А.И. Орлов

ОРГАНИЗАЦИОННО-ЭКОНОМИЧЕСКОЕ МОДЕЛИРОВАНИЕ:

ТЕОРИЯ ПРИНЯТИЯ РЕШЕНИЙ

Допущено УМО вузов
по университетскому политехническому образованию
в качестве **учебника** для студентов вузов,
обучающихся по направлению 220700 «Организация
и управление наукоемкими производствами»
специальности 220701 «Менеджмент высоких технологий»

КноРус

МОСКВА
2011

УДК 519.2:330.4(075.8)
ББК 65.04я73
О-66

Рецензенты:
Д.А. Новиков, заместитель директора Института проблем управления Российской академии наук, чл.-корр. РАН, д-р техн. наук, проф.,
В.Н. Лившиц, заведующий кафедрой «Оценка эффективности инвестиционных проектов» факультета инноваций и высоких технологий МФТИ, засл. деятель науки РФ, д-р экон. наук, проф.

Орлов А.И.
О-66 Организационно-экономическое моделирование: теория принятия решений : учебник / А.И. Орлов. — М. : КНОРУС, 2010. — 568 с.
ISBN 978-5-406-00275-9

Представлены теория и практика разработки управленческих решений на основе организационно-экономического моделирования. Рассмотрены основы теории принятия решений, технология и процедуры разработки и принятия управленческих решений. Разобраны оптимизационные и вероятностно-статистические методы принятия решений. Большое внимание уделено экспертным технологиям. Приведены как традиционные, так и недавно разработанные методы принятия решений, даны примеры их применения для решения практических задач.

Для студентов и преподавателей вузов, слушателей бизнес-школ, программ MBA, институтов повышения квалификации и структур второго образования, менеджеров, экономистов, инженеров, научных и практических работников, связанных с принятием решений на основе анализа экономических и управленческих данных.

УДК 519.2:330.4(075.8)
ББК 65.04я73

Орлов Александр Иванович

ОРГАНИЗАЦИОННО-ЭКОНОМИЧЕСКОЕ МОДЕЛИРОВАНИЕ: ТЕОРИЯ ПРИНЯТИЯ РЕШЕНИЙ

Санитарно-эпидемиологическое заключение
№ 77.99.60.953.Д.003365.04.09 от 01.04.2009 г.

Изд. №. 1720. Подписано в печать 31.03.2010. Формат 60×90/16.
Гарнитура «PetersburgC». Печать офсетная. Бумага газетная.
Усл. печ. л. 35,5. Уч.-изд. л. 30,0. Тираж ?000 экз. Заказ №

ООО «Издательство КноРус».
129110, Москва, ул. Большая Переяславская, 46.
Тел.: (495) 680-7254, 680-0671, 680-1278.
E-mail: office@knorus.ru http://www.knorus.ru

Отпечатано в полном соответствии с качеством
предоставленного издательством электронного оригинал-макета
в ОАО «Московская типография № 2».
129085, Москва, пр. Мира, 105.

ISBN 978-5-406-00275-9

Оглавление

Предисловие

Решения принимают все — инженеры, менеджеры, экономисты, домохозяйки и космонавты. Принятие решений — основа любого управления. Поэтому знакомство с современной теорией принятия решений необходимо всем, связанным с системами управления. А управляет каждый из нас — хотя бы самим собой.

Содержание учебника. Первая из четырех его частей раскрывает теории принятия решений, глава 1 — введение в теорию принятия управленческих решений. На примере типовой задачи о запуске в серию того или иного типа автомобиля показаны возникающие проблемы. Рассмотрены четыре аналитических и три практических критерия принятия решений, в качестве восьмого — голосование как один из методов экспертных оценок. Обсуждены свойства процедур голосования, в том числе парадокс Кондорсе. Введены основные понятия теории принятия решений: лица, принимающие решения, порядок подготовки решения (регламент), цели и ресурсы, риски и неопределенности, критерии оценки решения. Обсуждаются реальные процедуры принятия решений и их математико-компьютерная поддержка.

Глава 2 посвящена широко применяющимся простым методам принятия решений. Разобраны подходы в стратегическом менеджменте, оперативные приемы, способы декомпозиции задач принятия решения.

Основы теории управления — предмет главы 3. Введены основные понятия, выявлена проблема многокритериальности реальных задач управления, обсуждены подходы к оптимальному управлению экономическими системами.

В дальнейших частях учебника рассматривается научный инструментарий современной теории принятия решений.

Математические методы разработки и принятия решений рассмотрены в части II учебника.

Содержание главы 4 — задачи оптимизации. В линейном программировании последовательно рассматриваются упрощенная производственная задача (с графическим решением) и двойственная к ней, задачи о диете, планировании номенклатуры и объемов выпуска, транспортная задача. Дается первоначальное представление о линейном программировании как научно-практической дисциплине. Рассмотрены методы решения задач линейного программирования: простой перебор, направленный перебор, симплекс-метод. К целочисленному программированию относятся задача о выборе оборудования и задача о ранце. К ним примыкает тематика бинарных отношений и дискретной оптимизации в экспертных оценках — одном из инструментов принятия решений. Обсуждаются подходы к решению задач целочисленного программирования — метод приближения непрерывными задачами и методы направленного перебора. Заключительный подраздел главы 4 — оптимизация на графах. Рассмотрены задачи коммивояжера, о кратчайшем пути, о максимальном потоке. Сформулирована задача линейного программирования при максимизации потока.

Современная прикладная статистика и другие статистические методы активно используются при разработке и принятии решений в различных областях. Систематическому изложению этих весьма эффективных методов посвящен ряд книг автора (учебники «Эконометрика», «Прикладная статистика», «Нечисловая статистика», справочник «Вероятность и прикладная статистика — основные факты» и др.). В данном учебнике в качестве базовых примеров разобраны две центральные темы из рассматриваемой области — организационно-экономическое моделирование на основе восстановления зависимостей методом наименьших квадратов и анализ динамики цен и использование индексов инфляции при принятии хозяйственных решений. Приведены результаты научных исследований, выполненных в 2008 г.

Глава 5 посвящена вопросам регрессии, корреляции и прогнозирования. В непараметрической постановке рассмотрены методы восстановления линейной зависимости между двумя переменными. Обсуждаются основы линейного регрессионного анализа, коэффициенты корреляции Пирсона и Спирмена. Практический пример — прогнозирование в отрасли лома черных металлов. Завершают главу 5 подразделы, повествующие о выборе вида регрессионной модели, непараметрическом оценивании точки пересечения регрессионных прямых, анализе статистических данных в рамках модели с периодической составляющей.

Анализ динамики цен и использование индексов инфляции при принятии управленческих решений — предмет главы 6, которая начинается с определения индексов инфляции и правил их расчетов на основе практически используемых на практике потребительских корзин. Рассмотрены свойства индексов инфляции, доказаны теоремы умножения и сложения. Использование индексов инфляции в экономических расчетах продемонстрировано на примере анализа динамики цен на продовольственные товары.

Часть III учебника посвящена методам и технологиям сбора и анализа мнений экспертов, применению экспертных оценок. Она

состоит из восьми глав. Экспертные технологии пока недостаточно представлены в литературе, поэтому мы вынуждены уделить им большое внимание.

Примеры процедур экспертных оценок даны в главе 7. Рассмотрены индивидуальные и коллективные экспертные оценки, методы оценки и выбора вариантов с помощью экспертов, процедуры экспертного прогнозирования, место экспертных оценок в теории и практике принятия решений на современном этапе.

Организационной стороне работы экспертной комиссии посвящена глава 8. Обсуждаются основные стадии экспертного опроса, в том числе выбор цели экспертизы и подбор экспертов. Выделены основания для классификации экспертных методов. Рассмотрена роль интуиции эксперта и использование информационных технологий.

Теория измерений и ее применение для обоснования экспертных процедур — предмет главы 9. Введены основные шкалы измерения. Поставлена задача поиска инвариантных алгоритмов. В качестве примера разобраны методы усреднения. Дан анализ различных видов средних, введены средние по Коши и по Колмогорову. Установлено, какими средними величинами следует пользоваться при анализе данных, измеренных в порядковой шкале (из средних по Коши), шкалах интервалов и отношений (из средних по Колмогорову).

В главе 10 для нахождения коллективного мнения по экспертным ранжировкам предложены методы средних арифметических и медиан рангов, а также процедура согласования кластеризованных ранжировок.

Глава 11 посвящена основным математическим задачам анализа экспертных оценок. На основе систем аксиом введено понятие «расстояние между экспертными мнениями». Мнение экспертной комиссии предложено определять с помощью медианы Кемени. Коэффициенты корреляции и конкордации рассмотрены в связи с проверкой согласованности мнений экспертов.

Важный частный вид экспертных оценок — бинарные данные и результаты парных сравнений — разобраны в главе 12. Дано обоснование «турнирного» метода ранжирования вариантов. Построена теория случайных толерантностей. Предложен метод проверки гипотез по совокупности малых выборок. Развита теория люсианов, соответствующая непараметрической модели парных сравнений. Рассмотрены и параметрические модели в этой области.

