное влияние на развитие этой теории не только на всей территории СССР, но и далеко за ее пределами. Другой мощной школой в теории надежности является североамериканская. Две школы отличались по тематике исследований и во многом дополняли друг друга. Достижения этих школ 1960–1980-х гг. до сих пор предопределяют мировое развитие теории надежности.

Продвижению результатов математической теории надежности в практику Б.В. придавал не меньшее значение, чем развитию самой математической теории. По его мнению, важнейшими аспектами востребованности и успешного применения практикой являются:

а) наличие в теории богатого набора математических моделей, отражающих разнообразные явления предметной области;

б) наличие в предметной области специалистов, способных понять математические модели и превратить их в «руководящие указания» на производстве;

в) наличие литературы самого разного уровня, отражающей достижения теории и практику ее применения;

г) возможность прямого контакта между создателями теории и специалистами предметной области для взаимной корректировки задач теории и методов ее приложения в предметной области.

Все перечисленные выше моменты нашли счастливое сочетание в работе огромного незримого коллектива ученых и практиков, имевших отношение к созданию и приложению теории надежности и управлению качеством в СССР. Усилиями Б.В., его сотрудников и учеников с 1960 по 1985 гг. была разработана весьма разветвленная математическая теория надежности и математическая теория контроля качества. Была налажена широкая пропаганда необходимости практического использования теоретических результатов, в том числе по линии общества «Знание». Организованы семинары и лекционные курсы в Политехническом музее, в МГУ им. М.В. Ломоносова, а затем и во многих городах СССР, где инженерный состав получал необходимую математическую подготовку для понимания и применения методов теории надежности и контроля качества. В кабинете надежности при Политехническом музее все заинтересованные лица могли получить консультации у ведущих специалистов, включая и самого Б.В. Издательства «Советское радио» и «Знание» выпустили серию книг, посвященных различным аспектам теории надежности и контроля качества. Огромное влияние оказала основополагающая монография [68], а также ряд других монографий с участием Б.В., в частности, небольшая яркая книга [47].

Была развернута большая работа по подготовке специалистов высшей категории в области теории надежности. В руководстве ряда отраслей промышленности оказались специалисты, хорошо понимающие необходимость внедрения современных методов теории надежности и контроля качества. И во всем этом самое непосредственное участие принимал Б.В. В результате, достижения математической теории надежности и контроля качества нашли широкое признание, как в научных кругах, так и среди прикладников. Правда, с сожалением приходится констатировать, что в целом на реальный подъем качества продукции в стране, за исключением предприятий оборонно-промышленного комплекса, эти достижения сказались мало.

Развитие теории управления качеством и надежностью активно продолжается и в настоящее время. В частности, в журнале «Заводская лаборатория» постоянно обсуждаются различные прикладные и теоретические проблемы управления качеством [7, 19]. В современных условиях реализация накопленного научного потенциала может дать значительное ускорение экономического роста как отдельных предприятий, так и страны в целом.

Конечно, нельзя не отметить и огромный личный вклад Б.В. в математическую теорию надежности. Предметом его наибольшего интереса была теория резервированных систем с восстановлением. Здесь им была поставлена задача, которая имела многочисленные продолжения в работах других математиков, а именно — задача об асимптотическом распределении момента первого отказа резервной группы с быстрым восстановлением. Б.В. удалось установить связь с асимптотической теорией суммирования случайного числа случайных слагаемых. И эта задача была им с блеском решена. Отметим, что подобные суммы используются не только в теории надежности, но и в различных иных прикладных областях, в частности, в логистике, т.е. науке о движении материальных, финансовых и информационных потоков (см., например, раздел 8.3 и монографию [45]).

И как здесь не вспомнить слова Б.В. о взаимообогащении фундаментальных и прикладных наук: «Я глубоко убежден в том, что прикладные проблемы не только дают возможность демонстрации силы математических методов и решения множества задач, необходимых для жизненной практики, но имеют огромное значение для развития самой математики. Дело в том, что в прикладных задачах часто приходится сталкиваться с совсем новыми ситуациями, о которых математик-теоретик не может догадаться. Традиционные методы математики недостаточны для решения возникающих вопросов, требуется разработка новых методов исследования и, возможно, — даже новых ветвей математики. Но практика важна для науки и тем, что именно практика выясняет возможности той или иной области математики для решения актуальных проблем других научных дисциплин и повседневных нужд общества. И, в конечном счете, ценность исследований математика будет определяться по тому, насколько широко и глубоко развиваемые им теории позволяют проникнуть в проблемы познания законов окружающего мира, помогают решению житейских проблем, касающихся всего общества. Чем теснее связана та или иная ветвь математики с практикой жизни, тем разнообразнее ее проблемы, тем быстрее она развивается. Так было, так есть и так будет» [69].

История математики и преподавание. Вскоре после создания Академии педагогических наук РСФСР (основана в 1943 г.) Б.В. был приглашен в Институт методов обучения. Итог его работы — книга [70], адресованная в первую очередь учителям и школьникам. Эта замечательная книга была первым достаточно полным исследованием истории математики в нашей стране.

Несомненной заслугой Б.В. является то, что он показал, что история математики необходима действующему математику. На Третьем Всесоюзном математическом съезде (1956 г.) Б.В. перечислил магистральные направления историко-научных исследований в этой области. Он подчеркнул значение истории математики «а) для целей выяснения общих закономерностей развития математики, б) для выявления общих перспектив ее последующего развития, для выявления методологических установок науки, г) для выяснения связей с другими науками и роли математики в истории культуры, д) для целей преподавания и воспитания» [71, с. 100].

Эти задачи Б.В. реализовывал на протяжении пятидесяти лет, написав более 180 работ по истории математики. Среди них — более 32 биографических статей, посвященных Н.И. Лобачевскому, П.Л. Чебышеву, М.В. Остроградскому, А.Н. Колмогорову и др. В фундаментальной работе [3] он проследил предысторию теории вероятностей, анализируя труды ученых, стоящих у истоков этой науки: Л. Пачолли (основатель бухгалтерского учета), Дж. Кардано, Н. Тартальи, Г. Галилея, Б. Паскаля, П. Ферма, Х. Гюйгенса. Б.В. мастерски умел показать в элементарных рассуждениях предшественников зерна более широких идей. Изложение столь понятно и интересно, что хочется заглянуть в первоисточники — труды Я. Бернулли, П.Л. Чебышева, П. Леви и других.

Наиболее известной книгой Б.В. — учебником «Курс теории вероятностей» — пользуются студенты университетов уже свыше полувека. Он выдержал несколько десятков изданий в СССР, США, ГДР, Японии и многих других странах. Совместно с А.Я. Хинчиным Б.В. написал научно-популярную книгу [72], которая также вот уже более пятидесяти лет пользуется огромной популярностью и выдержала множество изданий в СССР и за рубежом.

Б.В. уделял большое внимание вопросам преподавания. Он руководил научно-исследовательскими семинарами по программированному обучению, по вопросам преподавания в средней школе, был председателем секции теории вероятностей и математической статистики и секции средней школы Московского математического общества. Большое число статей было им опубликовано в журналах «Вестник высшей школы», «Математика в школе», в сборниках научно-методического совета Минвуза СССР.

Лекции Б.В. пользовались большим успехом в любой аудитории. Естественна попытка проанализировать те средства, которые использовал Б.В. для воздействия на слушателей во время лекций. Суть их в простоте, в уважении своих слушателей, в желании передать им те сведения, которые им необходимы; в демонстрации на ярких и доступных примерах важности того, о чем идет речь; в умении связывать общие идеи с различными частными задачами, которые близки интересам слушателей; в ненавязчивом, постоянном воспитании научного мировоззрения. И все это вместе взятое высказывалось Б. В. Гнеденко на лекциях так, что в каждый момент звучало нужное слово с нужной интонацией.

Охватывая в своем творчестве весь диапазон, который может попасть в поле зрения математика — от исходной практической проблемы до теоретической чисто математической задачи и затем от решения этой задачи обратно к практической проблеме — Б.В. вполне естественно обращался к осмыслению своего пути исследователя. Он посвящал методологическим исследованиям отдельные работы, постоянно обращался к проблемам таких исследований в книгах более общего характера [73]. Методологические вопросы постоянно обсуждались также в публикациях, посвященных роли математических методов исследования в научно-техническом прогрессе [74] или применению современных статистических методов в управлении качеством продукции [47].

Своей личностью, своей собственной научной, педагогической и организационной работой Б.В. Гнеденко показывал пример плодотворного единения теории и практики. И символично, что он в 1961 г. создал раздел «Математические методы исследования» в журнале «Заводская лаборатория» и возглавлял его более 30 лет. И сейчас для нас важны его выступления на страницах этого журнала [74, 75], в котором публикуются основные отечественные работы по статистическим методам.

В довоенный период советская вероятностно-статистическая наука прославилась двумя достижениями. Об одном — построении А.Н. Колмогоровым теории вероятностей на основе теории меры и интеграла Лебега — уже говорилось. Второе — разработка непараметрических критериев проверки согласия и однородности. Сначала фундаментальный результат — критерий согласия эмпирического с распределения с теоретическим (критерий Колмогорова) — был получен А.Н. Колмогоровым [30, с. 134–141], затем дело взял в свои руки член-корреспондент АН СССР Николай Васильевич Смирнов (1900–1966).

О работах Н.В. Смирнова. Его основные научные труды опубликованы в сборнике [76], на который и будем ссылаться. Наиболее ценная книга XX в. по статистическим методам, на наш взгляд, подготовлена членами-корреспондентами АН СССР Л.Н. Большевым и Н.В. Смирновым. Это — «Таблицы математической статистики» [14]. Название не должно обманывать — весьма полезна начинающая книгу пояснительная часть (разделы с кратким и строжайше выверенным описанием классических статистических методов, примерами их применения, комментариями к таблицам). Учебники Н.В. Смирнова по статистическим методам и по сей день остаются среди лучших [77, 78].

Мы уже упоминали, что с работы Н.В. Смирнова 1951 г. «О приближении плотностей распределения случайной величины» [15; 76, с. 205–223] началось развитие такого перспективного, в том числе в статистике нечисловых данных [33, гл. 11], направления, как непараметрические оценки плотности. Однако с его именем связывают «критерии Смирнова». Пусть $F_n(t)$ — эмпирическая функция распределения, построенная по выборке объема n из непрерывной функции распределения $F(t)$. Напомним, что согласно Л.Н. Большеву и Н.В. Смирнову [14] значение эмпирической функции распределения в точке x равно доле результатов наблюдений в выборке, меньших x. Одновыборочные критерии Смирнова, введенные в статье 1939 г. «Об уклонениях эмпирической функции распределения» [76, с. 88–107], основаны на статистиках

$$D_n^- = \inf_{-\infty<t<+\infty}(F_n(t)-F(t)), \quad D_n^+ = \sup_{-\infty<t<+\infty}(F_n(t)-F(t)).$$

Очевидно, критерий Колмогорова есть максимум этих двух статистик. Поэтому возникает желание объединить все три критерия в одну группу — группу критериев Колмогорова — Смирнова. Однако разработанные Н.В. Смирновым методы рассуждений, использованные для получения распределений рассматриваемых статистик, совершенно оригинальны. Они не имеют ничего общего с подходом А.Н. Колмогорова. Поэтому мы считаем, что надо говорить отдельно о критерии Колмогорова и отдельно о критериях Смирнова, а если уж надо объединить их вместе, то говорить о критериях *типа* Колмогорова-Смирнова, но не о критериях Колмогорова-Смирнова, поскольку употребление последнего выражения приводит к искажению исторической правды [35].

Двухвыборочные критерии Смирнова однородности двух независимых выборок были им предложены и изучены в 1939 г. (см. [76, с. 117–127]). Единственное ограничение — функции распределения $F(x)$ и $G(x)$ должны быть непрерывными. Критерии Смирнова основан на использовании эмпирических функций распределения $F_m(x)$ и $G_n(x)$, построенных по первой и второй выборкам соответственно. Значение двухсторонней статистики Смирнова

$$D_{m,n} = \sup_x |F_m(x) - G_n(x)|$$

сравнивают с соответствующим критическим значением (см., например, [14]) и по результатам сравнения принимают или отклоняют гипотезу H_0 о совпадении (однородности) функций распределения (подробнее — в главе 5). Практически значение статистики $D_{m,n}$ рекомендуется согласно [14] вычислять по формулам

$$D_{m,n}^+ = \max_{1\le r\le n}\left[\frac{r}{n} - F_m(y_r')\right] = \max_{1\le r\le m}\left[G_n(x_s') - \frac{s-1}{m}\right],$$

$$D^-_{m,n} = \max_{1\le r\le n}\left[F_m(y'_r) - \frac{r-1}{n}\right] = \max_{1\le r\le m}\left[\frac{s}{m} - G_n(x'_s)\right],$$

$$D_{m,n} = \max(D^+_{m,n}, D^-_{m,n}),$$

где $x'_1<x'_2<\ldots<x'_m$ — элементы первой выборки $x_1, x_2, \ldots, x_m$, переставленные в порядке возрастания, а $y'_1<y'_2<\ldots<y'_n$ — элементы второй выборки $y_1, y_2, \ldots, y_n$, также переставленные в порядке возрастания. Поскольку функции распределения $F(x)$ и $G(x)$ предполагаются непрерывными, то вероятность совпадения каких-либо выборочных значений равна 0. Статистики $D^-_{m,n}$, $D^+_{m,n}$ также могут быть использованы для проверки однородности двух независимых выборок. Их называют двухвыборочными односторонними статистиками Смирнова.

Статистика омега-квадрат (подробнее см. о ней в [45, глава 2.3])

$$\omega_n^2 = n\int_{-\infty}^{+\infty}(F_n(x) - F(x))^2 dF(x)$$

также используется для проверки согласия эмпирического распределения с фиксированным теоретическим. Эту статистику в 1928–1931 гг. предлагали использовать Г. Крамер и Р. фон Мизес, однако ее предельное распределение вычислил в 1937 г. Н.В. Смирнов в статье «О распределении ω^2 — критерия Мизеса» [76, с. 60–78], что и позволило использовать эту статистику в практических расчетах. Поэтому статистику ω^2 обычно называют также статистикой Крамера — Мизеса — Смирнова. Имеющаяся в статье [76, с. 60–78] погрешность в формулировке леммы 6 (с. 75, формула (97)) (пропущен множитель $(-1)^k$ из-за неправильного применения теории функций комплексного переменного) исправлена нами в статье [79].

Как следует из сказанного выше, А.Н. Колмогоров и Б.В. Гнеденко внесли огромный вклад в развитие статистических методов. Однако они занимались и многими другими проблемами (особенно А.Н. Колмогоров). Полностью посвятили себя статистическим методам в XX в. только два исследователя с академическими званиями — члены-корреспонденты АН СССР Н.В. Смирнов и Л.Н. Большев.

Логин Николаевич Большев (1922–1978) до конца Великой Отечественной войны участвовал в боевых действиях как летчик-истребитель. В 1951 г. окончил механико-математический факультет МГУ им. М.В. Ломоносова, будучи учеником А.Н. Колмогорова. Затем стал сотрудником Математического института АН СССР, в котором работал бок о Н.В. Смирновым, которого и сменил в 1966 г. на посту руководителя отдела математической статистики. Для работ Л.Н. Большева [80] характерно сочетание высокого математического уровня с направленностью на практические приложения статистических методов. Его безвременная кончина обозначила рубеж, после которого разрыв между математической статистикой и статистическими методами (включая прикладную статистику) стал в сложившихся отечественных условиях неизбежным.

Профессор В.В. Налимов как организатор науки. Профессор МГУ им. Ломоносова, доктор технических наук Василий Васильевич Налимов (1910–1997), далее В.В., —создатель и руководитель нескольких новых научных направлений: метрологии количественного анализа, химической кибернетики, математической теории эксперимента и наукометрии. Занимался проблемами математизации биологии, анализом оснований экологического прогноза, вероятностными аспектами эволюции, проблемами языка и мышления, философией и методологией науки, проблемами человека в современной науке, вероятностной теорией смыслов. Свой жизненный путь описал в книге [81].

Известность пришла к В.В. после выхода книги «Применение математической статистики при анализе вещества» [82] — справочника по применению классических статистических методов в работе химиков-аналитиков. Поскольку В.В. пришел в статистические методы не из математики, а из практической деятельности в заводских лабораториях, то и книга его была ориентирована на потребности практики.

Следующим шагом было создание секции «Математические методы исследования» в журнале «Заводская лаборатория». Сейчас под названием журнала стоит: «Ежемесячный научно-технический журнал по аналитической химии, физическим, математическим и механическим методам исследования, а также сертификации материалов». У истоков секции стояли Б.В. Гнеденко и В.В., однако реально работу секции организовывал В.В. Налимов. Под его руководством она стала и остается поныне штабом развертывания исследований по статистическим методам в нашей стране.

В соответствии с тематикой журнала публикации секции посвящены в основном статистическим методам анализа данных измерений, наблюдений, испытаний, анализов, опытов. Большое значение придается математическим методам планирования экспериментов. В частности, при оптимизации технологических процессов в металлургической, химико-технологической, фармацевтической и иных отраслях промышленности применение методов экстремального планирования экспериментов позволяет заметно повысить выход продукта, обычно на 30–300 %.

Основные направления работы секции — прикладная статистика и планирование эксперимента. В первом из них принимается, что экспериментатор не может выбирать точки (значения факторов), в которых проводятся измерения, во втором, напротив, выбор возможен, и основная задача — оптимальный подбор таких точек. Большое внимание уделяется вопросам оптимального управления технологическими процессами, в частности, статистическим методам управления качеством продукции. Рассматриваются также теория и практика экспертных оценок, применение нечетких множеств и др.

Заслугой В.В. является то, что в 1960–1970-е гг. в нашей стране была создана мощная научно-практическая школа в области планирования эксперимента. Перу В.В. принадлежит длинный ряд статей и книг, посвященный развитию теории и практики планирования эксперимента [83–85]. Итоги развития этой области статистических методов подведены учениками В.В. [86], ее математическим основам посвящен справочник [87].

В 1961 г. была создана секция «Химическая кибернетика» (под председательством В.В.) в Научном совете по комплексной проблеме «Кибернетика» при Президиуме АН СССР. С 1971 г. В.В. возглавлял секцию «Математическая теория эксперимента». Она объединяла более 500 активно действующих специалистов, работавших в академических и отраслевых институтах, вузах и на промышленных предприятиях. Развитие новой отрасли науки отслеживалось методами наукометрии [12], во многом созданной трудами В.В.

В 1965 г. А.Н. Колмогоров организовал в МГУ им. М.В. Ломоносова межфакультетскую Лабораторию статистических методов и пригласил В.В. стать первым его заместителем. Задачи, поставленные перед Лабораторией, формулировались примерно так: изучение и дальнейшая разработка вероятностно-статистических методов; их пропаганда и широкое внедрение в научную, инженерную и медицинскую практику; хоздоговорная деятельность; педагогическая и издательская деятельность; проведение общемосковских семинаров, летних научных школ, участие в конференциях [81, с. 272]. Штатный состав достигал 130 человек. Такого мощного научного института-лидера не было в нашей стране. Нет и сейчас.

Организационным структурам, занимавшимся развитием статистических методов в нашей стране, не удалось укрепиться.

Большим успехом было введение в начале 1970-х гг. преподавания в вузах химической кибернетики и создание соответствующих кафедр. Однако через год последовало решение о сокращении штатов, и эти вновь введенные кафедры перестали существовать.

Ректор МГУ им. М.В. Ломоносова академик И.Г. Петровский поддерживал создание и развитие межфакультетской Лаборатории статистических методов А.Н. Колмогорова. Однако после его смерти выяснилось, что эта Лаборатория существует нелегально, не входит в структуру университета. И в 1975 г. Лаборатория была расформирована. Ее сотрудники были распределены между пятью факультетами университета. Оказался уничтоженным единственный в нашей стране центр, занимавшийся методологическими аспектами вероятностно-статистического моделирования [81, с. 291]. И это резко отрицательно сказалось на уровне отечественных прикладных работ.

В июле 1959 г. при Президиуме АН СССР был создан Совет по кибернетике, который возглавил академик А.И. Берг. Инженер-адмирал (высшее флотское звание) Аксель Иванович Берг (1893–1979) работал в области создания, развития и применения радиолокации и современных систем радионавигации, над проблемами кибернетики, став крупнейшим специалистом в основных областях этой отрасли науки. Как уже отмечалось, около 20 лет А.И. Берг поддерживал развитие статистических методов. А после его смерти новое руководство Совета «перекрыло кислород» этой тематике.

После смерти в 1978 г. члена-корреспондента АН СССР Л.Н. Большева резко сократилось сотрудничество между математиками и статистиками, разошлись пути математической и прикладной статистики.

Все эти события второй половины 1970-х гг. способствовали тому, что интересы В.В. сместились из научно-организационной деятельности в сферу его личных научных интересов. В книге «Вероятностная модель языка» [88] В.В. развивает мысль о нечеткости слов в естественном языке (ср. с констатацией «Мы мыслим нечетко» в статье [89]). Затем в длинной серии публикаций В.В. развивает *вероятностно ориентированную философию,* включая вероятностное исчисление смыслов [90]. Последняя научная книга В.В. «В поисках иных смыслов» [91] начинается так: «Основная задача автора состоит в том, чтобы показать, что в наше время — в век утраты фундаментальных смыслов и всеобщей разбросанности знаний по отдельным закромам многоликой культуры — все же возможно построение единых, по-прежнему целостно звучащих метафизических систем».

Мы познакомились с основными достижениями пяти выдающихся исследователей — А.Н. Колмогорова, Б.В. Гнеденко, Н.В. Смирнова, Л.Н. Большева, В.В. Налимова. Вместе с ними работали тысячи специалистов. Нельзя не назвать А.Я. Хинчина, С.Н. Бернштейна, Е.Е. Слуцкого, В.С. Немчинова, В.И. Романовского, Г.К. Круга, А.А. Любищева. И многих, многих других. История русской и советской статистики требует дальнейшего изучения, прежде всего потому, что старые дискуссии продолжаются и сейчас. Так, в настоящем учебнике обсуждаются многие из тех проблем, которые волновали В.В. Налимова [81].

15.3. Дискуссия о прикладной статистике

Глазами американцев: российская дискуссия о прикладной статистике Развитие прикладной статистики в нашей стране сопровождалось бурными дискуссиями. Объективный анализ их начального этапа был дан на страницах органа Американской статистической ассоциации. Статья Сэмюеля Котца и Кэтлин Смит «Пространство Хаусдорфа и прикладная статистика: точка зрения ученых СССР» [24] описывает различные взгляды, имеющие распространение и в XXI веке. Чтобы «взглянуть со стороны» на споры 1980-х гг., используем эту статью.

Статья [24] посвящена дискуссии, развернувшейся на страницах советского статистического журнала «Вестник статистики» по вопросам существования и релевантности (уместности) прикладной статистики как самостоятельной научной дисциплины. В ней анализируется содержание четырех писем редактору и редакционных комментариев к ним, которые были опубликованы в этом журнале в период с октября 1985 г. по июнь 1987 г. Основная задача статьи состоит в том, чтобы осветить длительную (продолжающуюся по крайней мере 40 лет) полемику в советской (и российской) статистике между «идеологическими пуристами» и «прагматиками», которая в 1980-е гг. значительно усилилась. Существование разногласий, безусловно, не является новым явлением среди статистиков и в определенной степени оно носит здоровый характер, способствуя выработке критического отношения к предмету. Полемика в 1980-х гг. затрагивает суть предмета в отличие от более ранних этапов, когда она отличалась идеологической направленностью. В 1950–1960-е гг., в период хрущевской оттепели, когда в СССР более свободно начали публиковать статистические данные, в журнале *The American Statistician* («Американский статистик») — органе Американской статистической ассоциации — было опубликовано несколько статей, посвященных различным аспектам советской статистики, как организационным, так и затрагивающим существо предмета.

Советская статистика: 1917–1964 гг. Вопросы развития статистики в СССР с 1917 по 1964 г. довольно подробно освещены в статьях С. Котца [92, 93], прежде всего борьба двух противоположных мнений по вопросу о роли и содержании статистической науки в СССР. Между официальными статистиками Центрального статистического управления (ЦСУ, ныне Росстат) и статистиками — экономистами математической направленности во главе с В.С. Немчиновым (1890–1964) возникли разногласия.