Построению рейтингов (обобщенных показателей) посвящена глава 13. В качестве основной модели выбраны бинарные рейтинги, тесно связанные с теорией классификации (диагностики, дискримиции, распознавания образов). В задачах сравнения рейтингов основное внимание уделено линейным рейтингам. Обосновано применение прогностической силы как показателя качества алгоритма диагностики, построена асимптотическая теория для этого показателя и разработаны методы проверки обоснованности пересчета на модель линейного дискриминантного анализа.

Примеры разработки управленческих решений на основе экспертных оценок приведены в главе 14. Речь идет о маркетинговых исследованиях, экспертных технологиях в системе «Шесть сигм», применении иерархической системы показателей технического уровня и качества продукции. Рассказано о применении экспертных оценок при упорядочении системы государственных стандартов по статистическим методам управления качеством продукции. Обсуждается роль экспертных оценок в оценочной деятельности и инвестиционном менеджменте. Применена сценарная технология экспертного прогнозирования.

Часть IV учебника посвящена применению метода моделирования в теории принятия решений и обсуждению нескольких конкретных семейств моделей. В главе 15 рассмотрены основные понятия общей теории моделирования, в том числе математического, и методология моделирования, а также два примера — модель управления обучением (на основе системы дифференциальных уравнений) и вероятностно-статистическая модель помех, создаваемых электровозами.

Глава 16 посвящена типовым моделям в теории принятия решений, в том числе макроэкономическим моделям экономики отдельных стран и мирового хозяйства в целом, моделированию процессов налогообложения в России и других странах. Рассмотрено применение микроэкономических моделей: модели функционирования промышленного предприятия, экономико-математической модели проблем малого бизнеса, принятия решений в задачах логистики (управления запасами).

Модели обеспечения качества — предмет главы 17. Даны основы теории статистического контроля, построена асимптотическая теория одноступенчатых планов. Рассмотрены методы обнаружения разладки (отклонения течения процессов от планового) с помощью контрольных карт. Обсуждаются практические вопросы принятия решений при статистическом контроле качества продукции и услуг. Показано, что выходной контроль качества продукции нужен не всегда. Приведены удачные и неудачные примеры принятия решений в области качества и сертификации. Проанализирован опыт работ Всесоюзного центра статистических методов и информатики и Российской ассоциации статистических методов.

В заключении кратко рассказано о перспективных направлениях теории и практики принятия решений на основе организационно-экономического моделирования.

Как был написан учебник? У автора было два стимула.

Во-первых, необходимо сделать доступным широкой массе читателей более чем тридцатилетний опыт междисциплинарного научного коллектива, действующего вокруг семинара «Экспертные оценки и анализ данных». Семинар был организован в 1973 г. и работал сначала в МГУ им. М.В. Ломоносова, а затем в Институте проблем управления Российской академии наук. Некоторое время автор руководил семинаром (вместе с коллегами). Именно в рамках этого междисциплинарного коллектива была создана отечественная научная школа в области современной теории принятия решений и экспертных оценок.

Во-вторых, требуется подготовить учебник по теории принятия решений для обеспечения различных видов образовательных услуг. После сравнения различных подходов к преподаванию, многочисленных вариантов организации обучения автор решил взять за исходный пункт курс «Теория принятия решений» российско-французской программы МАСТЕР («Менеджмент промышленных систем»). Она с 1995 г. реализовывалась научно-учебным комплексом «Инженерный бизнес и менеджмент» Московского государственного технического университета им. Н.Э. Баумана совместно с Высшими техническими школами Парижа и Лиона.

Итак, учебник опирается на научные разработки последних лет и практику преподавания в России и во Франции с учетом достижений специалистов других стран.

Учебник А.И. Орлова «Теория принятия решений» был выпущен в 2006 г. издательством «Экзамен», а предварительно, годом ранее — в 2005 г. — появился его существенно сокращенный вариант — учебное пособие «Принятие решений. Теория и методы разработки управленческих решений» (Издательский центр «МарТ»).

Развитие научно-технического прогресса ставит перед инженерами, управленцами и экономистами новые задачи. В соответствии с потребностями практики в 2005 г. введена новая учебная специальность 220701 «Менеджмент высоких технологий», относящаяся к тогда же введенному направлению подготовки 220700 «Организация и управление наукоемкими производствами», предназначенному для обеспечения инженерами-менеджерами высокотехнологичных предприятий оборонно-промышленного комплекса. Для новой специальности понадобилась разработка нового научно-методического обеспечения, в том числе новых учебных дисциплин и соответствующих учебников (в частности, по организационно-экономическому моделированию), основанных на последних научно-технических разработках, подкрепленных практическим опытом. Был сформирован блок учебных дисциплин «Организационно-экономическое моделирование», в который были включены интеллектуальные инструменты современного менеджмента высоких технологий.

Организационно-экономическое моделирование — научная, практическая и учебная дисциплина, посвященная разработке, изучению и применению математических и статистических методов и моделей в экономике и управлении ее субъектами, прежде всего промышленными предприятиями и их объединениями. Ее существенная часть — теория принятия решений.

Понадобился новый учебник. Он перед вами. По сравнению с вариантами 2005—2006 гг. он имеет иную структуру. Гораздо больше рассказано об экспертных оценках (часть III вместо одной главы). Расширен объем материала, посвященного работе менеджера. В связи с выходом книг по статистическим методам соответствующие разделы настоящего учебника существенно сокращены. Включены результаты научных исследований последних лет, а устаревшие материалы исключены. Имеется целый ряд иных изменений по всему тексту. Перед вами — новая книга, а не новое издание учебника 2006 г.

Для кого эта книга? Учебник может быть использован различными категориями читателей. Студенты дневных отделений управленческих и экономических специальностей найдут в нем весь необходимый материал для изучения различных вариантов курсов типа «Теория принятия решений», «Управленческие решения», «Организационно-экономическое моделирование», «Экспертные оценки», «Экономико-математическое моделирование» и др. Особенно хочется порекомендовать учебник тем, кто получает наиболее ценимое в настоящее время образование — на экономических факультетах в технических вузах. Слушатели вечерних отделений, в том числе получающие второе образование по экономике и менеджменту, смогут изучить основы теории принятия решений и познакомиться с вопросами ее практического использования. Менеджерам, экономистам и инженерам, изучающим теорию принятия решений самостоятельно или в институтах повышения квалификации, учебник позволит познакомиться с ее ключевыми идеями и выйти на современный уровень.

Включенные в учебник материалы будут полезны не только студентам дневных и вечерних факультетов и слушателям системы второго высшего образования, но и тем, кто обучается по программам переподготовки, «Мастер (магистр) делового администрирования» (MBA)

и иным программам, в том числе международным. Специалистам по теории принятия решений, экспертным оценкам, теории управления, теории вероятностей и математической статистике эта книга также может быть интересна и полезна. В ней описан современный взгляд на рассматриваемую тематику, ее основные подходы и результаты, открывающие большой простор для дальнейших математических исследований.

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

К каждой главе прилагаются контрольные вопросы, примерные темы докладов и рефератов.

О роли литературных ссылок в учебнике необходимо сказать подробнее. Книга представляет собой замкнутый текст, не требующий для своего понимания ничего, кроме знания стандартных учебных курсов высшей математики и основ экономической теории. Зачем же нужны ссылки? Доказательства всех приведенных в учебнике теорем приведены в ранее опубликованных статьях и монографиях. Дотошный читатель при подготовке рефератов и при желании глубже проникнуть в материал книги может обратиться списку цитированной литературы. Каждая из глав учебника — это только введение в большую область теории принятия решений и организационно-экономического моделирования, и вполне естественным является желание выйти за пределы введения. Приведенная литература может этому помочь. За многие десятилетия накопились большие книжные богатства, 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.

Автор благодарен своим многочисленным коллегам, слушателям и студентам, прежде всего различных образовательных структур Московского государственного технического университета им. Н.Э. Баумана, Московского физико-технического института, Российской экономической академии им. Г.В. Плеханова и Академии народного хозяйства при Правительстве Российской Федерации (программа «Топ-Менеджер»), за полезные обсуждения.

Автор благодарен сотрудникам издательства «КноРус» за поддержку нашего научного направления и большую работу по подготовке рукописи к изданию.

С текущей научной информацией по организационно-экономическому моделированию и теории принятия решений можно познакомиться на сайте автора http://orlovs.pp.ru (и его версиях www.antorlov.nm.ru, www.antorlov.chat.ru, www.newtech.ru/~orlov, www.antorlov.euro.ru), его форуме http://forum.orlovs.pp.ru/, а также на странице Лаборатории экономико-математических методов в контроллинге http://www.ibm.bmstu.ru/nil/lab.html (на сайте научно-учебного комплекса «Инженерный бизнес и менеджмент» Московского государственного технического университета им. Н.Э. Баумана).

Большой объем информации по современным научным исследованиям по тематике учебника содержит электронный еженедельник «Эконометрика» (http://subscribe.ru/catalog/science.humanity.econometrika), выпускаемый с июля 2000 г. Автор искренне благодарен разработчику сайтов и редактору электронного еженедельника А.А. Орлову за многолетний энтузиазм.

В книге раскрыто представление о теории принятия решений, соответствующее общепринятому в мире. Сделана попытка довести рассказ 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 направят свои вопросы и замечания по адресу издательства или непосредственно автору на форуме сайта http://orlovs.pp.ru или по электронной почте E-mail: prof-orlov@mail.ru.

Часть I

ОСНОВЫ ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

ГЛАВА 1
ВВЕДЕНИЕ В ТЕОРИЮ ПРИНЯТИЯ РЕШЕНИЙ

1.1. ПРИМЕР ЗАДАЧИ ПРИНЯТИЯ РЕШЕНИЯ КОМИССИЕЙ ЭКСПЕРТОВ

Разберем несколько упрощенный пример задачи принятия решений при управлении организацией.

Суть задачи. Совет директоров фирмы «Русские автомобили» должен принять важное решение. Какой образец нового автомобиля запускать в серию — маленького верткого «Алешу» или представительного «Добрыню»? Отличаются эти типы автомобилей прежде всего расходом бензина на 100 км пробега: «Добрыня» длиннее, шире, выше, тяжелее, а потому и бензина ему надо больше, чем «Алеше». Зато «Добрыня» гораздо солиднее и вместительнее. Как показывают маркетинговые исследования, при дешевом бензине потребители предпочтут «Добрыню», при дорогом — «Алешу». Будущая цена бензина неизвестна, это — фактор риска для фирмы «Русские автомобили».

Итак, каждый из двух вариантов решения имеет плюсы и минусы. Для принятия решения явно не хватает следующей количественной информации:

- насколько вероятна к моменту выхода продукции на рынок низкая цена бензина и насколько — высокая;
- каковы будут финансовые результаты работы фирмы при различных сочетаниях цены бензина и типа выпускаемого автомобиля (а таких сочетаний четыре: низкая цена бензина и выпуск «Алеши», низкая цена бензина и выпуск «Добрыни», высокая цена бензина и выпуск «Алеши», высокая цена бензина и выпуск «Добрыни»).

На эти вопросы генеральный директор фирмы заранее поручил ответить соответствующим специалистам. Перед началом заседания члены совета директоров получают нужные для принятия решения количественные данные, сведенные в табл. 1.1. В частности, за то, что цена бензина окажется низкой, есть 60 шансов из 100, т.е. 60%. А за то, что она окажется высокой, — 40 шансов из 100.

Дискуссия на совете директоров. На заседании совета директоров началась дискуссия. Сначала выступили четыре высококвалифицированных эксперта, каждый из которых использовал свою экономико-математическую модель.

Таблица 1.1

Прибыль фирмы «Русские автомобили» при выпуске автомобилей двух типов, млн руб.

Цена бензина	Тип автомобиля	
	«Алеша»	«Добрыня»
Низкая (60%)	750	1 000
Высокая (40%)	500	200

— Полагаю, надо получить максимум в самом плохом случае, — сказал осторожный Воробьев. — А хуже всего будет при высокой цене бензина — прибыль фирмы по сравнению со случаем низкой его цены уменьшается при любом нашем решении. Выпуская «Алешу», заработаем 500 миллионов, а «Добрыню» — 200 миллионов. Значит, надо выпускать «Алешу», и как минимум 500 миллионов нам обеспечены.

— Нельзя быть таким пессимистом, — заявил горячий Лебедев. — Скорее всего, цена бензина будет низкой (за это — 60 шансов из 100, т.е. больше половины), а высокой — лишь как исключение. Надо быть оптимистами — исходить из того, что все пойдет, как мы хотим, цена бензина будет низкой. Тогда, выпуская «Добрыню», получим миллиардную прибыль.

— На мой взгляд, и пессимист Воробьев, и оптимист Лебедев обсуждают крайние случаи — самую худшую ситуацию и самую лучшую. А надо подходить системно, обсудить ситуацию со всех сторон, учесть обе возможности, — начал свое выступление обстоятельный Чибисов, когда-то изучавший теорию вероятностей. — Давайте рассчитаем среднюю прибыль. Рассмотрим сначала первый вариант — выпуск «Алеши». Мы получим 750 миллионов в 60% случаев (при низкой цене бензина) и 500 миллионов в 40% случаев (при высокой его цене), значит, в среднем $750 \times 0{,}6 + 500 \times 0{,}4 = 450 + 200 = 650$ миллионов. А для варианта «Добрыни» аналогичный расчет дает $1000 \times 0{,}6 + 200 \times 0{,}4 = 600 + 80 = 680$ миллионов, т.е. больше. Значит, надо выпускать «Добрыню».

— Предыдущий оратор рассуждает так, как будто мы будем выбирать тип автомобиля на каждом заседании совета директоров, да и все данные в табл. 1.1 лет сто не изменятся, — вступил в дискуссию реалист Куликов. — Но нам предстоит принять решение только один раз, и сделать это надо так, чтобы потом не жалеть об упущенных возможностях. Если мы решим выпускать «Добрыню», а к моменту выхода на рынок цена бензина окажется высокой, то получим 200 миллионов вместо 500 миллионов при решении, соответствующем будущей высокой цене бензина. Если же цена бензина будет низкой, мы прибыль не упускаем. Значит, максимально возможная упущенная выгода составит 500 – 200 = 300 миллионов. При выпуске «Алеши» в случае низкой цены бензина упущенная выгода составит 1000 – 750 = 250 миллионов, а при высокой цене бензина мы прибыль не упускаем. Значит, при выпуске «Алеши» максимально возможная упущенная выгода равна 250 миллионам, т.е. будет меньше, чем если мы решим выпускать «Добрыню». Значит, надо выпускать «Алешу», если мы хотим минимизировать максимально возможную упущенную выгоду.

После экспертов захотели высказаться трое членов совета директоров. Финансовый директор Волков потребовал:

— Любой проект, который совет утвердит, должен давать не менее 400 миллионов. Иначе у нас будут трудности в работе с банками. Значит, «Добрыня» не годится. Надо выпускать «Алешу».

С других позиций выступил директор по развитию Вепрев:

— Наша фирма должна развиваться устойчиво. Мы должны иметь надежный прогноз. Чем меньше разброс результатов у проекта, тем лучше. У «Алеши» разброс 750 – 500 = 250 миллионов, а у «Добрыни» 1000 – 200 = 800 миллионов. Я за «Алешу».

У директора по маркетингу Лисицына иное мнение.

— Нельзя добиться успеха без риска. Как говорят, «кто не рискует, тот не пьет шампанское!». Как пишут в западных учебниках, предприниматель и менеджер должны рисковать! Лучше журавль в небе, чем синица в руках. Зачем нам синица? Мечта и победа — вот наш путь. Я — за «Добрыню»!

— Подведем итоги, — сказал председательствующий Медведев-Пчелкин. — Выступили четверо экспертов, каждый привел убедительные доводы в пользу того или иного решения, каждый исходил из той или иной теоретической концепции. При этом за выпуск «Алеши» выступили двое — Воробьев и Куликов, а за выпуск «Добрыни» также двое — Лебедев и Чибисов. Мнения выступивших членов совета директоров также разделились — осторожные Волков и Вепрев против игрока Лисицына. Решение надо принять сегодня, иначе понесем большие убытки. Будем голосовать.

Результаты голосования: 15 членов совета директоров — за выпуск «Добрыни», 8 (в основном более осторожные представители старшего поколения) — за выпуск «Алеши». Большинством голосов решение принято — фирма «Русские автомобили» будет выпускать «Добрыню».

Какие выводы может извлечь менеджер из стенограммы хода заседания совета директоров фирмы «Русские автомобили»? Критерии принятия решения, выдвинутые четырьмя выступавшими экспертами,

противоречили друг другу: два из них приводили к выводу о выгодности выпуска автомобиля «Алеша», а два — автомобиля «Добрыня». И совет директоров решил вопрос голосованием. При этом каждый из голосовавших интуитивно оценивал достоинства и недостатки вариантов, т.е. выступал как эксперт, а весь совет в целом — как экспертная комиссия. Рассмотренный пример наглядно демонстрирует, что экспертные оценки — один из универсальных методов принятия решений.

Из каких экономико-математических моделей исходили эксперты? Обсудим исходные позиции четырех экспертов. Пессимистическая позиция Воробьева основывается на вполне естественном предположении, что внешние силы действуют против нас, в частности, хотят нанести нам ущерб, подняв цены на бензин. Так бывает, когда мы ведем борьбу с непримиримым противником, когда наш успех означает такой же по величине проигрыш противника, а полученный нами ущерб — такой же по величине выигрыш противника. Одна из наиболее известных экономико-математических концепций — теория антагонистических игр — исходит из таких соображений. Эта концепция хорошо приспособлена для моделирования хода войны, поскольку в ходе боевых действий победа одной армии — это поражение армии противника. Хотя в современной теории игр рассматриваются возможности коалиций и сотрудничества, исходной точкой продолжает служить представление об антагонистических интересах игроков, как в шахматной партии.