Официальные статистики считали, что статистика представляет собой описательную науку, в задачи которой входит сбор данных по плановой экономике, и что в условиях коммунизма статистику в конечном счете заменит простая бухгалтерия. Противоположных взглядов придерживались практики и статистики теоретической направленности. Они считали, что статистика и теория вероятностей важны в любой области. В 1954 г. на Всесоюзном научном совещании по теоретическим вопросам статистики (см. о нем в [5, с. 243–247]), в работе которой приняли участие ведущие ученые, известный советский математик А.Н. Колмогоров (1903–1987) помог представителям этих двух противостоящих школ прийти к прагматическому компромиссу. На совещании 1954 г. было заявлено, что статистика является самостоятельной общественной наукой и что «она изучает количественный аспект массовых социальных явлений в неразрывном единстве с их качественным аспектом» (см. Котц, [93, с. 136]). Был сделан вывод, что советскую статистику от «буржуазной» статистики отличает акцент на качественном аспекте явлений. Для «буржуазной» статистики, согласно официальной оценке в Советском Союзе, характерен формальный, чисто математический подход к изучению социальных явлений, при котором количественный показатель рассматривается отдельно от качественной основы.

Однако на математическую статистику как часть математики «официальные статистики» покушаться не решились, поскольку математическая статистика использовалась для решения задач обороноспособности страны. Вместе с тем статистические методы в промышленности и технических исследованиях, статистические методы в медицине, химии, геологии, экономике, социологии, психологии, истории и в других конкретных областях остались вне официальной структуры науки и образования. В результате решений совещания 1954 г. работы по этим направлениям шли под иными именами. Использовались термины типа «экономическая кибернетика», «математическое моделирование в медицине» и др. Недаром сборник «Современные проблемы кибернетики (прикладная статистика)», с момента выхода которого говорим о самостоятельности прикладной статистики, имеет название, свидетельствующее о «крыше» нашей науки. Соответственно в вузах не велась подготовка специалистов по статистическим методам в конкретных областях. В результате констатируем отставание на порядок от англоговорящих стран по числу специалистов. В США статистиков больше, чем математиков, а у нас создание Лаборатории статистических методов под руководством А.Н. Колмогорова со штатом в 130 человек рассматривалось как большой успех (в то время в Индии, в институте, которым руководил Махаланобис, работало около 2 000 человек [81, с. 271]).

Разногласия в 1980-е годы. Появление статьи [24] с началом публикации в 1965 г. полупериодического журнала «Ученые записки по статистике» под редакцией Немчинова (т.е. серии сборников статей, выпускавшихся издательствами «Наука», «Статистика», «Финансы и статистика»).

В 1986 г. вышел юбилейный 50-й выпуск «Ученых записок по статистике». В нем опубликовали свои статьи статистики математической ориентации. Многие из них — выпускники и кандидаты наук престижной школы теории вероятностей и математической статистики при МГУ, которую первоначально возглавлял А.Н. Колмогоров, и такой же школы при Ленинградском университете, во главе которой некоторое время стоял Ю.В. Линник. Эти ученые работали в больших городах, в различных институтах, занимающихся вопросами применения прикладной статистики. Ученые выполняли ориентированные на практическое применение работы в теории управления запасами, прикладном многомерном анализе и т.д., однако создается впечатление, что они испытывали желание заниматься вопросами, носящими более математический характер. Эта тенденция нашла свое отражение на страницах сборника «Ученые записки по статистике», в котором постепенно, но постоянно начали публиковать статьи математического и абстрактного характера, что вызвало недовольство среди статистиков различных научно-исследовательских институтов, связанных с органами официальной государственной статистики (в то время — ЦСУ).

В 1983 г. в издательстве «Наука» вышел в свет 45-й том «Ученых записок по статистике», который был скромно озаглавлен «Прикладная статистика», и разразился скандал. Опишем ход развития полемики, проанализировав содержание четырех писем редактору, которые были опубликованы с октября 1985 г. по июль 1987 г. в ежемесячном журнале «Вестник статистики» — органе ЦСУ.

В ответ на публикацию в сборнике «Ученые записки по статистике» многочисленных математических статей абстрактного характера К. Тимофеев (псевдоним) написал сердитое письмо под заголовком «Что же такое прикладная статистика?» [94]. Он утверждал, что термин «прикладная статистика» является абсурдным, так как то, что она якобы описывает, является одной из областей статистической науки, а не новым направлением. Тимофеев заявил: «Из содержания представленных в 45-м томе статей становится совершенно очевидным: название «Прикладная статистика» использовано для того, чтобы в «Ученых записках по статистике» опубликовать материалы, которые к ней (т.е. к статистике) не имеют ни прямого, ни даже косвенного отношения» [94, с. 66]. Кроме этого, он выразил несогласие с рядом приведенных в сборнике математических формул и абстрактных концепций. В частности, он привел цитату из статьи, в которой говорится, что статья посвящена «измеримым отображениям произвольного вероятностного пространства в множество непустых компактов плоскости, снабженное метрикой Хаусдорфа» (метрика Хаусдорфа — одно из расстояний между множествами; критикуемая статья была озаглавлена «Статистика случайных множеств»). Тимофеев не только не захотел перенестись «в другое измерение», он подверг автора критике за то, что он в своей статье сослался на работы зарубежных ученых, а не на работы классиков марксизма-ленинизма и советские статистические источники, а также за то, что он написал работу, не связанную с реальной жизнью. Он с неодобрением указал, что авторы статей, публикуемых в «Ученых записках по статистике», часто ссылаются на свои собственные работы. Он написал: «Создается впечатление, что книга «Прикладная статистика» использована не только для публикации не относящихся к статистике материалов, но и для рекламы и саморекламы некоторых математиков, решивших снискать себе славу в области экономики и статистики» [94, с. 67]. Тимофеев признал, что эти статьи могут представлять определенный интерес для математиков, однако он полагал, что они вряд ли будут полезны в практической работе тем специалистам, на службе у которых, по его мнению, должна быть статистическая наука, а именно статистикам, экономистам и социологам.

Через десять месяцев журнал «Вестник статистики» опубликовал ответ [20] на выступление Тимофеева. Один из авторов, которых критиковал Тимофеев, А. Орлов, написал ответ в таком же резком тоне, и он был опубликован в официальном органе ЦСУ. В своей статье, перед которой было напечатано вступление от редакции, Орлов пункт за пунктом критиковал позицию Тимофеева. Орлов представил себя, как современного статистика. Он написал, что Тимофеев запутался и не знаком с переменами, которые произошли в статистике, и отметил, что термин «прикладная статистика» не является ни новым, ни редко употребляемым. Он используется специалистами различных учреждений по всей стране. А. Орлов провел грань между математической статистикой и прикладной статистикой, добавив, что прикладная математическая статистика является «неотъемлемой частью» прикладной статистики, а прикладная математическая статистика наряду с аналитической статистикой (т.е. частью математической статистики, ориентированной на решение внутриматематических проблем, например, характеризационных задач [95]) составляют математическую статистику, которая является одной из областей математики. Однако Орлов подчеркнул, что прикладная статистика включает и нематематические области, такие, как «методология организации и проведения прикладного статистического исследования и применения его результатов (как планировать исследование, как выбирать вероятностно-статистическую модель, как собирать данные, как подготавливать их к обработке, как представлять результаты обработки и т.д.), а также соответствующее программное обеспечение» [20, с. 54)].

Далее он указал, что интенсивное использование компьютеров в прикладной статистике свидетельствует о том, что в действительности ее можно рассматривать как часть кибернетики.

Орлов привел много примеров использования прикладной статистики в народном хозяйстве, сделав акцент на планировании эксперимента и контроле качества. Он отметил, что благодаря прикладной статистике была получена большая экономия финансовых средств: «Высокая эффективность прикладной статистики естественна — она родилась из практических нужд» [20, с. 54]. Он охарактеризовал большой вклад в практическую работу, который внесли многие из тех статей, которые Тимофеев высмеял за абстрактные заголовки. В заключение статьи он привел таблицу, из которой видно, что ученые, публикующие свои работы в «Ученых записках по статистике», чаще ссылаются на работы советских авторов, чем зарубежных, и он подчеркнул, что эти авторы опираются на опыт своей практической работы, а не повторяют ранее опубликованный материал. Он составил эту таблицу на основе советского реферативного журнала «Математика», в котором «советские публикации составляют 1/6 мировых публикаций по прикладной статистике, реферируемых за год» [20, с. 56].

Однако, по-видимому, редакторов журнала «Вестник статистики» не убедили доводы Орлова. В дополнение к его письму они напечатали свое заявление о том, что письмо Тимофеева было опубликовано для того, чтобы показать, что сборник «Ученые записки по статистике» перестал отвечать своей цели и превратился в математический журнал и что содержание статей в «Прикладной статистике» (том 45 «Ученых записок по статистике») не отвечает названию сборника. Более того, редакторы добавили, что находят доводы Тимофеева убедительными. Выступив с критикой письма Орлова, они упрекнули его за то, что он пытается «опровергнуть содержание письма К. Тимофеева, а заодно изобразить его автора как человека, не сведущего в делах, которыми занимается А. Орлов, а с ним и ряд других математиков» (с. 57). Они продолжали утверждать, что многие леммы и теоремы, которыми оперирует Орлов и его коллеги, не используются в практической работе. В частности, они проявили упорное желание узнать, «каков экономический эффект (в миллионах рублей), который удалось извлечь из шума при помощи измеримых отображений произвольного вероятностного пространства в множество непустых компактов плоскости, снабженное метрикой Хаусдорфа» [20, с. 57]. Касаясь ссылок на работы зарубежных авторов, редакторы отметили, что из таблицы Орлова видно, что ученые действительно ссылаются на зарубежные источники, и таким образом они приходят к выводу, что их утверждение верно. Обширные политизированные тексты «редакторов», весьма враждебные, но не подписанные, демонстрируют распространенные в то время — да и сейчас — приемы борьбы, используемые врагами современной науки.

Подтверждением того, что спорные вопросы еще не решены, по крайней мере в умах читателей, явилась публикация третьего письма, написанного Н. Шереметом [96]. Шеремет, доцент Московского института инженеров железнодорожного транспорта (МИИТ), придерживалтся умеренных взглядов по вопросу об определении прикладной статистики и ее роли. В начале своего письма он отметил, что Тимофеев не ответил на свой собственный вопрос: «Что же такое прикладная статистика?» По мнению Шеремета, прикладные науки являются связующим звеном между чисто «инженерными» работами и научными исследованиями или чистой наукой. Он выступил в защиту необходимости стадии «корректировки» или «подстройки» между стадиями научных изысканий и применением научных теорий на практике. Затем он привел хорошо известное мнение Большева о том, что вся статистика является прикладной (Л.Н. Большев высказал это мнение в личной беседе с А.И. Орловым, цитата была включена в статью [20]), но не поддержал это утверждение, так как оно является слишком широким обобщением. Затем Шеремет проанализировал точку зрения, что каждая наука имеет свою собственную статистику (например, физическая статистика и биологическая статистика), но отверг ее, так как она противоречит мнению Ф. Энгельса, высказанному при подобных обстоятельствах в связи с механикой, физикой и химией. Шеремет критиковал Орлова за примеры из области экономики, т.к. эти примеры могли привести к ошибочному — по мнению Шеремета — предположению, что прикладная статистика является универсальной наукой.

Шеремет настаивал на определении статистики как общественной науки, однако признает возможность использования прикладной статистики в своей собственной области. Шеремет написал в свойственных ему неопределенных выражениях:

«Можно предположить, что предметом данной научной дисциплины являются «статистические данные»... Здесь уже не важно, от какого реального явления отвлечены данные абстрактные понятия... Математическая идеализация «статистических данных» и операций над ними дает возможность сводить известное разнообразие связей и закономерностей конкретной практической области к их определенному классу, производить необходимые расчеты» [96, с. 69].

Он заявил, что прикладная статистика пока еще не является — по его мнению — четко определенной областью, и в заключение написал, что «прикладной статистике» в большей степени присущи черты междисциплинарных исследований, чем исследований, проводимых в рамках самостоятельной дисциплины [96, с. 71].

В письме Шеремет допустил несколько неточностей, граничащих с дезинформацией. Он, кажется, не знает, что с 1973 г. журнал «Анналы статистики» (The Annals of Statistics — основной западный статистический журнал) является непосредственным продолжением журнала «Анналы математической статистики» (Annals of Mathematical Statistics) и не делает разницы между узким техническим термином «статистика» (как функция от результатов наблюдений) и термином «статистика» (как наука и методология). Ссылка на элементарный учебник Вайнберга и Шумахера 1969 г. [97] как на образцовую современную монографию по прикладной статистике в лучшем случае вызывает сомнение.

Показательным является сам факт публикации подобного письма без редакционного комментария в советском консервативном журнале по статистическим наукам — в журнале, который со времени своего возрождения в 1949 г. стал выразителем позиций официальных статистиков (многие из них строго придерживаются марксистко-ленинской ориентации), рассматривающих статистику только как описательную науку.

На страницах «Вестника статистики» письмо Шеремета было не единственным откликом на полемику между Тимофеевым и Орловым. По всей видимости, независимо от письма Шеремета в июле 1987 года «Вестник статистики» опубликовал письмо И. Манделя [98], доцента института Народного хозяйства в Алма-Ате (Казахстан). В качестве комментария на письма Тимофеева и Орлова Мандель составил развернутую схему, отражающую взаимосвязь теории статистики, прикладной статистики и математической статистики. Эта схема была представлена наряду с шестью другими методологическими приемами, чтобы показать, какое влияние оказывают теория статистики, прикладная статистика и математическая статистика на методы исследования массовых явлений. Главным в его доводах являлось положение о том, что в то время, как теория статистики в основном отражает социальную сферу массовых процессов, прикладная статистика должна быть направлена на отражение массовых явлений любого характера. Таким образом, прикладная статистика должна являться своего рода «буферной наукой», которая переводит результаты математической статистики на язык, понятный исследователям в различных областях науки и практики. Он высказал сожаление по поводу существующих расхождений во взглядах между чистыми математическими статистиками и чистыми «прикладными» и обратил внимание на многочисленные примеры неправильного использования статистической методологии. Он приветствовал усилия математиков (в СССР и за рубежом), направленные на ликвидацию разрыва между математикой и реальным миром. В заключение он посоветовал называть прикладную статистику в значении «буферной науки» «прикладной математической статистикой». На конкретный вопрос о том, является ли сборник «Ученые записки по статистике» подходящим изданием для публикации статей по прикладной (математической?) статистике, он дает категорический отрицательный ответ, полностью совпадающий с мнением Тимофеева по этому вопросу. Мандель составил таблицу, согласно которой в 4 выпусках «Записок» (1978–1985), подготовленных прикладными статистиками, опубликовано 85 статей (1 092 стр.). Из них 62 статьи (787 стр.), т.е. почти три четверти, по его мнению, по своему содержанию больше подходили для публикации в известном советском журнале «Теория вероятностей и ее применения», так как были посвящены чисто математическим результатам и написаны в виде теорем и доказательств. Мандель, увы, не знал, что к тому времени редколлегия и авторы журнала «Теория вероятностей и ее применения» уже полностью оторвались от практики анализа статистических данных. По мнению Манделя, отличительной чертой прикладной статистики является отсутствие доказательств; для нее характерны только ссылки на теоремы и обсуждение вопросов «истинно» прикладного характера.

Обсуждение было продолжено в феврале 1988 г., когда в очередном выпуске «Вестника статистики» было опубликовано письмо болгарского профессора, специалиста по статистике, В. Цонева [99]. Он предлагает коренным образом изменить терминологию, связанную со всей статистической наукой.

Перестройка в области статистики. «Перестройка» в политике отразилась и в области статистики. Это проявилось не только в публикации новых статистических данных по промышленному травматизму, алкоголизму, преступности и т.д., но также и в координации работы многочисленных учреждений, занимающихся обработкой статистических данных. Реорганизация ЦСУ явилась еще одним свидетельством озабоченности правительства недостатками в данной области. К примеру, статистические данные, связанные с производством черных металлов, собирались и обрабатывались тремя учреждениями — Госпланом, ЦСУ и Институтом экономики министерства черной металлургии. На Всесоюзной конференции статистиков в мае 1985 г. выяснилось, что данные по прокату черных металлов, поступающие из этих трех источников, «совершенно разные» [100]. В феврале 1987 г. литературно-художественный журнал «Новый мир» выступил с открытой и резкой критикой отсутствия достоверных статистических данных. Несколько статистиков, среди них — Н. Шеремет и Т. Козлов, заведующий кафедрой статистики МИИТ — выступили с резким опровержением. Дискуссии продолжаются, поскольку ситуация в XXI в. аналогична — статистические данные разных ведомств и организаций не совпадают.

Как отмечено в [24], разногласия между учеными, о которых говорилось выше, характерны не только для Советского Союза. Американские и другие западные статистики также сталкиваются с проблемой определения роли прикладной статистики и, в более широком плане, с проблемой определения статистики как науки.

Попытки объединения отечественных статистиков. В марте 1989 г. в Центральном экономико-математическом институте АН СССР состоялся Всесоюзный круглый стол «Статистика и перестройка», на котором собрались представители различных направлений в статистике — впервые в отечественной истории! Выступления были опубликованы в виде 55-го тома «Ученых записок по статистике» [101].

Высшей точкой общественного движения, ставящего целью объединение отечественных статистиков, было создание в 1990 г. Всесоюзной статистической ассоциации (ВСА), объединившей статистиков всех направлений — специалистов по прикладной и математической статистике, по надежности (в основном представителей оборонно-промышленного комплекса), преподавателей экономико-статистических дисциплин, работников официальной государственной статистики [102, 103]. Ведущую роль в создании ВСА сыграли работники Всесоюзного центра статистических методов и информатики (о его конкретной деятельности рассказано в главе 10). Наша платформа была изложена в статье [8], опубликованной, несмотря на ее весьма резкую форму, в «Вестнике статистики». Устав ВСА, решения Учредительного съезда и Пленума правления ВСА предусматривали различные формы работы [104].

Однако в 1991 г. СССР прекратил свое существование. ВСА, как и другие союзные организации, перестала действовать. И наметившееся единство статистиков распалось. Госкомстат РФ полностью «закрылся» от статистической науки, перестал даже отвечать на обращения профессиональных статистических организаций. Произошел окончательный отрыв специалистов математической статистики от практики. В настоящее время журнал «Теория вероятностей и ее применения» не представляет никакого интереса для тех, кто обрабатывает конкретные данные. Отметим, однако, выпуск энциклопедии «Вероятность и математическая статистика» [105], содержащей массу полезной информации для специалистов по статистическим методам.

Работы по прикладной статистике и другим статистическим методам продолжались в рамках Российской ассоциации статистических методов (созданной на базе одноименной секции ВСА) и Российской академии статистических методов, а также в рамках Белорусской статистической ассоциации. Основным местом публикации отечественных работ по статистическим методам является секция «Математические методы исследования» журнала «Заводская лаборатория», созданная в 1961 г. Б.В. Гнеденко и В.В. Налимовым. В ней за более чем 60 лет помещено около 1 000 статей по различным направлениям прикладной статистики, прежде всего по статистическому анализу числовых величин, статистике нечисловых данных, многомерному статистическому анализу, планированию эксперимента, опыту применения статистических методов при решении конкретных прикладных задач.

Литература к главе 15

1. *Никитина, Е.П.* Коллекция определений термина «статистика» / Е.П. Никитина, В.Д. Фрейдлина, А.В. Ярхо. — Москва : МГУ, 1972. — 46 с.

2. *Ленин, В.И.* Развитие капитализма в России. Процесс образования внутреннего рынка для крупной промышленности / В.И. Ленин. — Т. XII. — Москва : Политиздат, 1986. — 610 с.

3. *Гнеденко, Б.В.* Очерк по истории теории вероятностей / Б.В. Гнеденко. — Москва :УРСС, 2001. — 88 с.

4. *Клейн, Ф.* Лекции о развитии математики в XIX столетии / Ф. Клейн. — Ч. I. — Москва : Ленинград : Объединенное научно-техническое издательство НКТП СССР, 1937. — 432 с.

5. *Плошко, Б.Г.* История статистики : учебное пособие / Б.Г. Плошко, И.И. Елисеева. — Москва : Финансы и статистика, 1990. — 295 с.

6. *Бернштейн, С.Н.* Современное состояние теории вероятностей и ее приложений / С.Н. Бернштейн // Труды Всероссийского съезда математиков в Москве 27 апреля — 4 мая 1927 г. — Москва : Ленинград : ГИЗ, 1928. — С. 50–63.

7. *Орлов, А.И.* О современных проблемах внедрения прикладной статистики и других статистических методов / А.И. Орлов // Заводская лаборатория. — 1992. — Т. 58. — № 1. — С. 67–74.

8. *Орлов, А.И.* О перестройке статистической науки и ее применений / А.И. Орлов // Вестник статистики. — 1990. — № 1. — С. 65–71.

9. *Кендалл, М.* Теория распределений / М. Кендалл, А. Стьюарт. — Москва : Наука, 1966. — 566 с.

10. *Кендалл, М.* Статистические выводы и связи / М. Кендалл, А. Стьюарт. — Москва : Наука, 1973. — 899 с.

11. *Кендалл, М.* Многомерный статистический анализ и временные ряды / М. Кендалл, А. Стьюарт. — Москва : Наука, 1976. — 736 с.

12. *Налимов, В.В.* Наукометрия. Изучение развития науки как информационного процесса / В.В. Налимов, З.М. Мульченко. — Москва : Наука, 1969. — 192 с.

13. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения = Applied statistics. Regulations for determinations of estimates and confidence limits for parameters of gamma distribution : государственный стандарт Союза ССР : издание официальное : утвержден Постановлением Государственного комитета СССР по стандартам от 27 июня 1983 г. № 2684 : введен впервые : дата введения 1 января 1985 г. — Москва : Изд-во стандартов, 1984. — 53 с. (В настоящее время отменен как нормативный документ, но может использоваться как научная публикация.)

14. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1965. — 1-е изд.; 1968. — 2-е изд.; 1983. — 3-е изд.

15. *Смирнов, Н.В.* О приближении плотностей распределения случайных величин / Н.В. Смирнов // Ученые записки МГПИ им. В.П. Потемкина. — 1951. — Т. XVI. — Вып. 3. — С. 69–96.

16. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. —3-е изд., испр. и доп. — Москва : Экзамен, 2004. — 576 с.

17. *Орлов, А.И.* О развитии прикладной статистики / А.И. Орлов // Современные проблемы кибернетики (прикладная статистика). — Москва : Знание, 1981. — С. 3–14.

18. *Тутубалин, В.Н.* Границы применимости (вероятностно-статистические методы и их возможности) / В.Н. Тутубалин. — Москва : Знание, 1977. — 64 с.

19. *Орлов, А.И.* Сертификация и статистические методы / А.И. Орлов // Заводская лаборатория, 1997. — Т. 63. — № 3. — С. 55–62.

20. *Орлов, А.И.* Что дает прикладная статистика народному хозяйству? / А.И. Орлов // Вестник статистики. — 1986. — № 8. — С. 52–56.

21. *Орлов, А.И.* Применение эконометрических методов при решении задач контроллинга / А.И. Орлов, Л.А. Орлова // Контроллинг. — 2003. — № 4 (8). — С. 50–54.

22. *Комаров, Д.М.* Роль методологических исследований в разработке методоориентированных экспертных систем (на примере оптимизационных и статистических методов / Д.М. Комаров, А.И. Орлов // Вопросы применения экспертных систем. — Минск: Центросистем, 1988. — С. 151–160.

23. The teaching of statistics // Studies in mathematical education. — Vol. 7. — 1991. — 258 p.

24. *Kotz, S.* The Hausdorff Space and Applied Statistics: A View from USSR // S. Kotz, K. Smith // The American Statistician. —1988. — November. — Vol. 42. — № 4. — P. 241–244.

25. *Кудлаев, Э.М.* Вероятностно-статистические методы исследования в работах А.Н. Колмогорова / Э.М. Кудлаев, А.И. Орлов // Заводская лаборатория. — 2003. — Т. 69. — № 5. — С. 55–61.

26. *Корнев, В.П.* Видные деятели отечественной статистики. 1686–1990. Биографический словарь / В.П. Корнев. — Москва : Финансы и статистика, 1993. — 200 с.

27. *Чехов, А.П.* Остров Сахалин : сочинения / А.П. Чехов. — Т. 14–15. — Москва : Наука, 1978. — 928 с.

28. Колмогоров в воспоминаниях / под редакцией А.Н. Ширяева. — Москва : Физматлит, 1993. — 736 с.

29. *Колмогоров, А.Н.* Избранные труды: Математика и механика / А.Н. Колмогоров. — Москва : Наука, 1985. — 470 с.

30. *Колмогоров, А.Н.* Теория вероятностей и математическая статистика / А.Н. Колмогоров. — Москва : Наука, 1986. — 535 с.

31. *Колмогоров, А.Н.* Теория информации и теория алгоритмов / А.Н. Колмогоров. — Москва : Наука, 1987. — 304 с.

32. *Колмогоров, А.Н.* Основные понятия теории вероятностей / А.Н. Колмогоров. — Москва : Ленинград : ОНТИ, 1936. — 80 с. — 3-е изд. — Москва : Фазис, 1998. — 144 с.

33. *Орлов, А.И.* Прикладная статистика : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 672 с.

34. *Тюрин, Ю.Н.* Линейная модель в многомерной непараметрической статистике / Ю.Н. Тюрин // Многомерный статистический анализ в социально-экономических исследованиях. — Москва : Наука, 1974. — С. 7–24.

35. *Орлов, А.И.* О критериях Колмогорова и Смирнова / А.И. Орлов // Заводская лаборатория. — 1995. — Т. 61. — № 7. — С. 59–61.

36. *Орлов, А.И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат / А.И. Орлов // Заводская лаборатория». — 1985. — Т. 51. — № 1. — С. 60–62.

37. *Тюрин, Ю.Н.* Исследования по непараметрической статистике (непараметрические методы и линейная модель) : специальность 01.01.05 «Теория вероятностей и математическая статистика» : автореферат на соискание ученой степени доктора физико-математических наук / Юрий Николаевич Тюрин ; Московский государственный университет им. Н.Г. Ломоносова. — Москва, 1985. — 33 с.

38. *Орлов, А.И.* Методы поиска наиболее информативных множеств признаков в регрессионном анализе / А.И. Орлов // Заводская лаборатория. — 1995. — Т. 61. — № 1. — С. 56–58.

39. *Сердобольский, В.И.* Статистический анализ при большом числе параметров / В.И. Сердобольский, А.И. Орлов // Программно-алгоритмическое обеспечение прикладного многомерного статистического анализа : тезисы докладов III Всесоюзной школы-семинара. — Москва : ЦЭМИ АН СССР, 1987. — С. 151–160.

40. *Феллер, В.* Введение в теорию вероятностей и ее приложения : в 2 томах / В. Феллер ; перевод с английского. — Т. 1.— Москва : Мир, 1984. — 528 с.

41. *Гнеденко, Б.В.* Курс теории вероятностей : учебник / Б.В. Гнеденко. — 7-е изд., испр. — Москва : Эдиториал УРСС, 2001. — 320 с.

42. *Ширяев, А.Н.* Статистический последовательный анализ / А.Н. Ширяев. — Москва : Наука, 1976. — 240 с.

43. *Круглов, В.М.* Предельные теоремы для случайных сумм / В.М. Круглов, В.Ю. Королев. — Москва : МГУ, 1990. — 188 с.

44. *Gnedenko, B.V.* Random summation: limit theorems and applications / B.V. Gnedenko, V.Yu. Korolev. — Boca Raton, Fl. : CRC Press, 1996. — 268 p.

45. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

46. *Колмогоров, А.Н.* Статистический приемочный контроль при допустимом числе дефектных изделий, равном нулю / А.Н. Колмогоров. — Ленинград : Знание, 1951. — 24 с.

47. *Гнеденко, Б.В.* Математика и контроль качества продукции / Б.В. гнеденко. — Москва : Знание, 1978. — 64 с.

48. *Беляев, Ю.К.* Вероятностные методы выборочного контроля / Ю.К. Беляев. — Москва : Наука, 1975. — 407 с.

49. *Лумельский, Я.П.* Статистические оценки результатов контроля качества / Я.П. Лумельский. — Москва : Изд-во стандартов, 1979. — 200 с.

50. *Kac, M.* On test of normality and other tests of goodness of fit based on distance methods / M. Kac, J. Kiefer, J. Wolfowitz // Ann. Math. Statist. — 1955. — V. 26. — № 2. — P.189–211.

51. *Рунион, Р.* Справочник по непараметрической статистике / Р. Рунион. — Москва : Финансы и статистика, 1982. — 198 с.

52. *Ивченко, Г.И.* Математическая статистика / Г.И. Ивченко, Ю.И. Медведев. — Москва : Высшая школа, 1984. — 248 с.

53. *Лумельский, Я.П.* К вопросу сравнения несмещенных и других оценок / Я.П. Лумельский // Прикладная статистика. — Москва : Наука, 1983. — С. 316–319.

54. *Журбенко, И.Г.* О выявлении эффекта воздействия в рандомизированных экспериментах / И.Г. Журбенко, Э.М. Кудлаев // Успехи математических наук. — 1984. — Т. 39. — Вып. 1. — С. 3–38.

55. *Булинский, А.В.* Линейные выборочные оценки сумм / А.В. Булинский, А.Н. Колмогоров // Теория вероятностей и ее применения. — 1979. — Т. 24. — № 2. — С. 241–251.

56. *Гнеденко, Б.В.* Предельные распределения для сумм независимых случайных величин / Б.В. Гнеденко, А.Н. Колмогоров. — Москва : Ленинград : ГТТИ, 1949. — 264 с.

57. *Гнеденко, Б.В.* Элементы программирования (2-е изд.) / Б.В. Гнеденко, В.С. Королюк, Е.Л. Ющенко. — Москва : Физматгиз, 1963. — 348 с.

58. *Добровольская, Н.К.* Борис Владимирович Гнеденко / Н.К. Добровольская // Киевские математики-педагоги / под редакцией А.Н. Боголюбова. — Киев : Вища школа, 1979. — С. 37–60.

59. *Гнеденко, Б.В.* Предельные теоремы для максимального члена вариационного ряда / Б.В. Гнеденко // Доклады АН СССР. — 1941. — Т. 32. — С. 7–9.

60. *Кудлаев, Э.М.* Оценивание параметров распределения Вейбулла — Гнеденко / Э.М. Кудлаев // Техническая кибернетика. — 1986. — № 6. — С. 5–18.

61. *Гнеденко, Б.В.* О случайных величинах, обусловленных суммами независимых случайных величин / Б.В. Гнеденко, Э.М. Кудлаев // Вестник МГУ им. М.В. Ломоносова. Сер. мат. и мех. — 1995. — Вып. 1. — С. 23–31.

61. *Боев, Г.П.* Методика составления эмпирических зависимостей и номограмм в текстильном деле / Г.П. Боев, Ю.К. Виноградов, Б.В. Гнеденко. — Москва : Гизлегпром, 1936. — 128 с.

62. *Гнеденко, Б.В.* К теории счетчиков Гейгер-Мюллера / Б.В. Гнеденко // Журнал экспериментальной и теоретической физики. — 1941. — Т. 11. — Вып. 1. — С. 101–106.

63. *Гнеденко, Б.В.* Введение в теорию массового обслуживания / Б.В. Гнеденко, И.Н. Коваленко. — Москва : Наука, 1966. — 301 с.

64. *Гнеденко, Б.В.* Элементарное введение в теорию массового обслуживания / Б.В. Гнеденко, А.Я. Хинчин. — 6-е изд. — Москва : Наука, 1964. — 146 с.

65. Приоритетные системы обслуживания // Б.В. Гнеденко, Э.А. Даниелян, Б.Н. Димитров и др. — Москва : Изд-во МГУ, 1973. — 447 с.

66. *Гнеденко, Б.В.* Каждому специалисту нужно знать математическую статистику / Б.В. Гнеденко // Вестник высшей школы. — 1961. — № 12. — С. 29–30.

67. *Колмогоров, А.Н.* О работах Б.В. Гнеденко по теории вероятностей / А.Н. Колмогоров // Теория вероятностей и ее применения. — 1962. — Т. 7. — № 2. — С. 323–329.

68. *Гнеденко, Б.В.* Математические методы в теории надежности / Б.В. Гнеденко, Ю.К. Беляев, А.Д. Соловьев. — Москва : Наука, 1965. — 524 с.

69. *Гнеденко, Б.В.* Введение в специальность математика / Б.В. Гнеденко. — Москва : Наука, 1991. — 340 с.

70. *Гнеденко, Б.В.* Очерки истории математики в России / Б.В. Гнеденко. — Москва : ГТТИ, 1946. — 247 с.

71. *Гнеденко, Б.В.* О некоторых задачах истории математики / Б.В. Гнеденко // Труды третьего Всесоюзного математического съезда. (Москва, июнь-июль 1956). — Т.II. Краткое содержание обзорных и секционных докладов. — Москва : Изд-во АН СССР, 1956. — С. 100–101.

72. *Гнеденко, Б.В.* Элементарное введение в теорию вероятностей / Б.В. Гнеденко, А.Я. Хинчин. — Москва : ГТТИ, 1946. — 128 с.

73. *Гнеденко, Б.В.* О математике / Б.В. Гнеденко. — Москва : Эдиториал УРСС. 2000. — 208 с.

74. *Гнеденко, Б.В.* Роль математических методов исследования в кардинальном ускорении научно-технического прогресса / Б.В. Гнеденко, А.И. Орлов // Заводская лаборатория. — 1988. — Т. 54. — № 1. — С. 1–4.

75. *Гнеденко, Б.В.* Математическая статистика — мощное орудие в работе заводской лаборатории / Б.В. Гнеденко // Заводская лаборатория. — 1961. — Т. 27. — № 10. — С. 1251–1253.

76. *Смирнов, Н.В.* Теория вероятностей и математическая статистика. Избранные труды. — Москва : Наука, 1970. — 289 с.

77. *Смирнов, Н.В.* Курс теории вероятностей и математической статистики для технических приложений / Н.В. Смирнов, И.В. Дунин-Барковский. — 3-е изд., стер. — Москва : Наука, 1969. — 512 с.

78. *Смирнов, Н.В.* Теория вероятностей и математическая статистика в приложении к геодезии / Н.В. Смирнов, Д.А. Белугин. — Москва : Недра, 1969. — 380 с.

79. *Орлов, А.И.* Скорость сходимости распределения статистики Мизеса — Смирнова / А.И. Орлов // Теория вероятностей и ее применения. — 1974. — Т. 19. — № 4. — С. 766–786.

80. *Большев, Л.Н.* Избранные труды. Теория вероятностей и математическая статистика / Л.Н. Большев. — Москва : Наука, 1987. — 286 с.

81. *Налимов, В.В.* Канатоходец. Воспоминания / В.В. Налимов. — Москва : Прогресс, 1994. — 456 с.

82. *Налимов, В.В.* Применение математической статистики при анализе вещества / В.В. Налимов. — Москва : Физматгиз, 1960. — 430 с.

83. *Налимов, В.В.* Статистические методы планирования экстремальных экспериментов / В.В. Налимов, Н.К. Чернова. — Москва : Физматгиз, 1965. — 340 с.

84. *Налимов, В.В.* Теория эксперимента / В.В. Налимов. — Москва : Наука, 1971. — 207 с.

85. *Налимов, В.В.* Логические основания планирования эксперимента / В.В. Налимов, Т.И. Голикова. —2-е изд., перераб. и расшир. — Москва : Металлургия, 1981. — 151 с.

86. *Маркова, Е.В.* Математическая теория эксперимента: история, развитие, будущее / Е.В. Маркова, Е.П. Никитина // Заводская лаборатория. — 2002. — Т. 68. — № 1. — С. 112–118.

87. Математическая теория планирования эксперимента / под редакцией С.М. Ермакова. — Москва : Наука, 1983. — 392 с.

88. *Налимов, В.В.* Вероятностная модель языка / В.В. Налимов. —2-е изд., расширенное. — Москва : Наука, 1979. — 303 с.

89. *Орлов, А.И.* Математика нечеткости / А.И. Орлов // Наука и жизнь. — 1982. — № 7. — С. 60–67.

90. *Налимов, В.В.* Спонтанность сознания: Вероятностная теория смыслов и смысловая архитектоника личности / В.В. Налимов. — Москва : Прометей, 1989. — 288 с.

91. *Налимов, В.В.* В поисках иных смыслов / В.В. Налимов. — Москва : Прогресс, 1993. — 280 с.

92. *Kotz, S.* Statistical Terminology — Russian Vs. English — in the Light of the Development of Statistics in the USSR / S. Kotz // The American Statistician. — 1965. — Vol. 19. — № 3. — P. 16–22.

93. *Kotz, S.* Statistics in the USSR / S. Kotz // Survey. — 1965. — Vol. 57. — October. — P. 132–141.

94. *Тимофеев, К.* Что же такое прикладная статистика? / К. Тимофеев // Вестник статистики. — 1985. — № 10. — С. 66–67.

95. *Каган, А.М.* Характеризационные задачи математической статистики / А.М. Каган, Ю.В. Линник, С.Р. Рао. — Москва : Наука, 1972. — 656 с.

96. *Шеремет, Н.* О так называемой прикладной статистике / Н. Шеремет // Вестник статистики. — 1987. — № 2. — С. 67–71.

97. *Weinberg, J.H.* Statistics: An Intuitive Approach / J.H. Weinberg, J. Schumaker. — 2-nd ed. — Belmont, CA : Brooks-Cole, 1969.

98. *Мандель, И.* Теория статистики и прикладная статистика / И. Мандель // Вестник статистики. — 1987. — № 7. — С. 76–79.

99. *Цонев, В.* К дискуссии по вопросу: что же такое прикладная статистика / В. Цонев // Вестник статистики. — 1988. — № 2. — С. 67–68.

100. *Маркович, М.* Хроника и информация / М. Маркович // Вестник статистики. — 1986. — № 11. — С. 62–64.

101. Статистика и перестройка: Ученые записки по статистике. — Т. 55. — Москва : Наука, 1991. — 280 с.

102. *Орлов, А.И.* Создана единая статистическая ассоциация / А.И. Орлов // Вестник Академии наук СССР. — 1991. — № 7. — С. 152–153.

103. *Орлов, А.И.* Всесоюзная статистическая ассоциация — гарантия успешного внедрения современных статистических методов / А.И. Орлов // Надежность и контроль качества. — 1991. — № 6. — С. 54–55.

104. Устав Всесоюзной статистической ассоциации (ВСА). 1-й Пленум Правления ВСА // Вестник статистики. — 1991. — № 2. — С. 71–76.

105. Вероятность и математическая статистика : энциклопедия / главный редактор Ю.В. Прохоров. — Москва : Большая Российская Энциклопедия, 1999. — 910 с.

106. *Орлов, А.И.* Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат и ошибки при их применении / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 97. — С. 32–45.

107. *Орлов, А.И.* Основные этапы становления статистических методов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 97. — С. 73–85.

108. *Орлов, А.И.* Распределения реальных статистических данных не являются нормальными / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 117. — С. 71–90.

109. *Орлов, А.И.* Вероятностно-статистические методы в работах А.Н. Колмогорова / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 98. — С. 96–104.

110. *Орлов, А.И.* Вероятностно-статистические методы в работах Б.В. Гнеденко / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 100. — С. 1–30.

111. *Орлов, А.И.* Непараметрическая и прикладная статистика в нашей стране / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 101. — С. 197–226.

112. *Орлов, А.И.* Современное состояние непараметрической статистики / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 106. — С. 239–269.

113. *Орлов, А.И.* Состояние и перспективы развития прикладной и теоретической статистики / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 115. — С. 202–226.

114. *Орлов, А.И.* Прикладная статистика — состояние и перспективы / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 119. — С. 44–74.

115. *Орлов, А.И.* Статистика нечисловых данных — центральная часть современной прикладной статистики / А.И. Орлов // Научный журнал КубГАУ. — 2020. — № 156. — С. 111–142.

116. *Орлов, А.И.* О развитии статистики объектов нечисловой природы / А.И. Орлов // Научный журнал КубГАУ. — 2013. — № 93. — С. 41–50.

Контрольные вопросы

1. Расскажите о первых статистических исследованиях.

2. Как менялось представление о предмете статистики с течением времени?

3. Что такое непараметрическая статистика и чем она отличается от параметрической?

4. Расскажите о вкладе А.Н. Колмогорова в статистические методы.

5. Расскажите о вкладе Б.В. Гнеденко в статистические методы.

6. Расскажите о работах по внедрению статистических методов в 1960–1970-е гг., организованных в В.В. Налимовым.

7. Какие проблемы рассматривались в ходе дискуссии по прикладной статистике, прошедшей в 1980-е гг. на страницах журнала «Вестник статистики»?

8. Каково положение статистических методов как научно-практической области в начале XXI в.?

Темы докладов, рефератов, исследовательских работ

1. Статистические работы, описанные в Ветхом Завете.

2. Статистические методы в XIX в.

3. Переписи на Руси, их цели и методы организации.

4. Развитие теории вероятностей и статистические методы в дореволюционной России.

5. Работы члена-корреспондента АН СССР Н.В. Смирнова в области непараметрической статистики.

6. Почему некорректно говорить о «критерии Колмогорова — Смирнова»?

7. Научная школа Московского энергетического института в области статистических методов.

8. Работы отечественных ученых в области статистических методов управления качеством.

9. История развития работ по математическим методам планирования эксперимента в нашей стране.

10. История межфакультетской Лаборатории статистических методов МГУ им. Н.В. Ломоносова.

11. История создания Всесоюзной статистической ассоциации.

ГЛАВА 16. СОВРЕМЕННЫЕ СТАТИСТИЧЕСКИЕ МЕТОДЫ

Подведем итоги и наметим перспективы развития прикладной статистики и других статистических методов. Обсудим тенденции развития статистических методов, выделим пять основных «точек роста». В связи с внедрением современных статистических методов обоснуем полезность понятия «высокие статистические технологии». Рассмотрим технологии использования компьютеров в прикладной статистике и в других статистических методах. Обсудим основные нерешенные проблемы в области статистических методов.

16.1. Точки роста

Отечественная литература по прикладной статистике и другим статистическим методам столь же необозрима, как и мировая. Только в секции «Математические методы исследования» журнала «Заводская лаборатория» с 1960-х гг. опубликовано более 1 000 статей. Не будем даже пытаться перечислять коллективы исследователей или основные монографии в этой области. Отметим только одно издание. По нашему мнению, наилучшей отечественной книгой XX века по прикладной статистике является сборник статистических таблиц Л.Н. Большева и Н.В. Смирнова [1] с подробными комментариями, играющими роль сжатого учебника и справочника.

Выделим и обсудим «точки роста» прикладной статистики и других статистических методов, те их направления, которые представляются перспективными в будущем, в следующие десятилетия XXI века, но пока в большинстве учебных изданий отодвинуты на задний план традиционными постановками.

При описании современного этапа развития статистических методов целесообразно выделить пять актуальных направлений, в которых развивается современная прикладная статистика, т.е. пять «точек роста»: непараметрика (т.е. непараметрическая статистика), робастность, бутстреп, статистика интервальных данных, статистика нечисловых данных (в несколько иной терминологии — статистика объектов нечисловой природы). Обсудим их.

Непараметрическая статистика. В первой трети XX в., одновременно с параметрической статистикой, в работах Спирмена и Кендалла появились первые непараметрические методы, основанные на коэффициентах ранговой корреляции, носящих ныне имена этих статистиков. Но непараметрика, не делающая нереалистических предположений о том, что функции распределения результатов наблюдений принадлежат тем или иным параметрическим семействам распределений, стала заметной частью статистики лишь со второй трети XX века. В 1930-е гг. появились работы А.Н. Колмогорова и Н.В. Смирнова, предложивших и изучивших статистические критерии, носящие в настоящее время их имена. Эти критерии основаны на использовании так называемого эмпирического процесса. (Как известно, эмпирический процесс — это разность между эмпирической и теоретической функциями распределения, умноженная на квадратный корень из объема выборки.) В работе А.Н. Колмогорова 1933 г. изучено предельное распределение супремума модуля эмпирического процесса, называемого сейчас критерием Колмогорова. Затем Н.В. Смирнов исследовал супремум и инфимум эмпирического процесса, а также интеграл (по теоретической функции распределения) квадрата эмпирического процесса.

Следует отметить, что встречающееся иногда в литературе словосочетание «критерий Колмогорова-Смирнова» некорректно, поскольку эти два статистика никогда не печатались вместе и не изучали один и тот же критерий схожими методами. Корректно сочетание «критерий типа Колмогорова-Смирнова», применяемое для обозначения критериев, основанных на использовании супремума функций от эмпирического процесса [2].

После второй мировой войны развитие непараметрической статистики пошло быстрыми темпами. Большую роль сыграли работы американского статистика Ф. Вилкоксона и его школы. К настоящему времени с помощью непараметрических методов можно решать практически тот же круг статистических задач, что и с помощью параметрических. Однако для обеспечения широкого внедрения непараметрических методов необходимо провести еще целый комплекс теоретических и пилотных (т.е. пробных) прикладных работ. Все большую роль играют непараметрические оценки плотности, непараметрические методы регрессии и распознавания образов (дискриминантного анализа). В нашей стране непараметрические методы получили достаточно большую известность после выхода в 1965 г. первого издания упомянутого выше сборника статистических таблиц Л.Н. Большева и Н.В. Смирнова [1], содержащего подробные таблицы для основных непараметрических критериев.

Тем не менее параметрические методы все еще популярнее непараметрических, особенно среди тех прикладников, кто слабо знаком со статистическими методами. Неоднократно публиковались экспериментальные данные, свидетельствующие о том, что распределения реально наблюдаемых случайных величин, в частности, ошибок измерения, в подавляющем большинстве случаев отличны от нормальных (гауссовских). Тем не менее, математики-теоретики продолжают строить и изучать статистические модели, основанные на гауссовости, а практики — применять подобные методы и модели. Другими словами, «ищут под фонарем, а не там, где потеряли».

Устойчивость статистических процедур (робастность). Если в параметрических постановках на вероятностные модели статистических данных накладываются слишком жесткие требования — их функции распределения должны принадлежать определенному параметрическому семейству, то в непараметрических, наоборот, излишне слабые — требуется лишь, чтобы функции распределения были непрерывны. При этом игнорируется априорная информация о том, каков «примерный вид» распределения. Априори можно ожидать, что учет этого «примерного вида» улучшит показатели качества статистических процедур. Развитием этой идеи является теория устойчивости (робастности) статистических процедур, в которой предполагается, что распределение исходных данных мало отличается от некоторого параметрического семейства. За рубежом эту теорию разрабатывали П. Хубер, Ф. Хампель и многие другие. Из монографий на русском языке, трактующих о робастности и устойчивости статистических процедур, самой ранней и наиболее общей была книга [3], следующей — монография [4]. Частными случаями реализации идеи робастности (устойчивости) статистических процедур являются статистика объектов нечисловой природы и статистика интервальных данных.

Имеется большое разнообразие моделей робастности в зависимости от того, какие именно отклонения от заданного параметрического семейства допускаются. Среди теоретиков наиболее популярной оказалась модель выбросов, в которой исходная выборка «засоряется» малым числом «выбросов», имеющих принципиально иное распределение. Однако эта модель представляется «тупиковой», поскольку в большинстве случаев большие выбросы либо невозможны из-за ограниченности шкалы прибора либо интервала изменения измеряемой величины, либо от них можно избавиться, применяя лишь статистики, построенные по центральной части вариационного ряда. Кроме того, в подобных моделях обычно считается известной частота засорения, что в сочетании со сказанным выше делает их малопригодными для практического использования.

Более перспективным представляется, например, модель малых отклонений распределений, в которой расстояние между распределением каждого элемента выборки и базовым распределением не превосходит заданной малой величины, и модель статистики интервальных данных.