Критика позиции пессимиста Воробьева может исходить из того, что внешний мир отнюдь не стремится нанести нам ущерб. Ясно, что для тех сил, взаимодействие которых определяет цены на нефть и бензин, фирма «Русские автомобили» не является противником. Они ее попросту не замечают, поскольку по сравнению с ними фирма «Русские автомобили» слишком мала. Разумнее считать, что будущие цены на бензин определяют силы, которые можно сравнить с природными, от которых зависит будущая погода. С этой точки зрения, позиция оптимиста Лебедева более обоснована, чем позиция пессимиста Воробьева, поскольку низкая цена бензина ожидается в 1,5 раза чаще, чем высокая.

Именно такие оптимисты, как Лебедев, разрабатывают обычно инвестиционные проекты и составляют бизнес-планы. Они считают, что удастся реализовать намеченное в заданные сроки и с заданными затратами. Правда, в конце современных бизнес-планов обычно имеется раздел, посвященный анализу рисков и управлению ими, однако исходной точкой продолжает служить представление о том, что оптимистический взгляд на будущее оправдан, а возникающие препятствия удастся преодолеть, не меняя общего плана действий.

Позиции пессимиста и оптимиста описывают крайние точки возможного будущего. На поиск компромисса между этими позициями нацелены многие экономико-математические модели. В соответствии с данными табл. 1.1 при выборе «Алеши» фирма получит от 500 млн до 750 млн руб. прибыли, а при выборе «Добрыни» — от 200 млн до 1000 млн руб. При принятии решений можно исходить из среднего арифметического граничных значений. Тогда выбору «Алеши» соответствует среднее значение $(500 + 750)/2 = 625$, а выбору «Добрыни» — среднее значение $(200 + 1000)/2 = 600$. Значит, по этому критерию надо запускать в серию «Алешу».

Очевиден произвол в усреднении минимального и максимального значений. Почему среднее арифметическое, а не среднее геометрическое? Почему минимальное и максимальное значения берутся с равными весами? Обобщая, получаем семейство методов, задаваемых параметром $\alpha \in [0; 1]$. Пусть минимальное значение берется с весом α, а максимальное — с весом $1 - \alpha$. Тогда выбору «Алеши» соответствует средневзвешенное значение $500\alpha + 750 (1 - \alpha)$, а выбору «Добрыни» — среднее значение $200\alpha + 1000 (1 - \alpha)$. Сравнивая эти два значения, заключаем, что при $\alpha < 5/11$ выгоднее запустить в серию «Добрыню», а при $\alpha \geq 5/11$ — «Алешу». Итак, решение определяется выбором параметра α. Как же в практической ситуации задать значение этого параметра?

Чибисов предлагает исходить из шансов осуществления того или иного прогноза цены бензина. Субъективная (т.е. оцененная экспертами) вероятность того, что цена бензина будет высокой, равна 0,4. Соответственно субъективная вероятность того, что цена бензина будет низкой, есть $1 - 0{,}4 = 0{,}6$. Значит, целесообразно положить $\alpha = 0{,}4$.

Однако и к рассуждениям Чибисова надо подойти критически. Хорошо известно, что субъективные вероятности могут быть далеки от объективных, соответствующих осуществлению большого числа испытаний в одних и тех же условиях. К тому же большого числа испытаний в рассматриваемой ситуации нет и быть не может — выбор проводится только один раз. Эту тему мы продолжим обсуждать в заключительном разделе настоящей главы.

Подход Куликова довольно часто рекомендуется во вводных курсах экономической теории. Однако «упущенная выгода» — это условная величина, а не «живые» деньги. В отличие от прибыли или выручки от реализации, отражающихся на банковских счетах предприятия, невозможно непосредственно использовать «упущенную выгоду» для решения конкретных задач управления финансово-хозяйственной

деятельностью предприятия. Неясно поэтому, имеет ли смысл принимать конкретные решения на основе чисто расчетной «упущенной выгоды».

Выступавшие члены совета директоров достаточно подробно обосновали свои выводы практическими соображениями, в том числе психологическими. Оставим их мнения без комментариев.

Итак, выступавшие разошлись во мнениях. Однако то или иное решение необходимо принять. Председателю пришлось поставить вопрос на голосование.

Может показаться, что голосование — это универсальный инструмент для принятия решений. Однако это не так. Есть много «подводных камней». Обсудим свойства процедур голосования.

1.2. ГОЛОСОВАНИЕ — ОДИН ИЗ МЕТОДОВ ЭКСПЕРТНЫХ ОЦЕНОК

Голосование — один из методов принятия решения комиссией экспертов. Организация голосования, в частности, на собрании акционеров, имеет свои «подводные камни».

Роль регламента. Многое зависит от регламента (т.е. правил проведения) голосования. Например, традиционным является принятие решений по *большинству голосов*: принимается то из двух конкурирующих решений, за которое поданы, по крайней мере, 50% голосов и еще один голос. А вот от какого числа отсчитывать 50% — от присутствующих или от списочного состава? Каждый из вариантов имеет свои достоинства и недостатки.

Если от присутствующих, то одно из двух решений будет почти наверняка принято (исключение — когда голоса разделятся точно поровну). Однако те, кто не был на собрании, могут быть недовольны и опротестовать решение. Очевидно, в ситуации, когда отсутствовало 90% от списочного состава, протест обоснован.

Если при принятии решения по большинству голосов исходить из списочного состава, то возникает проблема явки на заседание. При слабой явке решения присутствующими должны приниматься почти единогласно, следовательно, в ряде случаев ни одно из конкурирующих решений не будет принято. А если придет меньше 50% от утвержденного списочного состава, то принятие решений станет вообще невозможным.

Перечисленные сложности увеличиваются, если регламентом предусмотрено квалифицированное большинство — $^2/_3$ и еще один голос.

Используют также и метод *относительного большинства.* В соответствии с ним из ряда вариантов решения принимается то, за которое проголосуют больше участников голосования, чем за другие варианты. Согласно методу относительного большинства могут быть приняты решения, поддержанные 10 или 5% тех, кто подал голос.

О воздержавшихся. Еще одна проблема — как быть с воздержавшимися? Причислять ли их к голосовавшим «за» или к голосовавшим «против»? Ответ зависит от того, как поставлен вопрос председателем собрания: «Кто за?» или «Кто против?». Рассмотрим условный пример — результат голосования по трем кандидатурам в совет директоров (табл. 1.2).

Таблица 1.2

Мнения участников голосования на выборах в совет директоров

Кандидатура	Число голосов		
	за	против	воздержались
И.И. Иванов	200	100	100
П.П. Петров	150	50	200
С.С. Сидоров	0	0	400

Наиболее активным и результативным менеджером является И.И. Иванов. У него больше всего сторонников, но и больше всего противников. Его соперник П.П. Петров меньше себя проявил, у него меньше и сторонников и противников. Третий — С.С. Сидоров — никому не известен, и относительно его кандидатуры все участники голосования воздержались.

Пусть надо выбрать одного человека в совет директоров. Если председатель заседания спрашивает: «Кто за Иванова?», «Кто за Петрова?», «Кто за Сидорова?», то проходит И.И. Иванов. Если же председатель, видя усталость зала от обсуждения предыдущих голосований, спрашивает: «Кто против Иванова?» и т.п. (обоснование: чем меньше голосующих «против», тем бóльшую поддержку имеет кандидат), то выбирают «темную лошадку» С.С. Сидорова, поскольку активные противники остальных менеджеров «выбивают» их из соревнования.

При выборе двух членов совета директоров вопрос председателя «Кто за?» приводит к выбору И.И. Иванова и П.П. Петрова, а вопрос «Кто против?» — к выбору С.С. Сидорова и П.П. Петрова. Поэтому, желая избавиться от И.И. Иванова, председатель может при выборах ставить вопрос так: «Кто против?».

Нетрудно видеть, что вопрос «Кто за?» автоматически относит всех воздержавшихся к противникам данного кандидата, а вопрос «Кто против?» — к сторонникам. Успех никому не известного С.С. Сидорова связан именно с этим — он не нажил себе врагов.

В Государственной Думе голос депутата, отсутствующего на заседании или воздерживающегося, прибавляется к числу голосующих «против», поскольку для принятия законопроекта необходимо набрать не менее 226 голосов «за» из 450 (а при голосовании наиболее важных «конституционных» проектов — не менее 301).

Спрашивая «Кто за?», фактически исходим из принципа «Кто не с нами, тот против нас». А спрашивая «Кто против?», исходим из другого принципа — «Кто не против нас, тот с нами».