Бутстреп (размножение выборок). Другое из упомянутых выше направлений — бутстреп — связано с интенсивным использованием возможностей компьютеров. Основная идея состоит в том, чтобы теоретическое исследование заменить вычислительным экспериментом. Например, вместо описания выборки распределением из параметрического семейства строим большое число «похожих» выборок, т.е. «размножаем» выборку. Затем вместо оценивания характеристик (и параметров) и проверки гипотез на основе свойств теоретического распределения решаем эти задачи вычислительным методом, рассчитывая интересующие нас статистики по каждой из «похожих» выборок и анализируя полученные при этом распределения. Например, вместо того, чтобы теоретическим путем находить распределение статистики, доверительные интервалы и другие характеристики, моделируют большое число выборок, похожих на исходную, затем рассчитывают соответствующие значения интересующей исследователя статистики и изучают их эмпирическое распределение. Квантили этого распределения задают доверительные интервалы, и т.д.

Термин «бутстреп» мгновенно получил широкую известность после первой же статьи Б. Эфрона 1979 г. по этой тематике. Он сразу же стал обсуждаться в массе публикаций, в том числе и научно-популярных. В «Заводской лаборатории» № 10 за 1987 г. была помещена подборка статей по бутстрепу. На русском языке выпущен сборник статей Б. Эфрона [5]. Основная идея бутстрепа по Б. Эфрону состоит в том, что методом Монте-Карло (статистических испытаний) многократно извлекаются выборки из эмпирического распределения. Эти выборки, естественно, являются вариантами исходной, напоминают ее.

Сама по себе идея «размножения выборок» была известна гораздо раньше. Одна из статей Б. Эфрона в сборнике [5] называется так: «Бутстреп-методы: новый взгляд на метод складного ножа». Упомянутый «метод складного ножа» (*jackknife*) предложен М. Кенуем еще в 1949 г., за 30 лет до появления статьи Б. Эфрона. «Размножение выборок» при этом осуществляется путем исключения одного наблюдения. Таким путем для выборки объема n получаем n «похожих» на нее выборок объема $(n - 1)$ каждая. Если же исключать по 2 наблюдения, то число «похожих» выборок возрастает до $n(n - 1)/2$ объема $(n - 2)$ каждая.

Преимущества и недостатки бутстрепа как статистического метода в сравнении с рядом аналогичных методов обсуждаются ниже. Необходимо подчеркнуть, что бутстреп по Эфрону — лишь один из вариантов методов «размножения выборки» (*resampling*), и, на наш взгляд, не самый удачный. Метод «складного ножа» представляется более полезным. На его основе можно сформулировать следующую простую практическую рекомендацию.

Предположим, что Вы по выборке делаете какие-либо статистические выводы. Вы хотите узнать также, насколько эти выводы устойчивы. Если у Вас есть другие (контрольные) выборки, описывающие то же явление, то Вы можете применить к ним ту же статистическую процедуру и сравнить результаты. А если таких выборок нет? Тогда Вы можете их построить искусственно. Берете исходную выборку и исключаете один элемент. Получаете похожую выборку (она взята из того же распределения, только объем на единицу меньше). Затем возвращаете этот элемент выборки и исключаете другой. Получаете вторую похожую выборку. Поступая таким образом со всеми элементами исходной выборки, получаете столько выборок, похожих на исходную, каков ее объем. Остается обработать их тем же способом, что и исходную, и изучить устойчивость получаемых выводов — разброс оценок параметров, частоты принятия или отклонения гипотез и т.д.

Можно изменять не выборку, а сами данные. Поскольку всегда имеются погрешности измерения, то реальные данные — это не числа, а интервалы (результат измерения плюс-минус погрешность). Нужна статистическая теория анализа таких данных.

Статистика интервальных данных. Перспективное и быстро развивающееся направление последних лет — статистика интервальных данных. Речь идет о развитии методов прикладной математической статистики в ситуации, когда статистические данные — не числа, а интервалы, в частности, порожденные наложением ошибок измерения на значения случайных величин.

Статистика интервальных данных идейно связана с интервальной математикой, в которой в роли чисел выступают интервалы. Это направление математики является дальнейшим развитием известных правил приближенных вычислений, посвященных выражению погрешностей суммы, разности, произведения, частного через погрешности тех чисел, над которыми осуществляются перечисленные операции. К настоящему времени удалось решить, в частности, ряд задач теории интервальных дифференциальных уравнений, в которых коэффициенты, начальные условия и решения описываются с помощью интервалов.

Одна из ведущих научных школ в области статистики интервальных данных — это школа проф. А.П. Вощинина, активно работающая с конца 1970-х гг. В частности, ее представителями изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности.

Рассмотрим другое направление в статистике интервальных данных, которое также представляется перспективным. В нем развиваются асимптотические методы статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом — уменьшаются до нуля погрешности. В частности, с помощью такой асимптотики в начале 1980-х годов были сформулированы правила выбора метода оценивания параметров гамма-распределения в ГОСТ 11.011-83 (в настоящее время отменен как нормативный документ, но может использоваться как научная публикация) [6].

В рамках рассматриваемого научного направления разработана общая схема исследования, включающая расчет нотны (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и рационального объема выборки (превышение которого не дает существенного повышения точности оценивания). Она применена к оцениванию математического ожидания, дисперсии, коэффициента вариации, параметров гамма-распределения и характеристик аддитивных статистик, при проверке гипотез о параметрах нормального распределения, в том числе с помощью критерия Стьюдента, а также гипотезы однородности с помощью критерия Смирнова. Разработаны подходы к рассмотрению интервальных данных в основных постановках регрессионного, дискриминантного и кластерного анализов. В частности, изучено влияние погрешностей измерений и наблюдений на свойства алгоритмов регрессионного анализа, разработаны способы расчета нотн и рациональных объемов выборок, введены и исследованы новые понятия многомерных и асимптотических нотн, доказаны соответствующие предельные теоремы. Начата разработка интервального дискриминантного анализа, в частности, рассмотрено влияние интервальности данных на введенный в главе 3.2 показатель качества классификации. Изучено асимптотическое поведение оценок метода моментов и оценок максимального правдоподобия (а также более общих оценок минимального контраста), проведено асимптотическое сравнение этих методов в случае интервальных данных. Найдены общие условия, при которых, в отличие от классической математической статистики, метод моментов дает более точные оценки, чем метод максимального правдоподобия (см. раздел 3.5).

В области асимптотической статистики интервальных данных российская наука имеет мировой приоритет. Во все виды статистического программного обеспечения включают алгоритмы интервальной статистики, «параллельные» обычно используемым алгоритмам прикладной математической статистики. Это позволяет в явном виде учесть наличие погрешностей у результатов наблюдений.

Статистика объектов нечисловой природы как часть прикладной статистики. Напомним, что согласно общепринятой в настоящее время классификации статистических методов прикладная статистика делится на следующие четыре области:

- статистика (числовых) случайных величин;
- многомерный статистический анализ;
- статистика временных рядов и случайных процессов;
- статистика объектов нечисловой природы.

Первые три из этих областей являются классическими. Они были хорошо известны еще в первой половине XX в. Остановимся на четвертой, сравнительно недавно вошедшей в массовое сознание специалистов. Ее именуют также статистикой нечисловых данных или попросту нечисловой статистикой. Анализ динамики развития прикладной статистики приводит к выводу, что в XXI в. она станет центральной областью прикладной статистики, поскольку содержит наиболее общие подходы и результаты.

Исходный объект в прикладной математической статистике — это выборка. В вероятностной теории статистики выборка — это совокупность независимых одинаково распределенных случайных элементов. Какова природа этих элементов? В классической математической статистике элементы выборки — это числа. В многомерном статистическом анализе — вектора. А в нечисловой статистике элементы выборки — это объекты нечисловой природы, которые нельзя складывать и умножать на числа. Другими словами, объекты нечисловой природы лежат в пространствах, не имеющих векторной структуры. Примерами объектов нечисловой природы являются:

- значения качественных признаков, т.е. результаты кодировки объектов с помощью заданного перечня категорий (градаций);
- упорядочения (ранжировки) образцов продукции (при оценке ее технического уровня и конкурентоспособности)) или заявок на проведение научных работ (при проведении конкурсов на выделение грантов), описывающие мнения экспертов;
- классификации, т.е. разбиения совокупности объектов на группы сходных между собой (кластеры);
- толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходство тематики научных работ, которое оценивается экспертами с целью рационального формирования экспертных советов внутри определенной области науки;

- результаты парных сравнений или контроля качества продукции по альтернативному признаку («годен» — «брак»), т.е. последовательности из 0 и 1;

- множества (обычные или нечеткие), например, зоны, пораженные коррозией; топокарты, полученные при кинетокардиографии; перечни возможных причин аварии, составленные экспертами независимо друг от друга; нечеткие экспертные оценки качества газовых плит;

- слова, предложения, тексты;

- векторы, координаты которых — совокупность значений разнотипных признаков, например, результат составления статистического отчета о научно-технической деятельности (т.н. форма № 1-наука) или заполненная компьютеризированная история болезни, в которой часть признаков носит качественный характер, а часть — количественный;

- ответы на вопросы экспертной, маркетинговой или социологической анкеты, часть из которых носит количественный характер (возможно, интервальный), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты; и т.д.

Интервальные данные также можно рассматривать как пример объектов нечисловой природы, а именно, как частный случай нечетких множеств.

С начала 1970-х гг. под влиянием запросов прикладных исследований в социально-экономических, технических, медицинских науках в России активно развивается статистика объектов нечисловой природы, известная также как статистика нечисловых данных или нечисловая статистика. В создании этой сравнительно новой области эконометрики и прикладной математической статистики приоритет принадлежит российским ученым.

Большую роль сыграл основанный в 1973 г. научный семинар «Экспертные оценки и анализ данных». В 1960-е гг. советское научное сообщество стало интересоваться методами экспертных оценок (об их истории и современном состоянии см. [7, 8]). Как следствие, началось знакомство с конкретными математизированными теориями, связанными с этими методами. Речь идет о репрезентативной теории измерений, ставшей известной в нашей стране по статье П. Суппеса и Дж. Зинеса в сборнике [9] и книге И. Пфанцагля [10], о теории нечеткости, современный этап которой начался с работ Л.А. Заде [11], теории парных сравнений, описанной в монографии Г. Дэвида [12]. К этому кругу идей примыкают теория случайных множеств (см., например, книгу Ж. Матерона [13]) и методы многомерного шкалирования (описаны, в частности, в монографиях А.Ю. Терехиной [14] и В.Т. Перекреста [15]). Но наибольшее влияние оказали идеи американского исследователя проф. Дж. Кемени, который аксиоматически ввел расстояние между ранжировками (теперь оно именуется в литературе расстоянием Кемени) и предложил использовать в качестве средней величины решение оптимизационной задачи (теперь — медиана Кемени). Его скромная книжка [16], написанная в соавторстве с Дж. Снеллом, породила большой поток исследований.

В течение 1970-х годов на основе запросов теории экспертных оценок (а также социологии, экономики, техники и медицины) развивались конкретные направления статистики объектов нечисловой природы. Были установлены связи между конкретными видами таких объектов, разработаны для них вероятностные модели. Научные итоги этого периода подведены в монографиях [3, 17, 18].

Следующий этап — выделение статистики объектов нечисловой природы в качестве самостоятельного направления в прикладной статистике, ядром которого являются методы статистического анализа данных произвольной природы. Программа развития этого нового научного направления впервые была сформулирована в статье [19]. Реализация этой программы была осуществлена в 1980-е годы. Для работ этого периода характерна сосредоточенность на внутренних проблемах нечисловой статистики. Ссылки на конкретные монографии, сборники, статьи и иные публикации нескольких десятков авторов приведены в главе 8 монографии [7]. Отметим лишь сборник научных статей [20], полностью посвященный нечисловой статистике.

К началу 1990-х гг. статистика объектов нечисловой природы с теоретической точки зрения была достаточно хорошо развита, основные идеи, подходы и методы были разработаны и изучены математически, в частности, доказано достаточно много теорем. Однако она оставалась недостаточно апробированной на практике. И в 1990-е гг. наступило время от теоретических математико-статистических исследований перейти к применению полученных результатов при решении конкретных задач в различных областях науки и практики.

Важно отметить, что в статистике нечисловых данных, как и в других областях прикладной статистики и прикладной математики вообще, одна и та же математическая схема может с успехом применяться при решении различных задач анализа конкретных данных. В технических исследованиях, и в менеджменте, и в экономике, и в геологии, и в медицине, и в социологии, и для анализа экспертных оценок, и во многих иных областях. А потому ее лучше всего формулировать и изучать в наиболее общем виде, для объектов произвольной природы.

Основные идеи статистики объектов нечисловой природы. В чем принципиальная новизна нечисловой статистики? Для классической математической статистики характерна операция сложения. При расчете выборочных характеристик распределения (выборочное среднее арифметическое, выборочная дисперсия и др.), в регрессионном анализе и других областях этой научной дисциплины постоянно используются суммы. Математический аппарат — законы больших чисел, Центральная предельная теорема и другие теоремы — нацелены на изучение сумм. В нечисловой же статистике нельзя использовать операцию сложения, поскольку элементы выборки лежат в пространствах, где нет операции сложения. Методы обработки нечисловых данных основаны на принципиально ином математическом аппарате — на применении различных расстояний в пространствах объектов нечисловой природы.

Кратко рассмотрим несколько идей, развиваемых в статистике объектов нечисловой природы для данных, лежащих в пространствах произвольного вида. Решаются классические задачи описания данных, оценивания, проверки гипотез — но для неклассических данных, а потому неклассическими методами.

Первой обсудим проблему определения средних величин. В рамках репрезентативной теории измерений удается указать вид средних величин, соответствующих тем или иным шкалам измерения (см. главу 11.3). В классической математической статистике эмпирические и теоретические средние величины вводят с помощью операций сложения (выборочное среднее арифметическое, математическое ожидание) или упорядочения (выборочная и теоретическая медианы). В пространствах произвольной природы средние значения нельзя определить с помощью операций сложения или упорядочения. Теоретические и эмпирические средние приходится вводить как решения экстремальных задач. Для теоретического среднего это — задача минимизации математического ожидания (в классическом смысле) расстояния от случайного элемента со значениями в рассматриваемом пространстве до фиксированной точки этого пространства (минимизируется указанная функция от этой точки). Для эмпирического среднего математическое ожидание берется по эмпирическому распределению, т.е. берется сумма расстояний от некоторой точки до элементов выборки и затем минимизируется по этой точке. При этом как эмпирическое, так и теоретическое средние как решения экстремальных задач могут быть не единственными элементами пространства, а описываться множествами таких элементов, которые могут оказаться и пустыми. Несмотря на возможность неоднозначности или пустоты решений экстремальных задач, удалось сформулировать и доказать законы больших чисел для средних величин, определенных указанным образом, т.е. установить сходимость эмпирических средних к теоретическим.

Хорошая теория дает больше того, что от нее вначале ожидалось. Удалось установить, что методы доказательства законов больших чисел допускают существенно более широкую область применения, чем та, для которой они были разработаны. А именно, с помощью этих методов удалось изучить асимптотику решений экстремальных статистических задач, к которым, как известно, сводится большинство постановок прикладной статистики. В частности, кроме законов больших чисел установлена и состоятельность оценок минимального контраста, в том числе оценок максимального правдоподобия и робастных оценок. К настоящему времени подобные оценки изучены также и в интервальной статистике.

В статистике в пространствах произвольной природы большую роль играют непараметрические оценки плотности, используемые, в частности, в различных алгоритмах регрессионного, дискриминантного, кластерного анализов. В нечисловой статистике предложен и изучен ряд типов непараметрических оценок плотности в пространствах произвольной природы, в частности, доказана их состоятельность, изучена скорость сходимости и установлен примечательный факт совпадения наилучшей скорости сходимости в произвольном случае с той, которая имеет быть в классической математико-статистической теории для числовых случайных величин.

Дискриминантный, кластерный, регрессионный анализы в пространствах произвольной природы основаны либо на параметрической теории — и тогда применяется подход, связанный с асимптотикой решения экстремальных статистических задач — либо на непараметрической теории — и тогда используются алгоритмы на основе непараметрических оценок плотности.

Для проверки гипотез могут быть использованы статистики интегрального типа, в частности, типа омега-квадрат. Любопытно, что предельная теория таких статистик, построенная первоначально в классической постановке [21], приобрела естественный (завершенный, изящный) вид именно для пространств произвольного вида [22], поскольку при этом удалось провести рассуждения, опираясь на базовые математические соотношения, а не на те частные (с общей точки зрения), что были связаны с конечномерным пространством.

Представляют практический интерес результаты, связанные с конкретными областями статистики нечисловых данных. В частности, со статистикой нечетких и случайных множеств (напомним, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств), с непараметрической теорией парных сравнений, с аксиоматическим введением метрик в конкретных пространствах объектов нечисловой природы, и с рядом других конкретных постановок.

Для анализа нечисловых, в частности, экспертных данных весьма важны методы классификации. С другой стороны, наиболее естественно ставить и решать задачи классификации, основанные на использовании расстояний или показателей различия, в рамках статистики нечисловых данных. Это касается как распознавания образов с учителем (другими словами, дискриминантного анализа), так и распознавания образов без учителя (т.е. кластерного анализа).

Статистические методы анализа нечисловых данных особенно хорошо приспособлены для применения в экономике, социологии и экспертных оценках, поскольку в этих областях от 50 до 90 % данных являются нечисловыми.

Итак, статистика нечисловых данных является центром прикладной статистики. А ее теоретическая основа — статистика в пространствах произвольной природы — является стержнем математической статистики.

Другие точки роста. Выше рассмотрены пять «точек роста» прикладной статистики. Разумеется, они не исчерпывают все многообразие фронта научных исследований в рассматриваемых областях. Кроме того, мы почти не затронули разнообразные применения статистических методов в конкретных прикладных исследованиях и разработках. Много интересных проблем есть в планировании экспериментов, особенно кинетических (см., например, [23]), при анализе проблем надежности, в новых статистических методах управления качеством продукции [7], в том числе в связи с идеями Г. Тагути, при анализе рисков, в вопросах экологии и промышленной безопасности и др.

В течение последних более чем 60 лет в России наблюдается огромный разрыв между государственной статистикой и научным сообществом специалистов по статистическим методам (подробнее об этом см. статью [24]). В учебнике по истории статистики [25] даже не упоминаются имена членов-корреспондентов АН СССР Н.В. Смирнова и Л.Н. Большева! А ведь они — единственные представители именно математической статистики как таковой в Академии наук в XX в. (еще ряд членов отечественной Академии наук имели математическую статистику среди своих интересов, но Н.В. Смирнов и Л.Н. Большев занимались практически только ею).

16.2. Высокие статистические технологии

При практическом использовании методов прикладной статистики применяются не отдельные методы описания данных, оценивания, проверки гипотез, а развернутые цельные процедуры — так называемые «статистические технологии». Понятие «статистическая технология» аналогично понятию «технологический процесс» в теории и практике организации производства.

Статистические технологии. Статистический анализ конкретных данных, как правило, включает в себя целый ряд процедур и алгоритмов, выполняемых последовательно, параллельно или по более сложной схеме. Как уже говорилось во Введении к настоящей книге, с точки зрения организатора прикладного статистического исследования можно выделить следующие этапы:

- планирование статистического исследования (включая разработку анкет, бланков наблюдения и учета и других форм сбора данных; их апробацию; подготовку сценариев интервью и анализа данных и т.п.);

- организация сбора необходимых статистических данных по оптимальной или рациональной программе (планирование выборки, создание организационной структуры и подбор команды статистиков, подготовка кадров, которые будут заниматься сбором данных, а также контролеров данных и т.п.);

- непосредственный сбор данных и их фиксация на тех или иных носителях (с контролем качества сбора и отбраковкой ошибочных данных по соображениям предметной области);

- первичное описание данных (расчет различных выборочных характеристик, функций распределения, непараметрических оценок плотности, построение гистограмм, корреляционных полей, различных таблиц и диаграмм и т.д.);

- оценивание тех или иных числовых или нечисловых характеристик и параметров распределений (например, непараметрическое интервальное оценивание коэффициента вариации или восстановление зависимости между откликом и факторами, т.е. оценивание функции);

- проверка статистических гипотез (иногда их цепочек — после проверки предыдущей гипотезы принимается решение о проверке той или иной последующей гипотезы; например, после проверки адекватности линейной регрессионной модели и отклонения этой гипотезы может проверяться адекватность квадратичной модели);

- более углубленное изучение, т.е. одновременное применение различных алгоритмов многомерного статистического анализа, алгоритмов диагностики и построения классификации, статистики нечисловых и интервальных данных, анализа временных рядов и др.;

- проверка устойчивости полученных оценок и выводов относительно допустимых отклонений исходных данных и предпосылок используемых вероятностно-статистических моделей, в частности, изучение свойств оценок методом размножения выборок и другими численными методами;

- применение полученных статистических результатов в прикладных целях, т.е. для формулировки выводов в терминах содержательной области (например, для диагностики конкретных материалов, построения прогнозов, выбора инвестиционного проекта из предложенных вариантов, нахождения оптимальных режима осуществления технологического процесса, подведения итогов испытаний образцов технических устройств и др.);

- составление итоговых отчетов, в частности, предназначенных для тех, кто не является специалистами в статистических методах анализа данных, в том числе для руководства — «лиц, принимающих решения».

Возможны и иные структуризации различных статистических технологий. Важно подчеркнуть, что квалифицированное и результативное применение статистических методов — это отнюдь не проверка одной отдельно взятой статистической гипотезы или оценка характеристик или параметров одного заданного распределения из фиксированного семейства. Подобного рода операции — только отдельные кирпичики, из которых складывается статистическая технология.

Итак, процедура статистического анализа данных — это информационный технологический процесс, другими словами, та или иная информационная технология. Статистическая информация подвергается разнообразным операциям (последовательно, параллельно или по более сложным схемам). В настоящее время об автоматизации всего процесса статистического анализа данных говорить было бы несерьезно, поскольку имеется слишком много нерешенных проблем, вызывающих дискуссии среди статистиков. Наличие разногласий — причина того, что так называемые «экспертные системы в области статистического анализа данных» пока не стали рабочим инструментом статистиков.

Проблема «стыковки» алгоритмов. В литературе статистические технологии рассматриваются явно недостаточно. В частности, обычно все внимание сосредотачивается на том или ином элементе технологической цепочки, а переход от одного элемента к другому остается в тени. Между тем проблема «стыковки» статистических алгоритмов, как известно, требует специального рассмотрения (см. главу 4.5), поскольку в результате использования предыдущего алгоритма зачастую нарушаются условия применимости последующего. В частности, результаты наблюдений могут перестать быть независимыми, может измениться их распределение и т.п.

Так, вполне резонной выглядит рекомендация: сначала разбейте данные на однородные группы, а потом в каждой из групп проводите статистическую обработку, например, регрессионный анализ. Однако эта рекомендация под кажущейся прозрачностью содержит подводные камни. Действительно, как поставить задачу в вероятностно-статистических терминах? Если, как обычно, примем, что исходные данные — это выборка, т.е. совокупность независимых одинаково распределенных случайных элементов, то классификация приведет к разбиению этих элементов на группы. В каждой группе элементы будут зависимы между собой, а их распределение будет зависеть от группы, куда они попали. Отметим, что в типовых ситуациях границы классов стабилизируются, а это значит, что асимптотически элементы кластеров становятся независимыми. Однако их распределение не может быть нормальным. Например, если исходное распределение было нормальным, то распределения в классах будет усеченным нормальным. Это означает, что необходимо пользоваться непараметрическими методами.

Разберем другой пример. При проверке статистических гипотез большое значение имеют такие хорошо известные характеристики статистических критериев, как уровень значимости и мощность. Методы их расчета и использования при проверке одной гипотезы обычно хорошо известны. Если же сначала проверяется одна гипотеза, а потом с учетом результатов ее проверки (конкретнее, если первая гипотеза принята) — вторая, то итоговую процедуру также можно рассматривать как проверку некоторой (более сложной) статистической гипотезы. Она имеет характеристики (уровень значимости и мощность), которые, как правило, нельзя простыми формулами выразить через характеристики двух составляющих гипотез, а потому они обычно неизвестны. Лишь в некоторых простых случаях характеристики итоговой процедуры можно рассчитать. В результате итоговую процедуру нельзя рассматривать как научно обоснованную, она относится к эвристическим алгоритмам. Конечно, после соответствующего изучения, например, методом Монте-Карло, она может войти в число научно обоснованных процедур прикладной статистики.

Термин «высокие статистические технологии». Термин «высокие технологии» популярен в современной научно-технической литературе. Он используется для обозначения наиболее передовых технологий, опирающихся на последние достижения научно-технического прогресса. Есть такие технологии и среди технологий статистического анализа данных — как в любой интенсивно развивающейся научно-практической области.

Примеры высоких статистических технологий и входящих в них алгоритмов анализа данных, подробный анализ современного состояния и перспектив развития даны выше при обсуждении «точек роста». В качестве «высоких статистических технологий» были выделены технологии непараметрического анализа данных; устойчивые (робастные) технологии; технологии, основанные на размножении выборок, на использовании достижений статистики нечисловых данных и статистики интервальных данных.

Обсудим пока не вполне привычный термин «высокие статистические технологии». Каждое из трех слов несет свою смысловую нагрузку.

«Высокие», как и в других областях, означает, что статистическая технология опирается на современные достижения статистической теории и практики, в частности, на достижения теории вероятностей и прикладной математической статистики. При этом «опирается на современные научные достижения» означает, во-первых, что математическая основа технологии получена сравнительно недавно в рамках соответствующей научной дисциплины, во-вторых, что алгоритмы расчетов разработаны и обоснованы в соответствии с нею (а не являются т.н. «эвристическими»). Со временем новые подходы и результаты могут заставить пересмотреть оценку применимости и возможностей технологии, привести к замене ее более современной. В противном случае «высокие статистические технологии» переходят в «классические статистические технологии», такие, как метод наименьших квадратов. Итак, высокие статистические технологии — плоды недавних серьезных научных исследований. Здесь два ключевых понятия — «молодость» технологии (во всяком случае, не старше 50 лет, а лучше — не старше 10 или 30 лет) и опора на «высокую науку».

Термин «статистические» привычен, но коротко разъяснить его нелегко. Проще сослаться на введение и все содержание настоящего учебника, на энциклопедию [26], книги [1, 7] и др. В частности, статистические данные — это результаты измерений, наблюдений, испытаний, анализов, опытов. А «статистические технологии» — это технологии анализа статистических данных.

Наконец, редко используемый применительно к статистике термин «технологии». Статистический анализ данных, как правило, включает в себя целый ряд процедур и алгоритмов, выполняемых последовательно, параллельно или по более сложной схеме. Структура типовой статистической технологии описана выше. Обработка статистических данных — это информационный технологический процесс.

Всегда ли нужны «высокие статистические технологии»? «Высоким статистическим технологиям» противостоят, естественно, «низкие статистические технологии» (а между ними расположены «классические статистические технологии»). «Низкие статистические технологии» — это те технологии, которые не соответствуют современному уровню науки и практики. Обычно они одновременно и устарели, и не вполне адекватны сути решаемых статистических задач.

Примеры таких технологий неоднократно критически рассматривались, в том числе и на страницах этой книги. Достаточно вспомнить критику использования критерия Стьюдента для проверки однородности при отсутствии нормальности и равенства дисперсии. Или применение критерия Вилкоксона для проверки совпадения теоретических медиан или функций распределения двух выборок. Или использование классических процентных точек критериев Колмогорова и омега-квадрат в ситуациях, когда параметры оцениваются по выборке и эти оценки подставляются в «теоретическую» функцию распределения. На первый взгляд вызывает удивление устойчивость «низких статистических технологий», их постоянное возрождение во все новых статьях, монографиях, учебниках. Поэтому, как ни странно, наиболее «долгоживущими» оказываются не работы, посвященные новым научным результатам, а публикации, разоблачающие ошибки, типа статьи [27]. Прошло около 20 лет с момента ее публикации, но она по-прежнему актуальна, поскольку ошибочное применение критериев Колмогорова и омега-квадрат по-прежнему распространено.

Целесообразно отметить по крайней мере четыре обстоятельства, которые определяют эту устойчивость ошибок. Во-первых, прочно закрепившаяся традиция. Так, многие учебники по курсам типа «Общей теории статистики», если беспристрастно проанализировать их содержание, состоят в основном из введения в прикладную статистику. Иногда изложение идет в стиле «низких статистических технологий», т.е. на уровне 1950-х гг., а во многом и на уровне начала XX в. К «низкой» прикладной статистике добавлена некоторая информация о деятельности органов Госкомстата РФ. Новое поколение, обучившись «низким» подходам, идеям, алгоритмам, их использует, а с течением времени и достижением должностей, ученых званий и степеней — пишет новые учебники со старыми ошибками.

Второе обстоятельство связано с большими трудностями при оценке экономической эффективности применения статистических методов вообще и при оценке вреда от применения ошибочных методов в частности. (А без такой оценки как докажешь, что «высокие статистические технологии» лучше «низких»?) При оценке вреда от применения ошибочных методов приходится учитывать, что общий успех в конкретной инженерной или научной работе вполне мог быть достигнут вопреки применению ошибочных методов, за счет «запаса прочности» других составляющих общей работы. Например, преимущество одного технологического приема над другим можно продемонстрировать как с помощью критерия Крамера — Уэлча проверки равенства математических ожиданий (что правильно), так и с помощью двухвыборочного критерия Стьюдента (что, вообще говоря, неверно, т.к. обычно не выполняются условия применимости этого критерия — нет ни нормальности распределения, ни равенства дисперсий).

Третье существенное обстоятельство — трудности со знакомством с высокими статистическими технологиями. В нашей стране в силу ряда исторических обстоятельств развития статистических методов в течение последних 10 лет только журнал «Заводская лаборатория» предоставлял такие возможности. К сожалению, поток современных отечественных и переводных статистических книг, выпускавшихся ранее, в частности, издательствами «Наука», «Мир», «Финансы и статистика», практически превратился в узкий ручеек… Возможно, более существенным является влияние естественной задержки во времени между созданием «новых статистических технологий» и написанием полноценной и объемной учебной и методической литературы. Она должна позволять знакомиться с новой методологией, новыми методами, теоремами, алгоритмами, методами расчетов и интерпретации их результатов, статистическими технологиями в целом не по кратким оригинальным статьям, а при обычном вузовском и последипломном обучении.

И, наконец, наиболее важное. Всегда ли нужны высокие статистические технологии? Приведем аналогию — нужна ли современная сельскохозяйственная техника для обработки приусадебного участка? Нужны ли трактора и комбайны? Может быть, достаточно технологий, основанных на использовании лопаты? Вернемся к данным государственной статистики. Применяются статистические технологии первичной обработки (описания) данных, основанные на построении разнообразных таблиц, диаграмм, графиков. Большинство потребителей статистической информации это представление данных удовлетворяет. Итак, чтобы высокие статистические технологии успешно использовались, необходимы два условия:

- чтобы они были *объективно* нужны для решения практической задачи;
- чтобы потенциальный пользователь технологий *субъективно* понимал это.

Таким образом, весь арсенал реально используемых в настоящее время эконометрических и статистических технологий можно распределить по трем потокам:

- высокие статистические технологии;
- классические статистические технологии;
- низкие статистические технологии.

Под классическими статистическими технологиями, как уже отмечалось, понимаем технологии почтенного возраста, сохранившие свое значение для современной статистической практики. Таковы технологии на основе метода наименьших квадратов (включая методы точечного оценивания параметров прогностической функции, непараметрические методы доверительного оценивания параметров и прогностической функции в целом, проверок различных гипотез о них), статистик типа Колмогорова, Смирнова, омега-квадрат, непараметрических коэффициентов корреляции Спирмена и Кендалла (относить их только к методам анализа ранжировок — значит делать уступку «низким статистическим технологиям») и многих других статистических процедур.

Основная проблема в области статистических технологий в настоящее время состоит в том, чтобы в конкретных эконометрических исследованиях использовались только технологии первых двух типов.

Каковы возможные пути решения этой проблемы? Бороться с конкретными невеждами — дело почти безнадежное. Конечно, необходима демонстрация квалифицированного применения высоких статистических технологий. В 1960–1970-х гг. этим занималась Лаборатория статистических методов акад. А.Н. Колмогорова в МГУ им. М.В. Ломоносова. В секции «Математические методы исследования» журнала «Заводская лаборатория» за последние 40 лет опубликовано более 1 000 статей, выполненных на уровне «высоких статистических технологий». В настоящее время действует Институт высоких статистических технологий и эконометрики и целый ряд других научных коллективов, работающих на уровне «высоких статистических технологий».

Очевидно, самое основное — это обучение. Какие бы новые научные результаты ни были получены, если они остаются неизвестными студентам, то новое поколение исследователей и инженеров, экономистов и менеджеров, других специалистов вынуждено осваивать их поодиночке, в порядке самообразования, а то и переоткрывать заново. Т.е. зачастую новые научные результаты практически исчезают из оборота научной и практической информации, едва появившись. Как ни странно, избыток научных публикаций превратился в тормоз развития науки. По нашим данным, к настоящему времени по статистическим технологиям опубликовано не менее миллиона статей и книг, в основном во второй половине XX в. Из них не менее 100 тысяч являются актуальными для современного специалиста. При этом реальное число публикаций, которые способен освоить исследователь за свою профессиональную жизнь, по нашей оценке, не превышает 2–3 тысяч. Итак, каждый специалист в области прикладной статистики знаком не более чем с 2–3 % актуальных для него литературных источников. Поскольку существенная часть публикаций заражена «низкими статистическими технологиями», то исследователь-самоучка, увы, имеет мало шансов выйти на уровень «высоких статистических технологий». С подтверждениями этого печального вывода постоянно приходится сталкиваться. Одновременно приходится констатировать, что масса полезных результатов погребена в изданиях прошлых десятилетий и имеет мало шансов пробиться в ряды используемых в настоящее время «высоких статистических технологий» без специально организованных усилий современных специалистов.

Итак, основное — обучение. Несколько огрубляя, можно сказать так: что попало в учебные курсы и соответствующие учебные пособия — то сохраняется, что не попало — то пропадает.

Необходимость высоких статистических технологий. Может возникнуть естественный вопрос: зачем нужны высокие статистические технологии, разве недостаточно обычных статистических методов? Специалисты по прикладной статистике справедливо считают и доказывают своими теоретическими и прикладными работами, что совершенно недостаточно. Так, совершенно очевидно, что многие данные в информационных системах имеют нечисловой характер, например, являются словами или принимают значения из конечных множеств. Нечисловой характер имеют и упорядочения, которые дают эксперты или менеджеры, например, выбирая главную цель, следующую по важности и т.д. Значит, нужна статистика нечисловых данных. Мы ее построили. Далее, многие величины известны не абсолютно точно, а с некоторой погрешностью — от и до. Другими словами, исходные данные — не числа, а интервалы. Нужна статистика интервальных данных. Мы ее развиваем. В широко известной монографии по контроллингу [28] на с. 138 хорошо сказано: «Нечеткая логика — мощный элегантный инструмент современной науки, который на Западе (и на Востоке — в Японии, Китае — А.О.) можно встретить в десятках изделий — от бытовых видеокамер до систем управления вооружениями, — у нас до самого последнего времени был практически неизвестен». Напомним, первая монография российского автора по теории нечеткости содержит основы высоких статистических технологий, связанные с анализом выборок нечетких множеств (см. книгу [29]). Ни статистики нечисловых данных, ни статистики интервальных данных, ни статистики нечетких данных не было и не могло быть в классической статистике. Все это — высокие статистические технологии. Они разработаны за последние десятилетия. А обычные вузовские курсы по общей теории статистики и по математической статистике разбирают научные результаты, полученные в первой половине XX в.

Важная и весьма перспективная часть прикладной статистики — применение высоких статистических технологий к анализу конкретных данных, что зачастую требует дополнительных теоретических исследований по доработке статистических технологий применительно к конкретной ситуации. Большое значение имеют конкретные статистические модели, например, модели экспертных оценок или эконометрики качества. И конечно, такие конкретные применения, как расчет и прогнозирование индекса инфляции. Сейчас уже многим экономистам и менеджерам ясно, что годовой бухгалтерский баланс предприятия может быть использован для оценки его финансово-хозяйственной деятельности только с привлечением данных об инфляции.

Институт высоких статистических технологий и эконометрики. Опишем опыт внедрения «высоких статистических технологий». Организованный нами в 1989 г. Институт высоких статистических технологий и эконометрики (ИВСТЭ) в настоящее время действует на базе кафедры ИБМ-2 «Экономика и организация производства» Московского государственного технического университета им. Н.Э. Баумана. Институт на хоздоговорных и госбюджетных началах занимается развитием, изучением и внедрением эконометрики и «высоких статистических технологий», т.е. наиболее современных технологий анализа экономических, технических, социологических, медицинских данных, ориентированных на использование в условиях современного производства и экономики. Основной интерес представляют применения «высоких статистических технологий» для анализа конкретных экономических данных, т.е. в эконометрике. Наиболее перспективным представляется применение «высоких статистических технологий» для поддержки принятия управленческих решений, прежде всего в таком новом (для России) современном направлении экономической науки и практики, как контроллинг (см., например, монографию [28]).

Вначале Институт действовал как Всесоюзный центр статистических методов и информатики Центрального правления Всесоюзного экономического общества. В 1990–1992 гг. было выполнено более 100 хоздоговорных работ, в том числе для НИЦентра по безопасности атомной энергетики, ВНИИ нефтепереработки, ПО «Пластик», ЦНИИ черной металлургии им. Бардина, НИИ стали, ВНИИ эластомерных материалов и изделий, НИИ прикладной химии, ЦНИИ химии и механики, НПО «Орион», ВНИИ экономических проблем развития науки и техники, ПО «Уралмаш», «АвтоВАЗ», МИИТ, Казахского политехнического института, Донецкого государственного университета, Института питания (Алма-Ата) и многих других.

Затем Институт разрабатывал эконометрические методы анализа нечисловых данных, а также процедуры расчета и прогнозирования индекса инфляции и валового внутреннего продукта. ИВСТЭ развивал методологию построения и использования математических моделей процессов налогообложения (для Министерства налогов и сборов РФ), методологию оценки рисков реализации инновационных проектов высшей школы (для Министерства промышленности, науки и технологий РФ). Институт оценивал влияние различных факторов на формирование налогооблагаемой базы ряда налогов (для Минфина РФ), прорабатывал перспективы применения современных статистических и экспертных методов для анализа данных о научном потенциале (для Министерства промышленности, науки и технологий РФ). Важное направление связано с эколого-экономической тематикой — разработка методологического, программного и информационного обеспечения анализа рисков химико-технологических объектов (для Международного научно-технического центра), методов использования экспертных оценок в задачах экологического страхования (совместно с Институтом проблем рынка РАН). Институт проводил маркетинговые исследования (в частности, для *Institute for Market Research GfK MR*, Промрадтехбанка, фирм, торгующих растворимым кофе, программным обеспечением, оказывающих образовательные услуги). Интерес вызывали работы Института по прогнозированию социально-экономического развития России методом сценариев, по экономико-математическому моделированию развития малых предприятий и созданию современных систем информационной поддержки принятия решений для таких организаций.

Институт ведет фундаментальные исследования в области высоких статистических технологий и эконометрики, в частности, в рамках МГТУ им. Н.Э. Баумана и Российского фонда фундаментальных исследований. Информация об Институте представлена на сайте «Высокие статистические технологии» (http://orlovs.pp.ru) и форуме (https://orlovs.pp.ru/forum/index.php). Институтом издается компьютерный еженедельник «Эконометрика» — https://subscribe.ru/catalog/science.humanity.econometrika (около 1 000 подписчиков). Архив выпусков еженедельника «Эконометрика» можно рассматривать как хрестоматию по различным разделам эконометрики, а также по высоким статистическим технологиям.

Современная эконометрика пока мало известна в России. А между тем в мировой науке эконометрика занимает достойное место. Напомним, что Нобелевские премии по экономике получили эконометрики Ян Тильберген, Рагнар Фриш, Лоуренс Клейн, Трюгве Хаавельмо, Джеймс Хекман и Дэниель Мак-Фадден, Энгл Грейнджер и Кеннет Риглз. Выпускается ряд научных журналов, полностью посвященных эконометрике, в том числе: *Journal of*

Econometrics (Швеция), Econometric Reviews (США), *Econometrica* (США), Sankhya (*Indian Journal of Statistics. Ser.D. Quantitative Economics.* Индия), *Publications Econometriques* (Франция). Применение эконометрики дает заметный экономический эффект. Например, в США — не менее 20 миллиардов долларов ежегодно только в области статистического контроля качества.

Однако в нашей стране по ряду причин прикладная статистика и эконометрика не были сформированы как самостоятельные направления научной и практической деятельности, в отличие, например, от Польши, не говоря уже об англосаксонских странах. В результате специалистов в области прикладной статистики и эконометрики у нас на порядок меньше, чем в США и Великобритании (Американская статистическая ассоциация включает более 20 000 членов).

О подготовке специалистов по высоким статистическим технологиям. Приходится с сожалением констатировать, что в России плохо налажена подготовка специалистов по высоким статистическим технологиям. В курсах по теории вероятностей и математической статистике обычно даются лишь классические основы этих дисциплин, разработанные в первой половине XX в., а преподаватели-математики свою научную деятельность предпочитают посвящать доказательству теорем, имеющих лишь внутриматематическое значение, а не развитию высоких статистических технологий. В настоящее время появилась надежда на эконометрику. В России начинают развертываться эконометрические исследования и преподавание эконометрики. Экономисты, менеджеры и инженеры, прежде всего специалисты по контроллингу, должны быть вооружены современными средствами информационной поддержки, в том числе высокими статистическими технологиями и эконометрикой. Очевидно, преподавание должно идти впереди практического применения. Ведь как применять то, чего не знаешь?

Приведем два примера — отрицательный и положительный, — показывающие связь преподавания с внедрением передовых технологий.

Один раз — в 1990–1992 гг. мы уже обожглись на недооценке необходимости предварительной подготовки тех, для кого предназначены современные программные продукты. Наш коллектив (Всесоюзный центр статистических методов и информатики Центрального Правления Всесоюзного экономического общества) разработал систему диалоговых программных систем обеспечения качества продукции. Их созданием руководили ведущие специалисты страны. Но распространение программных продуктов шло на 1–2 порядка медленнее, чем мы ожидали. Причина стала ясна не сразу. Как оказалось, работники предприятий просто не понимали возможностей разработанных систем, не знали, какие задачи можно решать с их помощью, какой экономический эффект они дадут. А не понимали и не знали потому, что в вузах никто их не учил статистическим методам управления качеством. Без такого систематического обучения нельзя обойтись — сложные концепции «на пальцах» за пять минут не объяснишь.

Есть и противоположный пример — положительный. В середине 1980-х гг. в советской средней школе ввели новый предмет «Информатика». И сейчас молодое поколение превосходно владеет компьютерами, мгновенно осваивая быстро появляющиеся новинки, и этим заметно отличается от тех, кому за 30–40 лет.

Если бы удалось ввести в средней школе курс теории вероятностей и статистики, то ситуация с внедрением высоких статистических технологий могла бы быть резко улучшена. Такой курс есть в Японии и США, Швейцарии, Кении и Ботсване, почти во всех странах (и ЮНЕСКО проводит всемирные конференции по преподаванию статистики в средней школе — см. сборник докладов [30]). Надо, конечно, добиться того, чтобы этот курс был построен на высоких статистических технологиях, а не на низких. Другими словами, он должен отражать современные достижения, а не концепции пятидесятилетней или столетней давности.

16.3. Компьютерно-статистические методы

Методы статистических испытаний (Монте-Карло). Многие информационные технологии в области прикладной статистики опираются на использование методов статистических испытаний. Этот термин применяется для обозначения компьютерных технологий, в которых в модель реального явления или процесса искусственно вводится большое число случайных элементов. Обычно моделируется последовательность независимых одинаково распределенных случайных величин или же последовательность, построенная на ее основе, например, последовательность накапливающихся (кумулятивных) сумм.

Необходимость в методе статистических испытаний возникает потому, что чисто теоретические методы дают точное решение, как правило, лишь в исключительных случаях. Либо тогда, когда исходные случайные величины имеют вполне определенные функции распределения, например, нормальные, чего, как правило, не бывает. Либо когда объемы выборок очень велики (с практической точки зрения — бесконечны).

Не только в задачах обработки данных возникает необходимость в методе статистических испытаний. Она не менее актуальна и при экономико-математическом моделировании технических, социально-экономических, медицинских и иных процессов. Представим себе всем знакомый объект — торговый зал самообслуживания по продаже продовольственных товаров. Сколько нужно работников в зале, сколько касс? Необходимо просчитать загрузку в разное время суток, в разные сезоны года, с учетом замены товаров и смены сотрудников. Нетрудно увидеть, что теоретическому анализу подобная система не поддается, а компьютерному — вполне.

Методы статистических испытаний стали развиваться после второй мировой войны с появлением компьютеров. Второе название — методы Монте-Карло — они получили по наиболее известному игорному дому, а точнее, по его рулетке, поскольку исходный материал для получения случайных чисел с произвольным распределением — это случайные натуральные числа.

В методах статистических испытаний можно выделить две составляющие. Базой являются датчики псевдослучайных чисел. Результатом работы таких датчиков являются последовательности чисел, которые обладают некоторыми свойствами последовательностей случайных величин (в смысле теории вероятностей). Надстройкой являются различные алгоритмы, использующие последовательности псевдослучайных чисел.

Что же это могут быть за алгоритмы? Приведем примеры. Пусть мы изучаем распределение некоторой статистики при заданном объеме выборки. Тогда естественно много раз (например, 100 000 раз) смоделировать выборку заданного объема (т.е. набор независимых одинаково распределенных случайных величин) и рассчитать значение статистики. Затем по 100 000 значениям статистики можно достаточно точно построить функцию распределения изучаемой статистики, оценить ее характеристики. Однако эта схема годится лишь для так называемой «свободной от распределения» статистики, распределение которой не зависит от распределения элементов выборки. Если же такая зависимость есть, то одной точкой моделирования не обойдешься, придется много раз моделировать выборку, беря различные распределения, меняя параметры. Чтобы общее время моделирования было приемлемым, возможно, придется сократить число моделирований в одной точке, зато увеличив общее число точек. Точность моделирования может быть оценена по общим правилам выборочных обследований.

Второй пример — частично описанное выше моделирование работы торгового зала самообслуживания по продаже продовольственных товаров. Здесь одна последовательность псевдослучайных чисел описывает интервалы между появлениями покупателей, вторая, третья и т.д. связаны с выбором ими первого, второго и т.д. товаров в зале (например, число — номер в перечне товаров). Короче, все действия покупателей, продавцов, работников предприятия разбиты на операции, каждая операция, в продолжительности или иной характеристике которой имеется случайность, моделируется с помощью соответствующей последовательности псевдослучайных чисел. Затем итоги работы сотрудников торговой организации и зала в целом выражаются через характеристики случайных величин. Формулируется критерий оптимальности, решается задача оптимизации и находятся оптимальные значения параметров. В частности, оптимальные планы статистического контроля строятся на основе вероятностно-статистических моделей (см. главу 10).

Датчики псевдослучайных чисел. Теперь обсудим свойства датчиков псевдослучайных чисел. Здесь стоит слово «псевдослучайные», а не «случайные». Это весьма важно. Дело в том, что за последние 50 лет обсуждались в основном три принципиально разных варианта получения последовательностей чисел, которые в дальнейшем использовались в методах статистических испытаний.

Первый — таблица случайных чисел. К сожалению, объем любой таблицы конечен, и сколько-нибудь сложные расчеты с ее помощью невозможны. Через некоторое время приходится повторяться. Кроме того, обычно обнаруживались те или иные отклонения от случайности.

Второй — физические датчики случайных чисел. Основной недостаток — нестабильность, непредсказуемые отклонения от заданного распределения (обычно — равномерного).

Третий — расчетный. В простейшем случае каждый следующий член последовательности рассчитывается по предыдущему. Например, так:

$$z_{n+1} \equiv Mz_n \pmod{P},$$

где z_0 — начальное значение (заданное целое положительное число), M — параметр алгоритма (заданное целое положительное число), $P = 2^m$, где m — число двоичных разрядов представления чисел, с которыми манипулирует компьютер. Знак ≡ здесь означает теоретико-числовую операцию сравнения, т.е. взятие дробной части от $\frac{Mz_n}{P}$ и отбрасывание целой части.

В настоящее время применяется именно третий вариант. Совершенно ясно, что он не соответствует интуитивному представлению о случайности. Например, интуитивно очевидно, что по предыдущему элементу случайной последовательности с независимыми элементами нельзя предсказать значение следующего элемента. А приведенная выше формула как раз и дает способ такого предсказания. Расчетный путь получения последовательности псевдослучайных чисел противоречит не только интуиции, но и подходу к определению случайности на основе теории алгоритмов, развитому акад. А.Н. Колмогоровым и его учениками в 1960-х гг. Однако во многих прикладных задачах он работает, и это основное.