Последовательность голосований. Рассмотрим простейшую ситуацию — при голосовании простым большинством голосов (от числа присутствующих) решают, принять или отклонить обсуждаемое решение. Тогда очевидно, что принятое решение улучшает ситуацию для большинства голосовавших, а ухудшить может лишь для меньшинства. А каков будет результат нескольких последовательных голосований? Оказывается, возможна ситуация, при которой положение всех без исключения голосовавших ухудшается.

Рассмотрим условный пример. Пусть в голосованиях участвуют трое — Иванов, Петров и Сидоров. Пусть первым на голосование выносится такой проект решения: «Выделить Иванову и Петрову по 10 000 руб., а Сидорова отправить в тюрьму». Иванову и Петрову такое решение выгодно — их положение улучшается. Поэтому они голосуют «за». Сидоров, естественно, голосует «против». Два против одного — решение принимается. Сидоров отправляется в тюрьму (но сохраняет право участвовать в голосовании), а Иванов и Петров получают по 10 000 руб.

Второе голосование проводится по проекту решения: «Иванову и Сидорову — по 10 000 руб., Петрова — в тюрьму». «За» — Иванов и Сидоров, «против» — Петров. Решение принято и выполнено. Петров присоединяется к Сидорову, сидящему в тюрьме.

Проект решения для третьего голосования таков: «Петрову и Сидорову — по 10 000 руб., Иванова — в тюрьму». Два «за», один — «против». Решение принято.

Каков итог? Все трое — в тюрьме, но каждому из них выделено по 20 000 руб. Положение всех троих значительно ухудшилось.

Нечто подобное бывает и в реальных ситуациях. Во время Великой французской революции (1789—1794) в результате серии последовательных голосований в высшем органе власти (Конвенте) большинство депутатов отправились на эшафот.

Как добиться нужного решения с помощью голосования? Англичанин С.Н. Паркинсон подробно исследовал ряд отрицательных явлений, широко распространенных в организационных системах. Его весьма критическая книга [98] необходима любому менеджеру, где бы он ни работал — в государственной организации или в частной фирме. Она поможет избежать многих ошибочных решений, распространенных в среде управленцев.

С типично английской иронией С.Н. Паркинсон обсуждает вопрос о том, как добиться принятия нужного менеджеру решения, например, о выделении 100 млн фунтов стерлингов на некоторый проект. Он советует поставить его примерно на 25-е место среди 30 вопросов, вынесенных на обсуждение той комиссии, которая должна принять решение, а начать обсуждение с малозначительного вопроса, например, у какой фирмы покупать бумагу для принтера, используемого секретаршей комиссии?

Что будет происходить? «Свеженькие» члены комиссии с интересом приступят к обсуждению и не более чем за полчаса досконально разберут достоинства и недостатки различных фирм, поставляющих канцелярские принадлежности. Каждый будет рад высказаться и продемонстрировать коллегам свои познания (при этом никто не подумает о том, что за время, потраченное на это обсуждение, члены комиссии получат суммарную оплату много бóльшую, чем возможная экономия при покупке бумаги на 500 лет вперед).

Второй вопрос будет обсуждаться с несколько меньшим пылом. К десятому вопросу члены комиссии окончательно выдохнутся, многие из них перестанут следить за обсуждением, им будет лень даже поднимать руки при голосовании. И председатель перейдет на голосование по принципу «Кто против?».

Только перед самым обедом члены комиссии начнут просыпаться и проявлять активность. Именно поэтому наиболее важный для председателя вопрос лучше ставить на 25-е место, а не на последнее 30-е. При такой тактике построения заседания есть все основания ожидать, что после формулировки 25-го вопроса повестки дня на возглас председателя «Кто против?» не последует никакой реакции и нужное председателю решение будет единогласно принято.

Из работ С.Н. Паркинсона можно извлечь весьма много ироничных рекомендаций. Современные научные разработки сведены вместе в монографии [12].

Подробнее теория и практика экспертных оценок рассмотрена в одной из следующих частей книги.

1.3. ПАРАДОКС КОНДОРСЕ

Рассмотрим научный результат, положивший начало теории принятия решений.

В 1785 г. французский философ и математик М.-Ж.-А. де Кондорсе (1743—1794) опубликовал одну из первых в мировой истории работ, посвященную проблемам принятия коллективных решений комиссиями экспертов. Речь шла о решениях в ходе выборов депутатов провинциальных ассамблей — тогдашних органов региональной власти.

В этой работе впервые были введены такие важные для современной теории принятия решений и его важнейшего раздела — теории экспертных оценок — понятия, как принцип Кондорсе и парадокс Кондорсе. Согласно принципу Кондорсе для определения воли большинства необходимо, чтобы каждый голосующий проранжировал всех кандидатов в порядке их предпочтения — вместо того, чтобы избирать депутатов относительным или абсолютным большинством голосов.

Рассмотрим пример из работы Кондорсе. Пусть запись $A > B > C$ означает, что голосующий предпочитает кандидата A кандидату B, а кандидата B — кандидату C. Пусть мнения 60 экспертов таковы:

$$23 \text{ человека: } A > C > B;$$

$$19 \text{ человек: } B > C > A;$$

$$16 \text{ человек: } C > B > A;$$

$$2 \text{ человека: } C > A > B.$$

Сравним отношения экспертов к кандидатам A и B. Имеем:

$$23 + 2 = 25 \text{ человек за то, что } A > B;$$

$$19 + 16 = 35 \text{ человек за то, что } B > A.$$

По мнению Кондорсе, итоговое решение экспертной комиссии должно состоять в том, что B лучше A.

Сравнивая A и C, имеем:

$$23 \text{ человека за то, что } A > C;$$

$$37 \text{ человек за то, что } C > A.$$

Отсюда, по Кондорсе, заключаем, что большинство предпочитает кандидата C кандидату A.

Наконец, сравним кандидатов C и B:

$$19 \text{ человек за то, что } B > C;$$

$$41 \text{ человек за то, что } C > B.$$

Следовательно, согласно логике рассуждений Кондорсе большинство предпочитает кандидата C кандидату B.

Таким образом, по Кондорсе воля большинства выражается в виде трех суждений: $C > B$; $B > A$; $C > A$, которые, очевидно, можно объединить в одно отношение предпочтения $C > B > A$. Если необходимо выбрать одного из кандидатов, то согласно принципу Кондорсе следует предпочесть кандидата C.

Сравним этот вывод с возможным исходом голосования по мажоритарной системе. За A — 23 человека (они назвали A первым среди кандидатов), за B — 19 человек, за C — 18 человек. Таким образом, по системе относительного большинства победит кандидат A. При голосовании по системе абсолютного большинства кандидаты A и B выйдут во второй тур, где кандидат A получит 25 голосов, а кандидат B — 35 голосов и победит.

Таким образом, регламент голосования (другими словами, правила игры) будут определять победителя, и эти победители будут разными при различных правилах голосования.

Рассмотрим еще один пример Кондорсе. Пусть мнения 60 экспертов таковы:

23 человека: $A > B > C$;

17 человек: $B > C > A$;

2 человека: $B > A > C$;

10 человек: $C > A > B$;

8 человек: $C > B > A$.

Легко подсчитать, что $B > C$ для 42 проголосовавших, $C > A$ — для 35 и $A > B$ — для 33 экспертов.

В соответствии с принципом Кондорсе имеют место три утверждения: $B > C$, $C > A$, $A > B$. Но вместе эти утверждения противоречивы — при попытке упорядочить кандидатов получаем порочный круг! В этом и состоит знаменитый парадокс (эффект) Кондорсе (или парадокс голосования). Итак, существуют такие конкретные результаты голосования, при которых в соответствии с принципом Кондорсе оказывается невозможным принять какое-то согласованное решение и определить волю большинства.

В другой форме парадокс Кондорсе возникает при постатейном принятии некоторого постановления или закона, когда каждая из статей закона принимается большинством голосов, а поставленный на голосование закон в целом отвергается (иногда даже стопроцентным большинством голосующих). Еще хуже, если закон оказывается внутренне

противоречивым, включая в себя «порочные круги» статей, принятых большинством голосов.

Третьей версией парадокса Кондорсе является принятие таких коллективных решений, которые на индивидуальном уровне не поддерживал ни один из голосующих. Пусть у нас имеются три человека, голосующих по трем вопросам. Первый их них голосует да—да—нет, второй: да—нет—да, третий: нет—да—да. Суммарный итог голосования подсчитывается как соотношение сумм голосов «да» и «нет» по каждому из вопросов. В рассмотренном случае суммарный итог голосования будет да—да—да. Этот итог не отражает мнения ни одного из голосовавших. Такой же парадокс сопутствует любому способу получения итогового мнения экспертной комиссии, при котором это итоговое мнение может не совпадать ни с одним из высказанных экспертами мнений. Впрочем, если итоговое мнение и совпадет с мнением конкретного эксперта, остальные эксперты могут счесть себя обиженными.