Методу статистических испытаний посвящена обширная литература. Время от времени обнаруживаются недостатки у популярных датчиков псевдослучайных чисел. Так, например, в середине 1980-х гг. выяснилось, что для одного из наиболее известных датчиков три последовательных значения связаны линейной зависимостью

$$Z_{n+2} = aZ_{n+1} + bZ_n, \quad n = 1,2,\ldots$$

После этого в 1985 г. в журнале «Заводская лаборатория» началась дискуссия о качестве датчиков псевдослучайных чисел, которая продолжалась до 1993 г. и закончилась статьей проф. С.М. Ермакова и нашим комментарием.

Итоги можно подвести так. Во многих случаях решаемая методом статистических испытаний задача сводится к оценке вероятности попадания в некоторую область в многомерном пространстве *фиксированной* размерности. Тогда из чисто математических соображений теории чисел следует, что с помощью датчиков псевдослучайных чисел поставленная задача решается корректно. Сводка соответствующих математических обоснований приведена, например, в работе С.М. Ермакова.

В других случаях приходится рассматривать вероятности попадания в области в пространствах *переменной* размерности. Типичным примером является ситуация, когда на каждом шагу проводится проверка, и по ее результатам либо остаемся в данном пространстве, либо переходим в пространство большей размерности. Например, в главе 6.3 при оценивании степени многочлена либо останавливались на данной степени, либо увеличивали степень, переходя в параметрическое пространство большей размерности. Так вот, вопрос об обоснованности применения метода статистических испытаний (а точнее, о свойствах датчиков псевдослучайных чисел) в случае пространств переменной размерности остается в настоящее время открытым. О важности этой проблемы вдохновенно говорил академик РАН Ю.В. Прохоров на Первом Всемирном Конгрессе Общества математической статистики и теории вероятностей им. Бернулли (Ташкент, 1986 г.).

Имитационное моделирование. Поскольку постоянно говорим о моделировании, приведем несколько общих формулировок.

Модель в общем смысле (обобщенная модель) — это создаваемый с целью получения и (или) хранения информации специфический объект (в форме мысленного образа, описания знаковыми средствами либо материальной системы), отражающей свойства, характеристики и связи объекта-оригинала произвольной природы, существенные для задачи, решаемой субъектом (это определение взято из монографии [10, с. 44]).

Например, в менеджменте производственных систем используют:

- модели технологических процессов (контроль и управление по технико-экономическим критериям, АСУ ТП — автоматизированные системы управления технологическими процессами);
- модели управления качеством продукции (в частности, модели оценки и контроля надежности);
- модели массового обслуживания (теории очередей);
- модели управления запасами (в современной терминологии — модели логистики, т.е. теории и практики управления материальными, финансовыми и информационными потоками);
- имитационные и эконометрические модели деятельности предприятия (как единого целого) и управления им (АСУ предприятием) и др.

Согласно академику РАН Н.Н. Моисееву, имитационная система — это совокупность моделей, имитирующих протекание изучаемого процесса, объединенная со специальной системой вспомогательных программ и информационной базой, позволяющих достаточно просто и оперативно реализовать вариантные расчеты. Другими словами, имитационная система — это совокупность имитационных моделей. А имитационная модель предназначена для ответов на вопросы типа: «Что будет, если…» Что будет, если параметры примут те или иные значения? Что будет с ценой на продукцию, если спрос будет падать, а число конкурентов расти? Что будет, если государство резко усилит вмешательство в экономику? Что будет, если остановку общественного транспорта перенесут на 100 м дальше от входа в торговый зал, о котором шла речь выше, и поток покупателей резко упадет? Кроме компьютерных моделей, на вопросы подобного типа часто отвечают эксперты при использовании метода сценариев (см. главу 11 выше).

При имитационном моделировании часто используется метод статистических испытаний (Монте-Карло). Теорию и практику машинных имитационных экспериментов с моделями экономических систем еще 35 лет назад подробно разобрал Т. Нейлор. Вернемся к внутристатистическому применению датчиков псевдослучайных чисел.

Методы размножения выборок (бутстреп-методы). Прикладная статистика бурно развивается последние десятилетия. Серьезным (хотя, разумеется, не единственным и не главным) стимулом является стремительно растущая производительность вычислительных средств. Поэтому понятен острый интерес к статистическим методам, интенсивно использующим компьютеры. Одним из таких методов является так называемый «бутстреп», предложенный в 1977 г. Б. Эфроном из Станфордского университета (США).

Сам термин «бутстреп» — это английское слово *bootstrap*, записанное русскими буквами.

В истории прикладной статистики было несколько более или менее успешно осуществленных рекламных кампаний. В каждой из них «раскручивался» тот или иной метод, который, как правило, отвечал нескольким условиям:

- по мнению его пропагандистов, полностью решал актуальную научную задачу;
- был понятен (при постановке задачи, при ее решении и при интерпретации результатов) широким массам потенциальных пользователей;
- использовал современные возможности вычислительной техники.

Пропагандисты метода, как правило, избегали беспристрастного сравнения его возможностей с возможностями иных эконометрических методов. Если сравнения и проводились, то с заведомо слабым «противником».

В нашей стране в условиях отсутствия систематического образования в области прикладной статистики подобные рекламные кампании находили особо благоприятную почву, поскольку у большинства затронутых ими специалистов не было достаточных знаний в области методологии построения моделей прикладной статистики для того, чтобы составить самостоятельное квалифицированное мнение.

Речь идет о таких методах и постановках, как бутстреп, нейронные сети, метод группового учета аргументов, робастные оценки по Тьюки — Хуберу, интеллектуальный анализ данных, генетические алгоритмы, логит, пробит, SVM, асимптотика пропорционального роста числа параметров и объема данных и др. Бывают локальные всплески неоправданного энтузиазма. Например, московские социологи в 1980-х гг. весьма активно пропагандировали так называемый «детерминационный анализ» — простой эвристический метод анализа таблиц сопряженности. Хотя в Новосибирске в это время давно уже было разработано продвинутое математическое и программное обеспечение анализа векторов разнотипных признаков, включающее в себя «детерминационный анализ» как весьма частный случай.

Однако даже на фоне всех остальных рекламных кампаний судьба бутстрепа исключительна. Во-первых, признанный его автор Б. Эфрон с самого начала признавался, что он ничего принципиально нового не сделал. Его исходная статья (первая в сборнике [3]) называлась: «Бутстреп-методы: новый взгляд на методы складного ножа». Тем самым Б. Эфрон честно признавал первенство за М. Кенуем — автором методов «складного ножа». Во-вторых, сразу появились статьи и дискуссии в научных изданиях, публикации рекламного характера, и даже в научно-популярных журналах. Бурные обсуждения на конференциях, спешный выпуск книг. В 1980-е гг. финансовая подоплека всей этой активности, связанная с выбиванием грантов на научную деятельность, содержание учебных заведений и т.п., была мало понятна отечественным специалистам.

В чем основная идея группы методов «размножения выборок», наиболее известным представителем которых является бутстреп?

Пусть дана выборка $x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n$. В вероятностно-статистической теории предполагаем, что это — набор независимых одинаково распределенных случайных величин. Пусть эконометрика интересует некоторая статистика $f_n(x_1, x_2, \ldots, x_n)$. Как изучить ее свойства? Подобными проблемами мы занимались на протяжении всей книги и знаем, насколько это непросто. Идея, которую предложил в 1949 г. М. Кенуй (это и есть «метод складного ножа») состоит в том, чтобы из одной выборки сделать много, исключая из нее по одному наблюдению (и возвращая ранее исключенные). Перечислим выборки, которые получаются из исходной:

$$x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n \,;$$

$$x_1, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n \,;$$

$$x_1, x_2, x_4, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}, x_n \,;$$

$$\ldots$$

$$x_1, x_2, x_3, \ldots x_{k-1}, x_{k+1}, \ldots, x_{n-1}, x_n \,;$$

$$\ldots$$

$$x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-2}, x_n \,;$$

$$x_1, x_2, x_3, \ldots x_{k-1}, x_k, x_{k+1}, \ldots, x_{n-1}.$$

Всего n новых (размноженных) выборок объемом (n – 1) каждая. По каждой из них можно рассчитать значение интересующей эконометрика статистики (с уменьшенным на 1 объемом выборки):

$$f_{n-1,1}(\omega) = f_{n-1}(x_2, x_3, .. x_{k-1}, x_k, x_{k+1}, ..., x_{n-1}, x_n);$$
$$f_{n-1,2}(\omega) = f_{n-1}(x_1, x_3, .. x_{k-1}, x_k, x_{k+1}, ..., x_{n-1}, x_n);$$
$$f_{n-1,3}(\omega) = f_{n-1}(x_1, x_2, x_4, .. x_{k-1}, x_k, x_{k+1}, ..., x_{n-1}, x_n);$$
$$\ldots$$
$$f_{n-1,k}(\omega) = f_{n-1}(x_1, x_2, x_3, .. x_{k-1}, x_{k+1}, ..., x_{n-1}, x_n);$$
$$\ldots$$
$$f_{n-1,n-1}(\omega) = f_{n-1}(x_1, x_2, x_3, .. x_{k-1}, x_k, x_{k+1}, ..., x_{n-2}, x_n);$$
$$f_{n-1,n}(\omega) = f_{n-1}(x_1, x_2, x_3, .. x_{k-1}, x_k, x_{k+1}, ..., x_{n-1}).$$

Полученные значения статистики позволяют судить о ее распределении и о характеристиках распределения — о математическом ожидании, медиане, квантилях, разбросе, среднем квадратическом отклонении. Значения статистики, построенные по размноженным подвыборкам, не являются независимыми. Однако, как мы видели в главе 6 на примере ряда статистик, возникающих в методе наименьших квадратов и в кластер-анализе (при обсуждении возможности объединения двух кластеров), при росте объема выборки влияние зависимости может ослабевать, а потому со значениями статистик типа $f_{n-1,k}(\omega), k = 1,2,...,n,$ можно обращаться как с независимыми случайными величинами (это утверждение строго доказано).

Однако и без всякой вероятностно-статистической теории разброс величин $f_{n-1,k}(\omega), k = 1,2,...,n,$ дает наглядное представление о том, какую точность может дать рассматриваемая статистическая оценка.

Сам М. Кенуй и его последователи использовали размножение выборок в основном для построения оценок с уменьшенным смещением. А вот Б. Эфрон предложил новый способ размножения выборок, существенно использующий датчики псевдослучайных чисел. А именно, он предложил строить новые выборки, *моделируя выборки из эмпирического распределения.* Другими словами, Б. Эфрон предложил взять конечную совокупность из n элементов исходной выборки $x_1, x_2, x_3, .. x_{k-1}, x_k, x_{k+1}, ..., x_{n-1}, x_n$ и с помощью датчика псевдослучайных чисел сформировать из нее любое число размноженных выборок. По сравнению с описанной выше процедурой Кенуя появляются новые недостатки — неизбежные совпадения элементов размноженных выборок и зависимость от качества датчиков псевдослучайных чисел. Однако существует математическая теория, позволяющая (при некоторых предположениях и безграничном росте объема выборки) обосновать процедуры бутстрепа (см. сборник статей [3]).

Есть много способов развития идеи размножения выборок. Можно по исходной выборке построить эмпирическую функцию распределения, а затем каким-либо образом от кусочно-постоянной функции перейти к непрерывной функции распределения, например, соединив точки $\left(x(i);\ \frac{i}{n}\right)$, $i=1,2,\ldots,n$, отрезками прямых. Другой вариант — перейти к непрерывному распределению, построив непараметрическую оценку плотности. После этого рекомендуется брать размноженные выборки из этого непрерывного распределения (являющегося состоятельной оценкой исходного), непрерывность защитит от совпадений элементов в этих выборках.

Другой вариант построения размноженных выборок — более прямой. Исходные данные не могут быть определены совершенно точно и однозначно. Поэтому предлагается к исходным данным добавлять малые независимые одинаково распределенные погрешности. При таком подходе соединяем вместе идеи устойчивости и бутстрепа. При внимательном анализе многие идеи прикладной статистики тесно друг с другом связаны.

В каких случаях целесообразно применять бутстреп, а в каких — другие методы прикладной статистики? В период рекламной кампании встречались, в том числе в научно-популярных журналах, утверждения о том, что и для оценивания математического ожидания полезен бутстреп. Это совершенно не так. При росте числа испытаний методом Монте-Карло бутстреп-оценка приближается к классической оценке — среднему арифметическому результатов наблюдений. Другими словами, бутстреп-оценка отличается от классической оценки только шумом псевдослучайных чисел.

Аналогичной является ситуация и в ряде других случаев. Там, где статистическая теория хорошо развита, где найдены методы анализа данных, в том или ином смысле близкие к оптимальным, бутстрепу делать нечего. А вот в новых областях со сложными алгоритмами, свойства которых недостаточно ясны, он представляет собой ценный инструмент для изучения ситуации.

Компьютерная статистика в контроллинге. В качестве примера применения компьютерной статистики рассмотрим конкретную прикладную область — контроллинг, т.е. современный подход к управлению организацией [8].

Контроллеру и сотрудничающему с ним статистику нужна разнообразная экономическая и управленческая информация, не менее нужны удобные инструменты ее анализа. Следовательно, информационная поддержка контроллинга необходима для успешной работы контроллера. Без современных компьютерных инструментов анализа и управления, основанных на продвинутых эконометрических и экономико-математических методах и моделях, невозможно эффективно принимать управленческие решения. Недаром специалисты по контроллингу большое внимание уделяют проблемам создания, развития и применения компьютерных систем поддержки принятия решений. Высокие статистические технологии и эконометрика — неотъемлемые части любой современной системы поддержки принятия экономических и управленческих решений.

Важная часть прикладной статистики — применение высоких статистических технологий к анализу конкретных экономических данных. Такие исследования зачастую требуют дополнительной теоретической работы по «доводке» статистических технологий применительно к конкретной ситуации. Большое значение для контроллинга имеют не только общие методы, но и конкретные эконометрические модели, например, вероятностно-статистические модели тех или иных процедур экспертных оценок или эконометрики качества, имитационные модели деятельности организации, прогнозирования в условиях риска. И конечно, такие конкретные применения, как расчет и прогнозирование индекса инфляции. Сейчас уже многим специалистам ясно, что годовой, квартальный или месячный бухгалтерский баланс предприятия может быть использован для оценки его финансово-хозяйственной деятельности только с привлечением данных об инфляции. Различные области экономической теории и практики в настоящее время еще далеко не согласованы. При оценке и сравнении инвестиционных проектов принято использовать такие характеристики, как чистая текущая стоимость, внутренняя норма доходности, основанные на введении в рассмотрение изменения стоимости денежной единицы во времени (это осуществляется с помощью дисконтирования). А вот при анализе финансово-хозяйственной деятельности организации на основе данных бухгалтерской отчетности изменение стоимости денежной единицы во времени по традиции не учитывают.

Специалисты по контроллингу должны быть вооружены современными средствами информационной поддержки, в том числе средствами на основе высоких статистических технологий и эконометрики. Очевидно, преподавание должно идти впереди практического применения. Ведь как применять то, чего не знаешь?

Статистические технологии применяют для анализа данных двух принципиально различных типов. Один из них — это результаты измерений (наблюдений, испытаний, анализов, опытов и др.) различных видов, например, результаты управленческого или бухгалтерского учета, данные Госкомстата и др. Короче, речь идет об объективной информации. Другой — это оценки экспертов, на основе своего опыта и интуиции делающих заключения относительно экономических явлений и процессов. Очевидно, это — субъективная информация. В стабильной экономической ситуации, позволяющей рассматривать длинные временные ряды тех или иных экономических величин, полученных в сопоставимых условиях, данные первого типа вполне адекватны. В быстро меняющихся условиях приходятся опираться на экспертные оценки. Такая новейшая часть прикладной статистики, как статистика нечисловых данных, была создана как ответ на запросы теории и практики экспертных оценок.

Для решения каких экономических задач могут быть полезны статистические методы? Практически для всех, использующих конкретную информацию о реальном мире. Только чисто абстрактные, отвлеченные от реальности исследования могут обойтись без нее. В частности, статистические методы необходима для прогнозирования, в том числе поведения потребителей, а потому и для планирования. Выборочные исследования, в том числе выборочный контроль, основаны на статистические методы. Но планирование и контроль — основа контроллинга. Поэтому статистические методы — важная составляющая инструментария контроллера, воплощенного в компьютерной системе поддержки принятия решений. Прежде всего оптимальных решений, которые предполагают опору на адекватные модели прикладной статистики. В производственном менеджменте это может означать, например, использование моделей экстремального планирования эксперимента (судя по накопленному опыту их практического использования, такие модели позволяют повысить выход полезного продукта на 30–300 %).

Высокие статистические технологии предполагают адаптацию применяемых методов к меняющейся ситуации. Например, параметры прогностического индекса меняются вслед за изменением характеристик используемых для прогнозирования величин. Таков метод экспоненциального сглаживания. В соответствующем алгоритме расчетов значения временного ряда используются с весами. Веса уменьшаются по мере удаления в прошлое. Многие методы дискриминантного анализа основаны на применении обучающих выборок. Например, для построения рейтинга надежности банков можно с помощью экспертов составить две обучающие выборки — надежных и ненадежных банков. А затем с их помощью решать для вновь рассматриваемого банка, каков он — надежный или ненадежный, а также оценивать его надежность численно, т.е. вычислять значение рейтинга.

Один из способов построения адаптивных статистических моделей — нейронные сети (см., например, монографию [39]). При использовании нейронных сетей упор делается не на формулировку адаптивных алгоритмов анализа данных, а — в большинстве случаев — на построение виртуальной адаптивной структуры. Термин «виртуальная» означает, что «нейронная сеть» — это специализированная компьютерная программа, «нейроны» используются лишь при общении человека с компьютером. Методология нейронных сетей идет от идей кибернетики 1940–1950-х годов. В компьютере создается модель мозга человека (весьма примитивная с точки зрения физиолога). Основа модели — весьма простые базовые элементы, называемые нейронами. Они соединены между собой, так что нейронные сети можно сравнить с хорошо знакомыми экономистам и инженерам блок-схемами. Каждый нейрон находится в одном из заданного множества состояний. Он получает импульсы от соседей по сети, изменяет свое состояние и сам рассылает импульсы. В результате состояние множества нейтронов изменяется, что соответствует проведению статистических вычислений.

Нейроны обычно объединяются в слои (как правило, два-три). Среди них выделяются входной и выходной слои. Перед началом решения той или иной задачи производится настройка. Во-первых, устанавливаются связи между нейронами, соответствующие решаемой задаче. Во-вторых, проводится обучение, т.е. через нейронную сеть пропускаются обучающие выборки, для элементов которых требуемые результаты расчетов известны. Затем параметры сети модифицируются так, чтобы получить максимальное соответствие выходных значений заданным величинам.

С точки зрения точности расчетов (и оптимальности в том или ином статистическом смысле) нейронные сети не имеют преимуществ перед другими адаптивными системами прикладной статистики. Однако они более просты для восприятия. Надо отметить, что в прикладной статистике используются и модели, промежуточные между нейронными сетями и «бычными» системами регрессионных уравнений (одновременных и с лагами). Они тоже используют блок-схемы, как, например, универсальный метод моделирования связей социально-экономических факторов ЖОК (этот метод описан в [7]).

Профессионалу в области контроллинга полезны многочисленные интеллектуальные инструменты анализа данных, относящиеся к высоким статистическим технологиям и эконометрике. В частности, заметное место в математико-компьютерном обеспечении принятия решений в контроллинге занимают методы теории нечеткости.

16.4. О методологии статистических методов

При разработке и применении статистических методов необходимо опираться на четкие методологические принципы, разработанные поколениями специалистов. Рассмотрим некоторые из них.

Задача — модель — метод — условия применимости. Разработка и применение статистических методов, в том числе прикладной статистики, предполагает последовательное осуществление трех этапов исследования. Первый — от исходной практической проблемы до теоретической чисто математической задачи. Второй — внутриматематическое изучение и решение этой задачи. Третий — переход от математических выводов обратно к практической проблеме.

В литературе вопросы методологии статистических методов обсуждаются явно недостаточно. Зато наблюдается поток публикаций, в которых постановки решаемых задач иногда выглядят весьма искусственно. Цель настоящего раздела — обосновать необходимость развития методологии статистических методов как самостоятельного научного направления, рассмотреть ряд проблем, относящихся к этому направлению.

В области моделирования задач прикладной статистики, как, впрочем, и в иных областях применения математики и кибернетики, целесообразно выделять четверки проблем:

ЗАДАЧА — МОДЕЛЬ — МЕТОД — УСЛОВИЯ ПРИМЕНИМОСТИ.

Обсудим каждую из только что выделенных составляющих.

Задача, как правило, порождена потребностями той или иной прикладной области. Вполне понятно, что при этом происходит одна из возможных математических формализаций реальной ситуации. Например, при изучении предпочтений потребителей у экономистов — маркетологов возникает вопрос: различаются ли мнения двух групп потребителей (см. главы 1 и 5). При математической формализации мнения потребителей в каждой группе обычно моделируются как независимые случайные выборки, т.е. как совокупности независимых одинаково распределенных случайных величин, а вопрос маркетологов переходит в рамках этой модели в вопрос о проверке той или иной статистической гипотезы однородности. Речь может идти об однородности характеристик, например, о проверке равенства математических ожиданий, или о полной (абсолютной однородности), т.е. о совпадении функций распределения, соответствующих двух совокупностям (см. главу 5).

Задача может быть порождена также обобщением потребностей ряда прикладных областей. Приведенный выше пример иллюстрирует эту ситуацию: к необходимости проверки гипотезы однородности приходят и медики при сравнении двух групп пациентов, и инженеры при сопоставлении результатов обработки деталей двумя способами, и т.д. Таким образом, одна и та же математическая модель может применяться для решения самых разных по своей прикладной сущности задач.

Важно подчеркнуть, что выделение перечня задач находится вне математики. Выражаясь инженерным языком, этот перечень является сутью технического задания, которое специалисты различных областей деятельности дают статистикам. Чтобы математик мог решать практическую задачу, необходимо ее суть выразить в математических терминах, т.е. построить математическую модель.

Метод, используемый в рамках определенной математической модели — это уже во многом, если не в основном, дело математиков. В моделях прикладной статистики речь идет, например, о методе оценивания, о методе проверки гипотезы, о методе доказательства той или иной теоремы, и т.д. В двух первых случаях алгоритмы разрабатываются и исследуются математиками, но используются прикладниками, в то время как метод доказательства касается лишь самих математиков.

Ясно, что для решения той или иной задачи в рамках одной и той же принятой исследователем модели может быть предложено много методов. Приведем примеры. Для специалистов по теории вероятностей и математической статистике наиболее хорошо известна история Центральной Предельной Теоремы теории вероятностей. Предельный нормальный закон был получен многими разными методами, из которых напомним теорему Муавра-Лапласа, метод моментов Чебышева, метод характеристических функций Ляпунова, завершающие эпопею методы, примененные Линдебергом и Феллером. В настоящее время для решения практически важных задач могут быть использованы современные информационные технологии на основе метода статистических испытаний и соответствующих датчиков псевдослучайных чисел. Они уже заметно потеснили асимптотические методы математической статистики. В рассмотренной выше проблеме однородности для проверки одной и той же гипотезы совпадения функций распределения могут быть применены самые разные методы — Смирнова, Лемана — Розенблатта, Вилкоксона и др. (см. главу 5).

Наконец, рассмотрим последний элемент четверки — условия применимости. Он — полностью внутриматематический. С точки зрения математика замена условия (кусочной) дифференцируемости некоторой функции на условие ее непрерывности может представляться существенным научным достижением, в то время как прикладник оценить это достижение не сможет. Для него, как и во времена Ньютона и Лейбница, непрерывные функции мало отличаются от (кусочно) дифференцируемых функций. Точнее, они одинаково хорошо (или одинаково плохо) могут быть использованы для описания реальной действительности.

Точно также прикладник не сможет оценить внутриматематическое достижение, состоящее в переходе от условия конечности четвертого момента случайной величины к условию конечности дисперсии. Поскольку результаты реальных измерений получены с помощью некоторого прибора (средства измерения), шкала которого конечна, то прикладник априори уверен, что все результаты измерений заведомо лежат на некотором отрезке (т.е. финитны). Он с некоторым недоумением наблюдает за математиком, который рассуждает о конечности тех или иных моментов — для прикладника они заведомо конечны.

Математики и прикладники. Таким образом, в настоящее время наблюдается значительное расхождение интересов «типового» математика и «типового» прикладника. Конечно, мы рассуждаем здесь, строя гипотетические модели восприятия и поведения того и другого. Опишем эти модели более подробно.

Прикладник заинтересован в научно обоснованном решении стоящих перед ним реальных задач. При этом при формализации задач он готов принять достаточно сильные математические предположения. Например, с точки зрения прикладника случайные величины могут принимать конечное множество значений, или быть финитными, или иметь нужное математику число моментов, и т.д. Как говорил А.Н. Колмогоров, переход от дискретности к непрерывности для прикладника оправдан только тогда, когда этот переход облегчает выкладки и расчеты, как в математическом анализе переход от сумм к интегралам облегчает рассуждения и вычисления. Если же при переходе к непрерывности возникают сложности типа необходимости доказательства измеримости тех или иных величин относительно тех или иных сигма-алгебр, то прикладник готов вернуться к постановке задачи с конечным вероятностным пространством.

Здесь уместно напомнить, что один из выдающихся вероятностников XX в. В. Феллер выпустил свой учебник по теории вероятностей в двух книгах, посвятив первую дискретным вероятностным пространствам, а вторую — непрерывным.

Другой пример — задачи оптимизации. Если оптимизация проводится по конечному множеству, то оптимум всегда достигается (хотя может быть не единственным). Если же множество параметров бесконечно, то задача оптимизации может и не иметь решения. Поэтому у прикладника есть стимул ограничиться математическими моделями с конечным множеством параметров. Напомним в связи с этим, что основные задачи прикладной статистики допускают оптимизационную постановку, а статистика объектов нечисловой природы как целое построена на решении оптимизационных задач (а не на суммировании тех или иных выражений, поскольку в пространствах объектов нечисловой природы нет операции сложения).

Модель поведения типового математика совершенно иная. Он, как правило, не обдумывает реальные задачи, поскольку не вникает в конкретные прикладные области. (Если же вникает, то является уже не только математиком, но и прикладником, и его поведение промоделировано в предыдущих абзацах.) Математик берет те задачи, которые уже ранее рассматривались, и старается получить для них математически интересные результаты. Зачастую это означает борьбу за ослабление математических условий, при которых были получены предыдущие результаты. При этом математика абсолютно не волнует, имеют ли какое-либо реальное содержание доказанные им теоремы, могут ли они принести какую-либо пользу прикладнику. Его интересует реакция математической общественности, а не реакция прикладников.

Сколько реально используется чисел? Для демонстрации разрыва между математиками и прикладниками обратим внимание на два парадокса.

Все реальные результаты наблюдений записываются рациональными числами (обычно десятичными числами с небольшим — от 2 до 5 — числом значащих цифр). Как известно, множество рациональных чисел счетно, а потому вероятность попадания значения непрерывной случайной величины в него равно 0. Следовательно, все рассуждения, связанные с моделированием непрерывными случайными величинами реальных результатов наблюдений — это рассуждения о том, что происходит внутри множества меры 0. Первый парадокс состоит в том, что множествами меры 0 в теории вероятностей принято пренебрегать. Другими словами, в точки зрения теории вероятностей всеми реальными данными можно пренебречь, поскольку они входят в одно фиксированное множество меры 0.

Глубже проанализируем ситуацию. Сколько всего чисел используется для записи реальных результатов наблюдений? Речь идет о типовых результатах наблюдений, измерений, испытаний, опытов, анализов. Они используются в технических, естественнонаучных, экономических, социологических, медицинских и иных исследованиях. Анализ практики показывает, что эти числа имеют вид $(a{,}bcde)10^k$. Здесь a принимает значения от 1 до 9, а стоящие после запятой b, c, d, e — от 0 до 9. В то же время показатель степени k меняется от (-100) до $+100$. Ясно, что общее количество возможных чисел равно $9\times10^4\times201$=18 090 000, т.е. меньше 20 миллионов.

Итак, второй парадокс, усиливающий первый, состоит в том, что для описания реальных результатов наблюдений вполне достаточно 20 миллионов отдельных символов. Бесконечность натурального ряда и континуум числовой прямой — это математические абстракции, надстроенные над дискретной и состоящей из конечного числа элементов реальностью. (При изменении числа значащих цифр, используемых для описания результатов наблюдений, принципиальный вывод не меняется.) Таким образом, реальные данные лежат не только во множестве меры 0, но и в конечном множестве, причем число элементов в этом множестве вполне обозримо.

Практические следствия методологии прикладной статистики. Из сказанного вытекают некоторые вполне определенные выводы, в том числе касающиеся преподавания и научных исследований.

Например, преподавание теории вероятностей может быть сосредоточено на случае конечного вероятностного пространства. Бесконечные вероятностные пространства могут при этом рассматриваться как удобные математические схемы. Их роль — давать возможность более легко и быстро получать полезные утверждения для конечных вероятностных пространств. Из сказанного вытекает, в частности, что различные параметрические семейства распределений (семейства нормальных, логарифмически нормальных, экспоненциальных, Коши, Вейбулла-Гнеденко, гамма-распределений) приобретают статус не более чем удобных приближений для распределений на конечных вероятностных пространствах. При таком подходе теряет свою парадоксальность тот эмпирически не раз проверенный факт, что распределение погрешностей измерений, как правило, не является гауссовым (см. главу 2.1).

В качестве другого примера рассмотрим методы оценивания параметров. По традиции много внимания в учебных курсах уделяется оценкам максимального правдоподобия (ОМП). Однако столь же хорошие асимптотические свойства имеют т.н. одношаговые оценки, гораздо более простые с вычислительной точки зрения (см. главу 3.2). Целесообразно их включить в учебные курсы, а ОМП исключить.

Целесообразно уделять внимание (репрезентативной) теории измерений, в частности, концепции шкал измерения. Необходимо знакомство с определениями и основными свойствами шкал наименований, порядковой, интервалов, отношений, разностей, абсолютной. Установлено, какими алгоритмами статистического анализа данных можно пользоваться в той или иной шкале, в частности, для усреднения результатов наблюдений. Так, для данных, измеренных в порядковой шкале, некорректно вычислять среднее арифметическое. В качестве средних величин для таких данных можно использовать порядковые статистики, в частности, медиану.

Статистические методы исследования часто опираются на использование современных информационных технологий. В частности, распределение статистики можно находить методами асимптотической математической статистики, а можно и путем статистического моделирования (метод Монте-Карло, он же — метод статистических испытаний).

Методологический анализ — первый этап моделирования задач принятия решений, да и вообще любого исследования. Он определяет исходные постановки для теоретической проработки, а потому во многом и успех всего исследования.

Методологический анализ — первый этап статистического исследования. Он определяет исходные постановки для теоретической проработки, а потому во многом и успех всего исследования [40]. Этот этап — один из наиболее важных [41]. Подчеркнем, что анализ динамики развития методов прикладной статистики выделить наиболее перспективные методы. В частности, в работе [42] установлено, что в настоящее время наиболее перспективными являются методы нечисловой статистики. Именно поэтому им уделено большое внимание в учебнике [7].

16.5. Основные нерешенные проблемы статистических методов

За последние тридцать лет выявился целый ряд нерешенных проблем статистических методов, как чисто научных, так и научно-организационных. Обсудим пять из них:

- влияние отклонений от традиционных предпосылок вероятностно-статистических моделей на свойства статистических процедур;

- оправданность использования асимптотических теоретических результатов при конечных объемах выборок;

- формулировки и обоснования правил выбора одного из многих критериев для проверки конкретных гипотез;

- конкретные способы организации теоретических работ в области статистических методов;

- организация и проведение прикладных работ с использованием статистических методов.

Приводимые ниже соображения отнюдь не претендует на решение перечисленных проблем. Их цель гораздо скромнее — обратить внимание на существование ряда нерешенных проблем в надежде, что коллективными усилиями удастся продвинуться в их решении.

Влияние отклонений от традиционных предпосылок. В вероятностной теории статистических методов выборка обычно моделируется как конечная последовательность независимых одинаково распределенных случайных величин или векторов. Часто предполагается, что эти величины или вектора имеют нормальное распределение.

На основе сформулированных классических предпосылок построено огромное здание классической математической статистики с большим числом теорем. Оно за последнее столетие обросло горой учебников и программных продуктов.

Однако при внимательном взгляде совершенно ясна нереалистичность классических предпосылок. Независимость результатов измерений обычно принимается «из общих предположений», между тем во многих случаях очевидна их коррелированность [43]. Одинаковая распределенность результатов измерений также вызывает сомнения из-за изменения во времени свойств измеряемых образцов, средств измерения и психофизического состояния специалистов, проводящих измерения (наблюдения, испытания, анализы, опыты). Даже обоснованность самой возможности применения вероятностных моделей также часто вызывает сомнения, например, при моделировании уникальных измерений (теорию вероятностей обычно привлекают при изучении массовых явлений). И уж совсем редко распределения результатов измерений можно считать нормальными (см. главу 2.1).

Итак, методы классической математической статистики обычно используют вне сферы их обоснованной применимости. Каково влияние отклонений от традиционных предпосылок на статистические выводы? В настоящее время об этом имеются лишь отрывочные сведения. Приведем три примера.

Пример 1. Построение доверительного интервала для математического ожидания обычно проводят с использованием распределения Стьюдента (при справедливости гипотезы нормальности). Как следует из Центральной Предельной Теоремы (ЦПТ) теории вероятностей, в асимптотике (т.е. при большом объеме выборки) такие расчетные методы дают правильные результаты. А именно, из ЦПТ вытекает использование квантилей нормального распределения, а из классической теории — квантилей распределения Стьюдента, но при росте объема выборки квантили распределения Стьюдента стремятся к соответствующим квантилям нормального распределения.

Пример 2. Для проверки однородности двух независимых выборок (на самом деле — для проверки равенства математических ожиданий) обычно рекомендуют использовать двухвыборочный критерий Стьюдента. Что будет при отклонении от нормальности распределений, из которых взяты выборки? Если объемы выборок равны или если дисперсии результатов наблюдений в выборках совпадают, то в асимптотике (когда объемы выборок безгранично возрастают) классический метод является корректным. Если же объемы выборок существенно отличаются и их дисперсии различны, то двухвыборочную статистику Стьюдента применять нельзя. Поскольку проверка равенства дисперсий — более сложная задача, чем проверка равенства математических ожиданий, то для выборок разного объема использовать двухвыборочную статистику Стьюдента не следует, лучше применять критерий Крамера-Уэлча, как это подробно обосновано в главе 5.

Пример 3. В задаче отбраковки (исключения) резко выделяющихся наблюдений (выбросов) расчетные методы, основанные на нормальности, являются крайне неустойчивыми по отношению к отклонениям от нормальности, что полностью лишает эти методы научной обоснованности (подробнее см. раздел 4.2).

Примеры 1–3 показывают весь спектр возможных свойств классических расчетных методов в случае отклонения от нормальности. Методы примера 1 оказываются вполне пригодными при таких отклонениях, примера 2 — пригодны в некоторых случаях, примера 3 — полностью непригодны.

Итак, имеется *необходимость изучения свойств расчетных методов классической математической статистики, опирающихся на предположение нормальности, в ситуациях, когда это предположение не выполнено.* Аппаратом для такого изучения наряду с методом Монте-Карло (статистических испытаний) могут послужить предельные теоремы теории вероятностей (и опирающиеся на них асимптотические методы математической статистики), прежде всего ЦПТ, поскольку интересующие нас расчетные методы обычно используют разнообразные суммы.

Пока подобное изучение не проведено, остается неясной научная ценность, например, применения факторного анализа к векторам из переменных, принимающих небольшое число градаций и к тому же измеренных в порядковой шкале. Этот пример показывает важность еще одного направления исследований — изучения свойств алгоритмов, предназначенных для анализа числовых данных, в случаях, когда данные измерены в шкалах, отличных от абсолютной, в частности, в порядковой шкале.

Из большого числа возможных постановок, относящихся к изучению влияния отклонений от традиционных предпосылок, укажем лишь на то, что реальные данные имеют небольшое число значащих цифр (обычно от 2 до 5), в то время как в классической математической статистике используются непрерывные случайные величины, для которых вероятность получения подобного результата наблюдения равна 0. Действительно, вероятность того, что хотя бы один элемент выборки из распределения с непрерывной функцией распределение попадет в заданное счетное множество, в частности, в множество рациональных чисел, равна 0 (согласно классическим свойствам вероятностной меры). Событиями, имеющими вероятность 0, принято пренебрегать. Следовательно, с точки зрения классической математической статистики любыми реальными данными нужно пренебречь! Выходов из этого парадокса несколько. Один из них — бурно развивающаяся в настоящее время статистика интервальных данных (см. главу 9 учебника [7]), другой — использование классических поправок Шеппарда для сгруппированных данных [44, 45]. Здесь еще много работы. Так, даже для такого широко используемого статистического показателя, как коэффициент корреляции, поправки на группировку (поправки Шеппарда) были получены сравнительно недавно — лишь в 1980 г. [46] (см. раздел 3.5).

Почему на первый план выдвинуто изучение классических алгоритмов, а не построение новых, специально предназначенных для работы в условиях отклонения от классических предпосылок? Во-первых, потому, что классические алгоритмы в настоящее время наиболее распространены (благодаря сложившейся системе образования как прикладников, так и математиков). Во-вторых, более новые подходы зачастую методологически уязвимы. Так, известная робастная модель засорения Тьюки — Хубера (см. раздел 3.4) нацелена на борьбу с большими выбросами, которые зачастую физически невозможны из-за ограниченности интервала возможных значений измеряемой характеристики, в котором работает конкретное средство измерения. Следовательно, модель Тьюки — Хубера имеет скорее теоретическое значение, чем практическое. Сказанное, конечно, не означает, что следует прекратить разработку, изучение и внедрение непараметрических и устойчивых методов, выделенных выше как «точки роста» современных эконометрики и прикладной статистики.

Использование асимптотических результатов при конечных объемах выборок. Как отмечено выше, изучение классических алгоритмов во многих случаях может быть проведено с помощью асимптотических методов математической статистики, в частности, с помощью ЦПТ и методов наследования сходимости (см. главу 14). Отрыв классической математической статистики от нужд прикладных исследований проявился, в частности, в том, что в распространенных монографиях недостает математического аппарата, необходимого, в частности, для изучения двухвыборочных статистик. Суть в том, что переходить к пределу приходится не по одному параметру, а по двум — объемам двух выборок. Пришлось разработать соответствующую теорию — теорию наследования сходимости, впервые изложенную в монографии [3, п. 2.4].

Однако применять результаты подобного изучения придется при конечных объемах выборок. Возникает целый букет проблем, связанных с таким переходом. Часть из них обсуждалась в главе 14.6 в связи с изучением свойств статистик, построенных по выборкам из конкретных распределений.

Однако при обсуждении влияния отклонений от исходных предположений на свойства статистических процедур возникают дополнительные проблемы. Какие отклонения считать типичными? Ориентироваться ли на наиболее «вредные» отклонения, в наибольшей степени искажающие свойства алгоритмов, или же сосредоточить внимание на «типичных» отклонениях?

При первом подходе получаем гарантированный результат, но «цена» этого результата может быть излишне высокой. В качестве примера укажем на универсальное неравенство Берри — Эссеена для погрешности в ЦПТ [47, 48]. Совершенно справедливо подчеркивает академик РАН А.А. Боровков [48, с. 172], что «скорость сходимости в реальных задачах, как правило, оказывается лучше».

При втором подходе возникает вопрос, какие отклонения считать «типичными». Попытаться ответить на этот вопрос можно, анализируя большие массивы реальных данных. Вполне естественно, что ответы различных исследовательских групп будут различаться.

Одна из ложных идей — использование при анализе возможных отклонений только какого-либо конкретного параметрического семейства. Например, семейств распределений Вейбулла — Гнеденко, экспоненциальных, нормальных, трехпараметрического семейства гамма — распределений и др. Как уже отмечалось во введении к настоящему учебнику, еще в 1927 г. акад. АН СССР С.Н. Бернштейн обсуждал методологическую ошибку, состоящую в сведении всех эмпирических распределений к четырехпараметрическому семейству Пирсона [49]. Однако и до сих пор параметрические методы статистики весьма популярны, особенно среди прикладников, и вина за это заблуждение лежит, прежде всего, на преподавателях статистических методов.

Выбор одного из многих критериев для проверки конкретной гипотезы. Во многих случаях для решения конкретной практической задачи разработано много методов, и специалист по прикладной статистике стоит перед проблемой: какой из них предложить прикладнику для анализа конкретных данных?

В качестве примера рассмотрим задачу проверки однородности двух независимых выборок. Как известно (см. главу 5), для ее решения можно предложить массу критериев: Стьюдента, Крамера-Уэлча, Лорда, хи-квадрат, Вилкоксона (Манна — Уитни), Ван-дер-Вардена, Сэвиджа, Н.В. Смирнова, типа омега-квадрат (Лемана — Розенблатта), Г.В. Мартынова и др. Какой из них выбрать?

Естественным образом приходит в голову идея «голосования»: провести проверку по многим критериям, а затем принять решение «по большинству голосов». С точки зрения статистической теории такая процедура приводит попросту к построению еще одного критерия, который априори ничем не лучше прежних (но и не хуже), но более труден для изучения. С другой стороны, если совпадают решения по всем рассмотренным статистическим критериям, исходящим из различных принципов, то в соответствии с концепцией устойчивости, впервые развитой в монографии [3] (см. также главу 14.6), это повышает доверие к полученному общему решению.

Распространено, особенно среди математиков, ложное и вредное мнение о необходимости поиска оптимальных методов, решений и т.д. Дело в том, что оптимальность обычно исчезает при отклонении от исходных предпосылок. Так, среднее арифметическое в качестве оценки математического ожидания является оптимальной оценкой тогда и только тогда, когда исходное распределение — нормальное (см., например, монографию [50]), в то время как состоятельной оценкой — всегда, лишь бы математическое ожидание существовало. С другой стороны, для любого произвольно взятого метода оценивания или проверки гипотез обычно можно так сформулировать понятие оптимальности, чтобы рассматриваемый метод стал оптимальным — с этой специально выбранной точки зрения. Возьмем, например, выборочную медиану как оценку математического ожидания. Она, разумеется, оптимальна, хотя и в другом смысле, чем среднее арифметическое (оптимальное для нормального распределения). А именно, для распределения Лапласа выборочная медиана является оценкой максимального правдоподобия, а потому оптимальной — в том смысле, в каком оптимальной является любая оценка максимального правдоподобия. Соответствующее понятие оптимальности требует аккуратных формулировок, оно строго изложено в монографии [51]. Как известно, оценки максимального правдоподобия удобны при теоретических рассмотрениях, а при анализе конкретных экономических, технических и иных данных следует применять одношаговые оценки (см. об этом раздел 3.2).

Проиллюстрируем сказанное примером. Критерии однородности двух выборок были проанализированы в монографии [52]. Естественных подходов к сравнению критериев несколько — на основе асимптотической относительной эффективности по Бахадуру, Ходжесу — Леману, Питмену и др. И выяснилось, что каждый обычно используемый критерий однородности является оптимальным при соответствующей альтернативе или подходящем распределении на множестве альтернатив. При этом математические рассуждения обычно опираются на альтернативу сдвига, сравнительно редко встречающуюся в практике анализа реальных статистических данных (в связи с критерием Вилкоксона эта альтернатива обсуждалась в главе 5.3). Итог печален — блестящая математическая техника, продемонстрированная в монографии [52], не позволяет дать рекомендации для выбора критерия проверки однородности при анализе реальных данных. Другими словами, с точки зрения работы прикладника, т.е. с точки зрения применимости полученных результатов при анализе конкретных данных, монография [52] бесполезна. Блестящее владение математикой и огромное трудолюбие, продемонстрированные автором этой монографии, увы, ничего не принесли практике.

Конечно, каждый практически работающий статистик так или иначе решает для себя проблему выбора статистического критерия. На основе ряда методологических соображений в главе 5 мы остановили свой выбор на состоятельном против любой альтернативы критерии типа омега-квадрат (Лемана — Розенблатта). Однако остается чувство неудовлетворенности в связи с недостаточной теоретической обоснованностью этого выбора.

Организация теоретических работ в области статистических методов. Выше продемонстрирована необходимость большой теоретической работы по развитию нацеленных на практическое использование статистических методов. В статье [53] 1992 г. обоснован вывод о необходимости создания сети научно-исследовательских организаций, которая выполняла бы такую работу. Как известно, количество научных работников к настоящему времени сократилось в несколько раза по сравнению с началом 1990-х годов. Так что на осуществление в ближайшие годы сформулированной в [53] научно-организационной программы надеяться не приходится.

Приходится с сожалением констатировать, что в рамках научной специальности «теория вероятностей и математическая статистика» наблюдается четко выраженное игнорирование проблем статистического анализа реальных данных и уход в глубь узкоматематических исследований, которые заведомо ничего не могут дать практике. Причины этого явления, типичного для математических дисциплин, обсуждались во введении к настоящему учебнику. Поэтому нет оснований ожидать, что при «естественном ходе событий» будут получены существенные продвижения в рассмотренных выше нерешенных проблемах статистических методов.

Помочь может выделение государственными структурами системы грантов, направленных на поддержку работ в области нерешенных проблем прикладной статистики. Принципиальным шагом явилось бы официальное выделение государственными органами «статистических методов» как самостоятельного научного направления. Отличного как от чисто математических дисциплин типа «теории вероятностей и математической статистики», так и от, например, ветви экономической теории, известной в официальных кругах под названием «статистика».

О прикладных работах с использованием статистических методов. Проблемы организации теоретических работ в области статистических методов лишь в перспективе важны для практической работы. Как правило, те, кто обрабатывает реальные данные, недостаточно знакомы с теоретическими основами алгоритмов и тем более не следят за событиями «на переднем крае» обсуждаемой научно-практической дисциплины. Это вполне естественно, поскольку основная специальность у таких специалистов — иная.

Несколько огрубляя, можно сказать, что реально используется только то, что имеется в учебниках и справочниках, в широко распространенных программных продуктах, а научные публикации с точки зрения прикладника представляют собой «информационный шум». Ситуация усугубляется традиционным ненормальным положением в отечественной статистике [24].

К сожалению, учебная и научная литература на русском языке (как, впрочем, и на иных языках) по статистическим методам далека от совершенства, переполнена устаревшими методологическими подходами и прямыми ошибками.

До сих пор наилучшим изданием остаются «Таблицы математической статистики» Л.Н. Большева и Н.В. Смирнова [1], созданные еще в 1960-х гг.

Хотя студенты почти всех специальностей изучают в конце курса высшей математики раздел «теория вероятностей и математическая статистика», реально они знакомятся лишь с некоторыми основными понятиями и результатами, которых недостаточно для практической работы. С некоторыми математическими методами исследования студенты встречаются в специальных курсах (например, таких, как «Прогнозирование и технико-экономическое планирование», «Технико-экономический анализ», «Контроль качества продукции», «Маркетинг», «Контроллинг, Математические методы прогнозирования», «Статистика» и др. — в случае студентов экономических специальностей), однако изложение в большинстве случаев носит весьма сокращенный и рецептурный характер. В результате подавляющую часть специалистов по прикладной статистике следует считать самоучками.

Поэтому большое значение имеет введение в технических вузах курсов «Высокие статистические технологии», «Статистические методы» и «Прикладная статистика», а на экономических факультетах таких вузов и в экономических вузах — курса «Эконометрика», поскольку эконометрика — это, как известно, статистический анализ конкретных экономических данных (см. [7]). Естественно, что курсы «Высокие статистические технологии», «Статистические методы», «Прикладная статистика» и «Эконометрика» должны быть обеспечены соответствующими учебниками и учебными пособиями, методическими материалами и обучающими компьютерными системами.

Только через систему образования можно поднять уровень массового применения статистических методов и сократить отставание от «переднего края» теории. А это отставание в настоящее время составляет не менее 20 (но и не более 100) лет [54].