Различные формы парадокса Кондорсе обсуждаются в связи с реалиями выборов и референдумов, иных форм современной деловой и общественной жизни.

1.4. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

Всем опытным управленцам хорошо известно, что один из наиболее эффективных интеллектуальных инструментов менеджера — это теория принятия решений. Подробно разобранный пример выбора типа автомобиля для запуска в серию наглядно демонстрирует ряд основных понятий теории принятия решений. Обсудим эти понятия в связи с организацией работы экспертной комиссии.

Кто принимает решения? Решение о выборе того или иного типа автомобиля для запуска в серию принимал совет директоров фирмы «Русские автомобили» большинством голосов. Однако в подготовке решения участвовали и другие люди — специалисты, подготовившие информацию, приведенную в табл. 1.1.

В теории принятия решений есть специальный термин — Лицо, принимающее решения (ЛПР). Это тот, на ком лежит ответственность за принятое решение, тот, кто подписывает приказ или иной документ, в котором выражено решение. Обычно это генеральный директор или председатель правления фирмы, командир воинской части, мэр города и т.п., словом, ответственный работник. Но иногда действует коллективное ЛПР, как в случае с советом директоров фирмы «Русские автомобили» или Государственной Думой.

Проект решения готовят специалисты, как говорят, «аппарат ЛПР», часто вместе с сотрудниками иных организаций. Если ЛПР доверяет своим помощникам, то может даже не читать текст, а просто подписать его. Но ответственность все равно лежит на ЛПР, а не на тех, кто участвовал в подготовке решения.

При практической работе важно четко отделять этап дискуссий, когда рассматриваются различные варианты решения, от этапа принятия решения, после которого надо решение выполнять, а не обсуждать.

В рассмотренном примере коллективное ЛПР включало в себя экспертов Воробьева, Лебедева, Чибисова и Куликова и членов совета директоров Волкова, Вепрева и Лисицына. Положительной стороной такого организационного решения является личная материальная и моральная заинтересованность перечисленных экспертов в принятии правильных решений. То же самое обстоятельство является и отрицательной стороной. Независимые эксперты часто предпочтительнее. Однако им гораздо труднее войти в реалии конкретной фирмы, чем менеджерам и акционерам этой фирмы.

Порядок подготовки решения (регламент). Часты конфликты между менеджерами по поводу сфер ответственности — кто за что отвечает, кто какие решения принимает. Поэтому очень важны регламенты, определяющие порядок работы. Недаром любое собрание принято начинать с утверждения председательствующего, секретаря и повестки заседания, а работу любого предприятия или общественного объединения — с утверждения его устава. Влияние регламента на результаты принятия решений показано ранее при обсуждении процедур голосования.

Цели. Каждое решение направлено на достижение одной или нескольких целей. Например, совет директоров фирмы «Русские автомобили» желал:

- продолжать выполнять миссию фирмы, т.е. выпуск автомобилей;
- получать максимальную возможную прибыль (в условиях неопределенности будущих цен на бензин).

Эти две цели можно достичь одновременно. Однако так бывает не всегда.

Например, часто встречающаяся формулировка «максимум прибыли при минимуме затрат» внутренне противоречива. Минимум затрат равен 0, когда работа не проводится, но и прибыль тогда тоже равна 0. Если же прибыль велика, то и затраты велики, поскольку и то и другое связано с объемом производства. Можно либо максимизировать прибыль при фиксированных затратах, либо минимизировать затраты при заданной прибыли, но невозможно добиться «максимума прибыли при минимуме затрат».

Одной и той же цели можно, как правило, добиться различными способами. Например, миссия фирмы «Русские автомобили» будет осуществляться и при выпуске машин типа «Алеша», и при выпуске «Добрыни».

Ресурсы. Каждое решение предполагает использование тех или иных ресурсов. Так, совет директоров фирмы «Русские автомобили» исходит из существования производства (системы предприятий), позволяющего выпускать автомобили типа «Алеша» и типа «Добрыня» (причем быстрый переход с выпуска одного типа автомобиля на выпуск другого типа невозможен по технологическим причинам). Если бы такого производства не было, то и дискуссия в совете директоров не имела бы смысла. Конечно, можно было бы сначала обсудить вопрос о строительстве заводов, о посильности таких затрат для фирмы...

Кроме того, предполагается, что у фирмы достаточно финансовых средств, материальных и кадровых ресурсов для массового выпуска автомобилей и того и другого типа. Ведь надо сначала подготовить производство и работников, закупить сырье и комплектующие изделия, произвести и реализовать продукцию и только потом получить прибыль (как разность между доходами и расходами). Для подготовки различных разделов бизнес-планов постоянно привлекаются эксперты.

В повседневной жизни мы чаще всего принимаем решения, покупая товары и услуги. И тут совершенно ясно, что такое ресурсы — это количество денег в нашем кошельке.

При практической работе над проектом решения важно все время повторять: «Чего мы хотим достичь? Какие ресурсы мы готовы использовать для этого?»

Эти соображения весьма важно учитывать рабочей группе для правильной ориентации экспертов, для обеспечения четкости постановок задач перед ними.

Риски и неопределенности. Почему четверо выступавших членов совета директоров разошлись во мнениях? В частности, потому что они по-разному оценивали риск повышения цен на бензин, влияние этого риска на успешность достижения цели.

Многие решения принимаются в условиях риска, т.е. при возможной опасности потерь. Связано это с разнообразными неопределенностями, окружающими нас. Кроме отрицательных (нежелательных) неожиданностей, бывают положительные — мы называем их удачами. Менеджеры стараются застраховаться от потерь и не пропустить удачу.

Внутренне противоречива формулировка: «Максимум прибыли и минимум риска». Обычно при возрастании прибыли возрастает и риск — возможность потерять многое или все.

Вернемся к табл. 1.1. Неопределенность не только в том, будет цена на бензин высокой или низкой. *Неопределенности — во всех числах этой таблицы*. Шансы низкой цены на бензин оценены в 60%. Этот прогноз, очевидно, не может быть абсолютно точным. Вместо 60% следовало бы поставить, скажем, (60 ± 3)%. Тем более содержат неустранимые неточности данные о предполагаемой прибыли. Ведь для того чтобы ее рассчитать, необходимо оценить:

- затраты на подготовку производства и выпуск продукции (это можно сделать достаточно точно, особенно при отсутствии инфляции);
- число будущих покупателей в зависимости от цены и установить оптимальную цену, обеспечивающую максимальную прибыль (отделу маркетинга сделать это достаточно трудно, хотя бы потому, что промежуточным этапом является прогноз социально-экономического развития страны, из которого вытекают финансовые возможности и предпочтения потребителей, размеры налогов и сборов и т.п.).

В результате вместо 1000 в таблице должно стоять, скажем, 1000 ± 200. Следовательно, рассуждения четырех членов совета директоров, опирающихся на числа из табл. 1.1, строго говоря, некорректны. Реальные числа иные, хотя и, надеемся, довольно близкие. Необходимо изучить устойчивость выводов по отношению к допустимым отклонениям исходных данных, а также по отношению к малым изменениям предпосылок используемой математической модели. Речь идет об общеинженерной и общеэкономической идее: любое измерение проводится с некоторой погрешностью, и эту погрешность необходимо указывать.

Оценка и анализ рисков, разработка рекомендаций по управлению ими — достойные задачи для экспертной комиссии. Могут быть полезны самые разные экспертные технологии, в том числе использующие «мозговой штурм».

Критерии оценки решения. Вспомним еще раз дискуссию в совете директоров фирмы «Русские автомобили». Каждый из выступавших использовал свой критерий для выбора наилучшего варианта решения. У каждого эксперта также может быть свой собственный критерий, отличный от критериев других экспертов, и организаторам дискуссии придется приложить немало усилий добиваясь, чтобы эксперты смотрели на проблему с близких позиций.

Воробьев предлагал исходить из наихудшего случая высокой цены бензина. Фактически он рассматривал внешний (для фирмы) мир как врага, который всячески будет стараться уменьшить прибыль фирмы. И в условиях жесткого противодействия со стороны внешнего мира он предлагал выбрать наиболее выгодный вариант решения — выпуск «Алеши». Подход Воробьева хорош при рассмотрении совершенно бескомпромиссного противостояния двух противников, имеющих противоположные интересы, например двух армий воюющих между собой государств. Как уже отмечалось, существует математизированная наука — так называемая «теория игр», в которой рассматриваются методы оптимального поведения в условиях антагонистического или иного конфликта. В дискуссии о выборе типа автомобиля для запуска в серию позиция Воробьева — это позиция крайнего пессимиста, поскольку нет оснований считать внешний мир активным сознательным противником фирмы. Отметим также, что наиболее плохой случай, на который ориентируется теория игр, встречается сравнительно редко (согласно табл. 1.1 — в 40% случаев). Вместе с тем подход теории игр достаточно популярен в литературе [61].

Подход оптимиста Лебедева прямо противоположен подходу Воробьева. Предлагается исходить из самого благоприятного стечения обстоятельств. Внешний мир для Лебедева — друг, а не враг. И надо сказать, что для такой позиции есть основания — низкая цена на бензин в полтора раза вероятнее высокой. С точки зрения теории планирования предложение Лебедева можно было бы взять за основу, добавив возможности коррекции плана в случае неблагоприятных обстоятельств, а именно повышения цены на бензин. И тут мы наталкиваемся на неполноту дискуссии в совете директоров — никто не рассмотрел принципиальную возможность подготовки производственной программы «двойного назначения». Выполнение такой программы обеспечивало бы гибкость управления — при низкой цене на бензин был бы налажен выпуск «Добрыни», а при высокой — «Алеши». В частности, такую гибкость обеспечивало бы повышение стандартизации автомашин фирмы, использование в них одних и тех же узлов и деталей, применение для их изготовления одних и тех же технологических процессов. Технологиям сравнения инвестиционных проектов и разработки бизнес-планов посвящено большое число пособий, например [52].

С чисто логической точки зрения оптимизм Лебедева не менее и не более оправдан, чем пессимизм Воробьева. Люди вообще и менеджеры в частности подразделяются на два типа — оптимистов и пессимистов. Особенно четко различие проявляется при вложении капитала, поскольку, как правило, увеличение прибыли связано с увеличением риска. Одни люди предпочтут твердый доход (да еще и застрахуются), отказавшись от соблазнительных, но рискованных предложений. Другой тип людей — оптимисты и авантюристы, они уверены, что им повезет. Такие люди надеются разбогатеть, играя в лотерею.

Надо иметь в виду, что на фирму, как и на отдельного человека, выигрыш или проигрыш одной и той же суммы могут оказать совсем разное влияние. Выигрыш приносит радость (но не счастье), в то время как проигрыш может означать разорение, полный крах, т.е. несчастье. Недаром в микроэкономической теории полезности рассматривают парадоксальное понятие — «полезность денег» — и приходят к выводу, что полезность равна логарифму имеющейся суммы.

Вернемся к совету директоров фирмы «Русские автомобили». Совсем с других позиций, чем Воробьев и Лебедев, подошел к делу Чибисов. Его подход фактически предполагает, что придется много раз принимать решения по аналогичным вопросам. Вот он и рассчитывает средний доход, исходя из того, что в 60% случаев цена бензина будет низкой, а в 40% случаев — высокой. Такой подход вполне обоснован, когда выбор технической политики проводится каждую неделю или каждый день. Например, к нему мог бы прибегнуть менеджер, проектирующий свой ресторан, — ориентироваться ли на открытые столики с видом на живописные окрестности или замкнуться в четырех стенах, отгородившись от дождя. Если события происходят много раз, то для принятия решений естественно использовать методы современной прикладной статистики [77] и эконометрики [89], как это делают, например, при статистическом контроле качества продукции и сертификации [87]. Тогда оценка математического ожидания дохода, проведенная Чибисовым, вполне корректна.

Однако совет директоров фирмы «Русские автомобили» решает вопрос об одном-единственном выборе. Поэтому 60 и 40% — это не вероятности как пределы частот, что обычно предполагается при применении теории вероятностей, это шансы низкой и высокой цены бензина (иногда употребляют термин «субъективные вероятности»). Использование подобных шансов полезно, чтобы в одном критерии свести вместе пессимистический и оптимистический подходы. Но ссылаться на всем известное практическое значение теории вероятностей в данном случае не приходится.

Четвертый оратор, Куликов, вводит в обсуждение новый критерий — «упущенная выгода». Обратите внимание: средний доход, рассчитанный Чибисовым, больше при выпуске «Добрыни», а упущенная выгода, наоборот, меньше при выпуске «Алеши». Эти два критерия в данном случае противоречат друг другу.

Каждому менеджеру, в том числе при разработке сценария экспертного исследования, приходится решать, какой из критериев для него важнее. В этом ему может помочь теория полезности, хорошо разработанная в экономике (в частности, так называемая «маржинальная полезность» в теории поведения потребителей и т.п.) и имеющая развитый математический аппарат [128].

Ряд примеров оптимизационных задач принятия решений рассмотрен далее.

1.5. СОВРЕМЕННЫЙ ЭТАП РАЗВИТИЯ ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

Теория принятия решений — быстро развивающаяся наука. Задачи, которыми она занимается, порождены практикой управленческих решений на различных уровнях — от отдельного подразделения или малого предприятия до государств и международных организаций. Кратко рассмотрим несколько проблем, активно обсуждающихся на современном этапе развития теории принятия решений. Это — системный подход при принятии решений, современные методы принятия решений и проблема горизонта планирования.

Системный подход при принятии решений. При обсуждении проблем принятия решений часто говорят о системном подходе, системе, системном анализе, системном управлении. Весьма важно учитывать требования системного подхода при проведении экспертного оценивания. Речь идет о том, что надо рассматривать проблему в целом, а не «выдергивать» для обсуждения какую-нибудь одну черту, хотя и важную.

Так, при массовом жилищном строительстве можно «выдернуть» черту (критерий оценки решения) — стоимость квадратного метра в доме. Тогда наиболее дешевые дома — пятиэтажки. Если же взглянуть системно, учесть стоимость транспортных и инженерных коммуникаций (подводящих электроэнергию, воду, тепло и др.), то оптимальное решение уже другое — девятиэтажные дома. В центрах мегаполисов, где весьма высока стоимость земельных участков, этажность должна быть еще выше.

Другой пример. Менеджер банка, отвечающий за распространение пластиковых карт, может сосредоточиться на рекламе. Между тем ему от системы «банк — владельцы карт» лучше перейти к системе «банк — руководители организаций — владельцы карт». Договоренность с руководителем учреждения, давшим в итоге приказ выплачивать заработную плату с помощью пластиковых карт, принесет менеджеру гораздо больший прирост численности владельцев карт, чем постоянная дорогая реклама. Его ошибка состояла в неправильном выделении системы, с которой он должен работать. Правильное решение могло быть найдено в результате применения экспертной технологии «мозгового штурма».

Менеджер банка будет не прав, оценивая работу подразделений банка в текущих рублях. Обязательно надо учитывать инфляцию [89]. Иначе мы сталкиваемся с парадоксальными явлениями, когда реальная ставка платы за кредит отрицательна; или же — рублевый оборот растет, банк якобы процветает, а после перехода к сопоставимым ценам путем деления на индекс инфляции становится ясно, что дела банка плохи.

Различных определений понятия «система» — десятки. Общим в них является то, что о системе говорят как о множестве, между элементами которого имеются связи. Целостность системы и ее «отделенность» от окружающего мира обеспечиваются тем, что взаимосвязи внутри системы существенно сильнее, чем связь какого-либо ее элемента с любым элементом, лежащим вне системы. По определению академика РАН Н.Н. Моисеева: «Системный анализ — это дисциплина, занимающаяся проблемами принятия решений в условиях, когда выбор альтернативы требует анализа сложной информации различной природы» [55].

Современные методы принятия решений. Кроме упомянутых или кратко рассмотренных ранее методов, прежде всего экспертных, при принятии решений применяют весь арсенал методов современной прикладной математики. Они используются для оценки ситуации и прогнозирования, при выборе целей, для генерирования множества возможных вариантов решений и выбора из них наилучшего [87].

Прежде всего надо назвать всевозможные методы оптимизации (математического программирования). Для борьбы с многокритериальностью используют различные методы свертывания критериев, а также интерактивные компьютерные системы, позволяющие вырабатывать решение в процессе диалога человека и ЭВМ. Применяют имитационное моделирование, базирующееся на компьютерных системах, отвечающих на вопрос: «Что будет, если...?», метод статистических испытаний (Монте-Карло), модели надежности и массового обслуживания. Часто необходимы статистические (эконометрические) методы, в частности методы выборочных обследований.

Особого внимания заслуживают проблемы неопределенности и риска, связанные как с природой, так и с поведением людей. Разработаны различные способы описания неопределенностей: вероятностные модели, теория нечеткости, интервальная математика. Для описания конфликтов (конкуренции) полезна теория игр. Для структуризации рисков используют деревья причин и последствий (диаграммы типа «рыбий скелет»). Менеджеру важно учитывать постоянные и аварийные

экологические риски. Плата за риск и различные формы страхования также постоянно должны быть в его поле зрения.

Экспертные технологии, основанные на сборе и анализе мнений специалистов, занимают центральное место как в теории, так и в практике разработки и принятия управленческих решений.

Словесные, графические, математические, компьютерные модели социально-экономических явлений и процессов — база для разработки конкретных методов подготовки решений.