Литература

1. *Большев, Л.Н.* Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. — Москва : Наука, 1965. — 2-е изд. — 1968. — 3-е изд. — 1983.

2. *Орлов, А.И.* О критериях Колмогорова и Смирнова / А.И. Орлов // Заводская лаборатория. — 1995. — Т. 61. — № 7. — С. 59–61.

3. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука, 1979. — 296 с.

4. *Смоляк, С.А.* Устойчивые методы оценивания: Статистическая обработка неоднородных совокупностей / С.А. Смоляк, Б.П. Титаренко. — Москва : Статистика, 1980. — 208 с.

5. *Эфрон, Б.* Нетрадиционные методы многомерного статистического анализа / Б. Эфрон. — Москва : Финансы и статистика, 1988. — 263 с.

6. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения = Applied statistics. Regulations for determinations of estimates and confidence limits for parameters of gamma distribution : национальные стандарт Союза ССР : издание официальное : утвержден и введен в действие постановлением Государственного комитета СССР по стандартам от 27 июня 1983 г. / А.И. Орлов, Н.Г. Миронова, М.Б. Невельсон и др. — Москва : Изд-во стандартов, 1984. — 53 с. (В настоящее время отменен как нормативный документ, но может использоваться как научная публикация.)

7. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 2-е изд., испр. и доп. — Москва : Экзамен, 2003. — 576 с.

8. *Орлов, А.И.* Менеджмент в техносфере : учебное пособие для студентов высших учебных заведений / А.И. Орлов, В.Н. Федосеев. — Москва : Академия, 2003. — 384 с.

9. *Суппес, П.* Основы теории измерений / П. Суппес, Дж. Зинес // Психологические измерения. — Москва : Мир, 1967. — С. 9–110.

10. *Пфанцагль, И.* Теория измерений / И. Пфанцагль. — Москва : Мир, 1976. — 166 с.

11. *Заде, Л.* Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. Заде. — Москва : Мир, 1976. — 168 с.

12. *Дэвид, Г.* Метод парных сравнений / Г. Дэвид. — Москва : Статистика, 1978. — 144 с.

13. *Матерон, Ж.* Случайные множества и интегральная геометрия / Ж. Матерон. — Москва : Мир, 1978. — 318 с.

14. *Терехина, А.Ю.* Анализ данных методами многомерного шкалирования / А.Ю. Терехина. — Москва : Наука, 1986. — 168 с.

15. *Перекрест, В.Т.* Нелинейный типологический анализ социально-экономической информации: Математические и вычислительные методы / В.Т. Перекрест. — Ленинград : Наука, 1983. — 176 с.

16. *Кемени, Дж.* Кибернетическое моделирование: Некоторые приложения / Дж. Кемени, Дж. Снелл. — Москва : Советское радио, 1972. — 192 с.

17. Анализ нечисловой информации / Ю.Н. Тюрин, Б.Г. Литвак, А.И. Орлов [и др.]. — Москва : Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1981. — 80 с.

18. *Литвак, Б.Г.* Экспертная информация: Методы получения и анализа / Б.Г. Литвак. — Москва : Радио и связь, 1982. — 184 с.

19. *Орлов, А.И.* Статистика объектов нечисловой природы и экспертные оценки / А.И. Орлов // Экспертные оценки. Вопросы кибернетики. — Вып. 58. — Москва : Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1979. — С. 17–33.

20. Анализ нечисловой информации в социологических исследованиях / под редакцией В.Г. Андреенкова, А.И.Орлова, Ю.Н. Толстовой. — Москва : Наука, 1985. — 220 с.

21. *Орлов, А.И.* Асимптотическое поведение статистик интегрального типа / А.И. Орлов // Доклады АН СССР. — 1974. — Т. 219. — № 4. — С. 808–811.

22. *Орлов, А.И.* Асимптотическое поведение статистик интегрального типа / А.И. Орлов // Вероятностные процессы и их приложения : межвузовский сборник. — Москва : МИЭМ, 1989. — С. 118–123.

23. *Горский, В.Г.* Современные статистические методы обработки и планирования экспериментов в химической технологии / В.Г. Горский // Инженерно-химическая наука для передовых технологий. Международная школа повышения квалификации Труды третьей сессии. 26–30 мая 1997, Казань, Россия / под редакцией В.А. Махлина. — Москва : Научно-исследовательский физико-химический институт им. Карпова, 1997. — С. 261–293.

24. *Орлов, А.И.* О перестройке статистической науки и ее применений / А.И. Орлов // Вестник статистики. — 1990. — № 1. — С. 65–71.

25. *Плошко, Б.Г.* История статистики : учебное пособие / Б.Г. Плошко, И.И. Елисеева. — Москва : Финансы и статистика. 1990. — 295 с.

26. Вероятность и математическая статистика : энциклопедия / главный редактор Ю.В. Прохоров. — Москва : Большая Российская энциклопедия, 1999. — 910 с.

27. *Орлов, А.И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат / А.И. Орлов // Заводская лаборатория. — 1985. — Т. 51. — № 1. — С. 60–62.

28. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях / А.М. Карминский, Н.И. Оленев, А.Г. Примак, С.Г. Фалько. — Москва : Финансы и статистика, 1998. — 256 с.

29. *Орлов, А.И.* Задачи оптимизации и нечеткие переменные / А.И. Орлов. — Москва : Знание, 1980. — 64 с.

30. The teaching of statistics // Studies in mathematics education. — Vol. 7. — 1989. — 258 p.

31. *Ермаков, С.М.* Метод Монте-Карло и смежные вопросы / С.М. Ермаков. — Москва : Наука, 1975. — 471 с.

32. *Ермаков, С.М.* Статистическое моделирование / С.М. Ермаков, Г.А. Михайлов. — Москва : Наука, 1982. — 296 с.

33. *Иванова, И.М.* Случайные числа и их применения / И.М. Иванова. — Москва : Финансы и статистика, 1984. — 111 с.

34. *Ермаков, С.М.* О датчиках случайных чисел / С.М. Ермаков // Заводская лаборатория. — 1993. — Т. 59. — № 7. — С. 48–50.

35. *Неуймин, Я.Г.* Модели в науке и технике. История, теория, практика. — Ленинград : Наука, 1984. — 190 с.

36. *Моисеев, Н.Н.* Математические задачи системного анализа / Н.Н. Моисеев. — Москва : Наука, 1981. — 488 с.

37. *Нейлор, Т.* Машинные имитационные эксперименты с моделями экономических систем / Т. Нейлор. — Москва : Мир, 1975. — 500 с.

38. *Орлов, А.И.* О реальных возможностях бутстрепа как статистического метода / А.И. Орлов // Заводская лаборатория. — 1987. — Т. 53. — № 10. — С. 82–85.

39. *Бэстенс, Д.Э.* Нейронные сети и финансовые рынки: принятие решений в торговых операциях / Д.Э. Бэстенс, В.М. ван дер Берт, Д. Вуд. — Москва : ТВП, 1998.

40. *Новиков, А.М.* Методология / А.М. Новиков, Д.А. Новиков. — Москва : СИНТЕГ, 2007. — 668 с.

41. *Орлов, А.И.* О методологии статистических методов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 104. — С. 53–80.

42. *Горский, В.Г.* Математические методы исследования: итоги и перспективы / В.Г. Горский, А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2002. — Т. 68. — № 1. — С. 108–112.

43. *Эльясберг, П.Е.* Измерительная информация. Сколько ее нужно, как ее обрабатывать? / П.Е. Эльясберг. — Москва : Наука, 1983. — 208 с.

44. *Крамер, Г.* Математические методы статистики / Г. Крамер. — Москва : Мир, 1975. — 648 с.

45. *Орлов, А.И.* О поправках на группировку / А.И. Орлов, И.В. Орловский // Прикладной многомерный статистический анализ. — Москва : Наука, 1978. — С. 339–342.

46. *Орлов, А.И.* Поправка на группировку для коэффициента корреляции / А.И. Орлов // Экономика и математические методы. — 1980. — Т. XVI. — № 4. — С. 800–801.

47. *Феллер, В.* Введение в теорию вероятностей и ее приложения / В. Феллер. — Т. 2. — Москва : Мир, 1984. — 751 с.

48. *Боровков, А.А.* Теория вероятностей / А.А. Боровков. — Москва : Наука, 1976. — 352 с.

49. *Бернштейн, С.Н.* Современное состояние теории вероятностей и ее приложений / С.Н. Бернштейн // Труды Всероссийского съезда математиков в Москве 27 апреля — 4 мая 1927 г. — Москва : Ленинград : ГИЗ, 1928. — С. 50–63.

50. *Каган, А.М.* Характеризационные задачи математической статистики / А.М. Каган, Ю.В. Линник, С.Р. Рао. — Москва : Наука, 1972. — 656 с.

51. *Ибрагимов, И.А.* Асимптотическая теория оценивания / И.А. Ибрагимов, Р.З. Хасьминский. — Москва : Наука, 1979. — 528 с.

52. *Никитин, Я.Ю.* Асимптотическая эффективность непараметрических критериев / Я.Ю. Никитин. — Москва : Наука, 1995. — 240 с.

53. *Орлов, А.И.* О современных проблемах внедрения прикладной статистики и других статистических методов / А.И. Орлов // Заводская лаборатория. — 1992. — Т. 58. — № 1. — С. 67–74.

54. *Орлов, А.И.* Высокие статистические технологии / А.И. Орлов. / Заводская лаборатория. — 2003. — Т. 69. — № 11. — С. 55–60.

55. *Орлов, А.И.* Точки роста статистических методов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 103. — С. 136–162.

56 *Орлов, А.И.* О высоких статистических технологиях / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 105. — С. 14–38.

57. *Лойко, В.И.* Высокие статистические технологии и системно-когнитивное моделирование в экологии : монография / В.И. Лойко, Е.В. Луценко, А.И. Орлов. — Краснодар : КубГАУ, 2019. — 258 с.

58. *Орлов, А.И.* Новая парадигма прикладной статистики / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2012. — Т. 78. — № 1. — Ч. I. — С. 87–93.

59. *Орлов, А.И.* Основные черты новой парадигмы математической статистики / А.И. Орлов // Научный журнал КубГАУ. — 2013. — № 90. — С. 45–71.

60. *Орлов, А.И.* Новая парадигма анализа статистических и экспертных данных в задачах экономики и управления / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 98. — С. 1254–1260.

61. *Орлов, А.И.* О новой парадигме математических методов исследования / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 122. — С. 807–832.

62. *Орлов, А.И.* Смена парадигм в прикладной статистике / А.И. Орлов // Заводская лаборатория. Диагностика материалов. — 2021. — Т. 87. — № 7. — С. 6–7.

63. *Орлов, А.И.* Компьютерно-статистические методы: состояние и перспективы / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 103. — С. 163–195.

64. *Орлов, А.И.* Взаимосвязь предельных теорем и метода Монте-Карло / А.И. Орлов // Научный журнал КубГАУ. — 2015. — № 114. — С. 27–41.

65. *Орлов, А.И.* Применение метода Монте-Карло при изучении свойств статистических критериев однородности двух независимых выборок / А.И. Орлов // Научный журнал КубГАУ. — № 154. — С. 55–83.

66. *Орлов, А.И.* О методологии статистических методов / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 104. — С. 53–80.

67. *Орлов, А.И.* О влиянии методологии на последствия принятия решений / А.И. Орлов // Научный журнал КубГАУ. — 2017. — № 125. — С. 319–345.

68. *Орлов, А.И.* Методология моделирования процессов управления в социально-экономических системах / А.И. Орлов // Научный журнал КубГАУ. — 2014. — № 101. — С. 166–196.

69. *Орлов, А.И.* Системная нечеткая интервальная математика. Монография (научное издание) / А.И. Орлов, Е.В. Луценко. — Краснодар : КубГАУ. 2014. — 600 с.

70. *Орлов, А.И.* Системная нечеткая интервальная математика — основа математики XXI века / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 165. — С. 111–130.

71. *Орлов, А.И.* Отечественная научная школа в области эконометрики / А.И. Орлов // Научный журнал КубГАУ. — 2016. — № 121. — С. 235–261.

72. *Орлов, А.И.* Отечественная научная школа в области организационно-экономического моделирования, эконометрики и статистики / А.И. Орлов // Контроллинг. — 2019. — № 73. — С. 28–35.

73. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в цифровой экономике (на примере управления качеством) / А.И. Орлов // Научный журнал КубГАУ. — 2021. — № 169. — С. 216–242.

74. *Орлов, А.И.* Организационно-экономическое моделирование и искусственный интеллект в организации производства в эпоху цифровой экономики / А.И. Орлов // Инновации в менеджменте. — 2021. — № 2 (28). — С. 36–45.

Контрольные вопросы

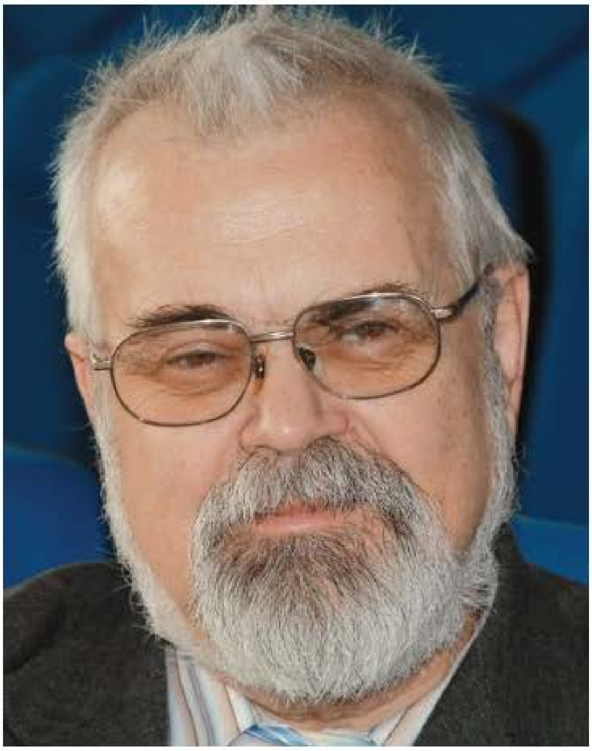
1. Какие «точки роста» статистических методов Вы знаете?
2. Чем непараметрическая статистика отличается от параметрической?
3. По каким признакам выделяются высокие статистические технологии?
4. Какова роль информационных технологий в статистических методах?
5. Какие вопросы обсуждаются в научном направлении «методология статистических методов»?
6. Какие нерешенные проблемы статистических методов являются наиболее актуальными в настоящее время?

Темы докладов, рефератов, исследовательских работ

1. Достоинства и недостатки различных датчиков псевдослучайных чисел.
2. Метод статистических испытаний (Монте-Карло) и его применения в теории и практике статистических методов.
3. Многообразие вариантов «бутстрепа».
4. Устойчивость и робастность
5. Основные идеи нечисловой статистики.
6. Нотна и рациональный объем выборки — основные понятия статистики интервальных данных.
7. Высокие статистические технологии в начале XXI века.
8. Актуальные методологические проблемы статистических методов.

ПРИЛОЖЕНИЕ

ОБ АВТОРЕ



Орлов Александр Иванович, 1949 г.р., профессор (1995 г. — по кафедре математической экономики), доктор экономических наук (2009 г. — по математическим и инструментальным методам экономики), доктор технических наук (1992 г. — по применению математических методов), кандидат физико-математических наук (1976 г. — по теории вероятностей и математической статистике).

Профессор кафедр «Экономика и организация производства» факультета «Инженерный бизнес и менеджмент» и «Вычислительная математика и математическая физика» факультета «Фундаментальные науки» Московского государственного технического университета им. Н.Э. Баумана, руководитель секции «Организационно-экономическое моделирование, эконометрика и статистика», директор Института высоких статистических технологий и эконометрики, заведующий Лабораторией экономико-математических методов в контроллинге.

Член редколлегий журналов «Заводская лаборатория. Диагностика материалов», «Контроллинг», «Инновации в менеджменте», «Социология: методология, методы, математическое моделирование», «Управление большими системами: сборник трудов». Главный редактор электронного еженедельника «Эконометрика».

Академик Международной академии исследований будущего, Российской Академии статистических методов. Вице-президент Всесоюзной Статистической Ассоциации, президент Российской ассоциации статистических методов.

Основные направления научной и педагогической деятельности: теория принятия решений, прикладная статистика и другие статистические методы, эконометрика, экономико-математические методы, экспертные оценки, менеджмент, экономика предприятия, макроэкономика, экология.

Автор более 1 100 научных и методических публикаций в России и за рубежом, в том числе более 50 книг. Один из наиболее цитируемых математиков и экономистов России.

Более подробная информация приведена на сайте «Википедия», в статье «Орлов, Александр Иванович (ученый)».

Основные книги профессора А.И. Орлова

1. *Орлов, А.И.* Устойчивость в социально-экономических моделях / А.И. Орлов. — Москва : Наука,1979. — 296 с.

2. *Орлов, А.И.* Задачи оптимизации и нечеткие переменные / А.И. Орлов. — Москва : Знание, 1980. — 64 с.

3. *Орлов, А.И.* Анализ нечисловой информации (препринт) / А.И. Орлов, Ю.Н. Тюрин, Б.Г. Литвак, Г.А. Сатаров, Д.С. Шмерлинг. — Москва : Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1981. — 80 с.

4. *Гусев, В.А.* Внеклассная работа по математике в 6–8 классах / В.А. Гусев, А.И. Орлов, А.Л. Розенталь. — Москва : Просвещение, 1977. — 288 с. (2-е изд., испр. и доп. — Москва : Просвещение, 1984.). Переводы на казахский, литовский, молдавский, таджикский языки.

5. *Орлов, А.И.* Пакет программ анализа данных «ППАНД» : учебное пособие / А.И. Орлов, И.Л. Легостаевая, О.М. Черномордик. — Москва : Сотрудничающий центр Всемирной организации здравоохранения по профессиональной гигиене, 1990. — 93 с.

6. *Орлов, А.И.* Математическое моделирование процессов налогообложения (подходы к проблеме) / А.И. Орлов, В.Г. Кольцов, Н.Ю. Иванова. — Москва : Изд-во ЦЭО Министерства общего и профессионального образования РФ, 1997. — 232 с.

7. *Орлов, А.И.* Экология : учебное пособие / А.И. Орлов, С.А. Боголюбов. — Москва : Знание, 1999. — 288 с.

8. *Орлов, А.И.* Менеджмент : учебное пособие / А.И. Орлов, С.А. Боголюбов, Ж.В. Прокофьева. — Москва : Знание, 2000. — 288 с.

9. *Орлов, А.И.* Управление качеством окружающей среды : учебник / А.И. Орлов, С.А. Боголюбов. — Т. 1. — Москва : МГИЭМ(ту), 2000. — 283 с.

10. *Орлов, А.И.* Системы экологического управления : учебник / А.И. Орлов, С.А. Боголюбов. Москва : Европейский центр по качеству, 2002. — 224 с.

11. *Орлов, А.И.* Эконометрика : учебник. — Москва : Экзамен, 2002 (1-е изд.), 2003 (2-е изд.), 2004 (3-е изд.). — 576 с.

12. Управление промышленной и экологической безопасностью : учебное пособие / А.И. Орлов, В.Н. Федосеев, В.Г. Ларионов, А.Ф. Козьяков. — Москва : УРАО, 2002 (1-е изд.), 2003 (2-е изд.). — 220 с.

13. *Орлов, А.И.* Менеджмент в техносфере : учебное пособие / А.И. Орлов, В.Н. Федосеев. — Москва : Академия, 2003. — 384 с.

14. *Орлов, А.И.* Теория и методы разработки управленческих решений : учебное пособие / А.И. Орлов. — Москва : ИКЦ «МарТ» ; Ростов-на-Дону : Издательский центр «МарТ», 2005. — 496 с.

15. *Орлов, А.И.* Прикладная статистика : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 672 с.

16. *Орлов, А.И.* Теория принятия решений : учебник / А.И. Орлов. — Москва : Экзамен, 2006. — 576 с.

17. *Орлов, А.И.* Проектирование интегрированных производственно-корпоративных структур: эффективность, организация, управление / А.И. Орлов, С.Н. Анисимов, А.А. Колобов [и др.] : под редакцией А.А. Колобова, А.И. Орлова. — Москва : Изд-во МГТУ им. Н.Э. Баумана, 2006. — 728 с.

18. *Колобов, А.А.* Менеджмент высоких технологий. Интегрированные производственно-корпоративные структуры: организация, экономика, управление, проектирование, эффективность, устойчивость / А.А. Колобов, И.Н. Омельченко, А.И. Орлов. — Москва : Экзамен, 2008. — 621 с.

19. *Орлов, А.И.* Организационно-экономическое моделирование : учебник : в 3 частях. Ч. 1: Нечисловая статистика / А.И. Орлов. — Москва : Изд-во МГТУ им. Н.Э. Баумана, 2009. — 542 с.

20. *Орлов, А.И.* Эконометрика : учебник для вузов / А.И. Орлов. — 4-е изд., доп. и перераб. — Ростов-на-Дону : Феникс, 2009. — 572 с.

21. *Орлов, А.И.* Менеджмент: организационно-экономическое моделирование : учебное пособие для вузов / А.И. Орлов. — Ростов-на-Дону : Феникс, 2009. — 475 с.

22. *Орлов, А.И.* Вероятность и прикладная статистика: основные факты : справочник / А.И. Орлов. — Москва : КноРус, 2010. — 192 с.

23. *Орлов, А.И.* Организационно-экономическое моделирование: теория принятия решений : учебник. — Москва : КноРус, 2011. — 568 с.

24. *Орлов, А.И.* Организационно-экономическое моделирование : учебник : в 3 частях. Ч. 2. Экспертные оценки / А.И. Орлов. — Москва : Изд-во МГТУ им. Н.Э. Баумана, 2011. — 486 с.

25. *Орлов, А.И.* Устойчивые экономико-математические методы и модели. Разработка и развитие устойчивых экономико-математических методов и моделей для модернизации управления предприятиями / А.И. Орлов. — Saarbrücken : Lambert Academic Publishing, 2011. — 436 с.

26. *Орлов, А.И.* Организационно-экономическое моделирование : учебник : в 3 частях. Ч. 3. Статистические методы анализа данных / А.И. Орлов. — Москва : Изд-во МГТУ им. Н.Э. Баумана, 2012. — 624 с.

27. *Орлов, А.И.* Проблемы управления экологической безопасностью. Итоги двадцати лет научных исследований и преподавания / А.И. Орлов. — Saarbrücken : Palmarium Academic Publishing, 2012. — 344 с.

28. *Орлов, А.И.* Системная нечеткая интервальная математика : монография / А.И. Орлов, Е.В. Луценко. — Краснодар : Изд-во КубГАУ, 2014. — 600 с.

29. *Орлов, А.И.* Перспективные математические и инструментальные методы контроллинга : монография / А.И. Орлов, Е.В. Луценко, В.И. Лойко ; под научной редакцией профессора С.Г. Фалько. — Краснодар : Изд-во КубГАУ, 2015. — 600 с.

30. *Орлов, А.И.* Организационно-экономическое, математическое и программное обеспечение контроллинга, инноваций и менеджмента : монография / А.И. Орлов, Е.В. Луценко, В.И. Лойко ; под общей редакцией С.Г. Фалько. — Краснодар : Изд-во КубГАУ, 2016. — 600 с.

31. *Лойко, В.И.* Современные подходы в наукометрии: монография / В.И. Лойко, Е.В. Луценко, А.И. Орлов ; под научной редакцией профессора С.Г. Фалько. — Краснодар : Изд-во КубГАУ, 2017. — 532 с.

32. *Орлов, А.И.* Методы принятия управленческих решений : учебник / А.И. Орлов. — Москва : КНОРУС, 2018. — 286 с.

33. *Лойко, В.И.* Современная цифровая экономика / В.И. Лойко, Е.В. Луценко, А.И. Орлов. — Краснодар : КубГАУ, 2018. — 508 с.

34. *Лойко, В.И.* Высокие статистические технологии и системно-когнитивное моделирование в экологии : монография / В.И. Лойко, Е.В. Луценко, А.И. Орлов. — Краснодар : Изд-во КубГАУ, 2019. — 258 с.

35. *Агаларов, З.С.* Эконометрика : учебник / З.С. Агаларов, А.И. Орлов. — Москва : Дашков и К°, 2021. — 380 с.