Необходимо подчеркнуть, что весьма полезны и различные «простые приемы» принятия решений [59]. Например, при сравнении двух возможных мест работы весьма помогает таблица из трех столбцов. В левом из них перечислены характеристики рабочего места: заработок, продолжительность рабочего времени, время в пути от дома до работы, надежность предприятия, возможности для профессионального роста, характеристики рабочего места и непосредственного начальства и т.п. А в двух других столбцах — оценки этих характеристик для двух возможных мест работы, в «натуральных» показателях или в процентах от максимума. Иногда при взгляде на подобную таблицу все сразу становится ясно. Но можно вычислить значения обобщенного показателя, введя весовые коэффициенты и сложив взвешенные оценки вдоль столбцов. Не менее полезно изобразить на бумаге возможные варианты решения, которое предстоит принять, а также возможные реакции лиц и организаций на те или иные варианты решения, а затем и возможные ответы на эти реакции. Полезны таблицы доводов «за» и «против» и др. Простые приемы принятия решений обсуждаются в главе 2.

Проблема горизонта планирования и контроллинг. Отнюдь не все проблемы теории принятия решений решены, многие ее разделы продолжают бурно развиваться. Приведем примеры.

Во многих ситуациях продолжительность проекта не определена либо горизонт планирования инвестора не охватывает всю продолжительность реализации проекта до этапа утилизации. В таких случаях важно изучить влияние горизонта планирования на принимаемые решения. Эта проблема обсуждается в [87].

В последние годы все большую популярность получает *контроллинг* — современная концепция системного управления организацией, в основе которой лежит стремление обеспечить ее долгосрочное эффективное существование [32; 248]. В конкретных прикладных работах успех достигается при комбинированном применении различных методов. В контроллинге активно используются эконометрические методы [90], прежде всего технологии экспертных исследований. Для подготовки решений создаются аналитические центры и «ситуационные комнаты», позволяющие соединять человеческую интуицию и компьютерные расчеты. Все шире используются в контроллинге информационные технологии поддержки принятия решений.

Теория принятия решений оценок — развитая научная и практическая дисциплина с большим числом подходов, идей, алгоритмов, теорем и способов их практического использования. Однако необходимо подчеркнуть: *менеджер отвечает за принятие решений и не имеет права перекладывать ответственность на специалистов.*

Контрольные вопросы

1. Какой из критериев принятия решения, высказанных на заседании совета директоров фирмы «Русские автомобили» Воробьевым, Лебедевым, Чибисовым и Куликовым, представляется вам наиболее естественным? Как бы вы сами поступили на месте совета директоров фирмы «Русские автомобили»?
2. Какой образец мотоцикла предпочтительней запустить в серию? Исходные данные для принятия решения приведены в таблице. Разберите четыре критерия принятия решения: пессимистичный, оптимистичный, средней прибыли, минимальной упущенной выгоды.

Прибыль фирмы при различном выборе образца мотоцикла для запуска в серию, млн руб.

Цена бензина	Мотоцикл «Витязь»	Мотоцикл «Комар»
Низкая (20%)	900	700
Средняя (60%)	700	600
Высокая (20%)	100	400

3. Чем отличаются правила голосования «Кто за?» и «Кто против?» Какова роль воздержавшихся в каждом из этих случаев?
4. В чем состоит парадокс Кондорсе? Какие формы (версии) парадокса вам известны?
5. Целесообразно ли, на ваш взгляд, купить 1000 билетов лотереи с целью разбогатеть?
6. В чем заключается системный подход к принятию решений на основе технологий экспертных оценок?

Темы докладов и рефератов

1. Теория конечных антагонистических игр и ее применение в экономике.
2. Теория статистических решений применительно к дискуссии на заседании совета директоров фирмы «Русские автомобили».
3. Различные методы организации голосования в малых группах (с использованием результатов научных исследований, приведенных в монографии [12]).
4. Роль теории принятия решений в неформальной информационной экономике будущего.

ГЛАВА 2
ПРОСТЫЕ МЕТОДЫ ПРИНЯТИЯ РЕШЕНИЙ

2.1. НЕКОТОРЫЕ МЕТОДЫ ПРИНЯТИЯ РЕШЕНИЙ В СТРАТЕГИЧЕСКОМ МЕНЕДЖМЕНТЕ

Простые методы принятия решений не требуют применения развитого математического аппарата. Тем не менее во многих важных случаях их применения вполне достаточно.

Рассмотрим несколько широко используемых практических инструментов принятия решений на примере задач стратегического менеджмента [134].

Информация и инструменты стратегического планирования. Исходными данными стратегического планирования являются:

- структура конкурентов;
- структура рынков сбыта;
- тенденции технического развития и эволюции моды;
- структура рынков снабжения;
- правовая, социальная, технологическая, экономическая, экологическая и политическая окружающая среда;
- собственные сильные и слабые стороны.

На основе перечисленных данных в соответствии с миссией фирмы выбираются цели на длительную перспективу и анализируются ресурсы, которые для этого необходимы. Инструментами стратегического планирования являются, кроме использованного в главе 1 метода экспертных оценок, анализ «разрывов», анализ шансов и рисков (сильных и слабых сторон), анализ портфеля, метод проверочного списка, метод оценки по системе баллов, концепция жизненного цикла товара, а также иные методы прогнозирования, планирования и принятия решений.

При анализе «разрывов» сравнивают три возможных сценария развития фирмы:

- какого оборота (прибыли и других характеристик работы предприятия) можно достичь, если в будущем в процессе продаж ничего не изменится (сценарий А);
- какого оборота можно достичь, если попытаться при максимальном напряжении сил с существующим продуктом занять

возможно большую нишу на существующих рынках (сценарий Б);
- если дополнительно (к сценарию Б) развивать новые продукты и (или) новые рынки (сценарий В).

Разницу между результатами по сценариям Б и А называют *оперативным разрывом*, а между результатами по сценариям В и Б — *стратегическим разрывом*. Эта терминология подчеркивает роль нововведений в стратегическом плане фирмы — разработки новых продуктов или выхода на новые рынки либо и того и другого вместе.

Матрица портфеля Бостонской консалтинговой группы. В стратегическом планировании может оказаться полезным анализ портфеля предприятия (табл. 2.1). Надо подчеркнуть, что речь идет о стратегическом планировании не для всего предприятия, а для его «стратегических подразделений». Они выделяются комбинациями «продукт—рынок», которые:

- однородны, т.е. нацелены на определенный достаточно однородный круг потребителей;
- могут действовать независимо от других подразделений предприятия;
- распоряжаются достаточно большой долей рынка, чтобы проведение исследований по разработке специфической стратегии было выгодным.

Таблица 2.1

Матрица портфеля Бостонской консалтинговой группы

Рост спроса	Рыночная доля	
	высокая	низкая
Высокий	1. Звезды	3. Знак вопроса
Низкий	2. Дойные коровы	4. Собаки

Внеся товары (с учетом их доли в обороте фирмы) в соответствующие клетки табл. 2.1, можно рассчитать долю особо успешных товаров типа 1 (Звезды), которые, возможно, нуждаются в дальнейшем финансировании для увеличения и закрепления успеха. Хотя рост спрос на товары типа 2 (Дойные коровы) низок, но из-за большой доли рынка они могут еще долго приносить хороший доход на мало меняющихся (стагнирующих) рынках. Судьба товаров типа 3 (Знак вопроса) неясна. Оправданы ли большие финансовые затраты на расширение их доли на рынке? Товары типа 4 (Собаки) «зарабатывают» лишь себе на жизнь.

На основе табл. 2.1 можно проанализировать несколько возможных стратегий:

- «строить», т.е. «Знаки вопроса» переводить в «Звезды»;

- «держать», т.е. «Дойные коровы» должны удерживать свои доли рынка и стремиться к росту прежде всего для поддержки «Звезд» и «Знаков вопроса»;
- «собирать урожай», т.е., не принимая во внимание долгосрочные последствия, снимать сиюминутные сливки (при этом идет речь о «слабых» — «Дойных коровах», «Собаках» и «Знаках вопроса»);
- «выселяться», т.е. «Собаки» и «Знаки вопроса» забираются с рынка (перестают выпускаться), поскольку они ничего не приносят и не ожидается их рост, и т.д.

При определении целей и стратегий дальнейшего развития стратегические подразделения нуждаются во взаимной координации, однако без подавления их самобытности (другими словами, со стороны руководства фирмы должно осуществляться контролируемое децентрализованное руководство). Руководство фирмы должно направить отдельные подразделения на привлекательные рынки, обнаружить и использовать синергетический эффект от их взаимодействия и рационально распределить ресурсы. Так, руководство фирмы должно способствовать тому, чтобы «Дойные коровы» передали часть дохода «Звездам».

В таблице 2.1 сопоставлены такие характеристики выпускаемого товара, как «рост спроса» и «рыночная доля». Ясно, что высокий рост соответствует ранней стадии жизненного цикла товара, а низкий — поздней. Обычно высокая доля рынка сигнализирует о продолжительном периоде получения прибыли, а низкая — о коротком. Так, высокая доля рынка может быть из-за слабой конкуренции. Рыночный лидер может иметь преимущество в издержках на одно изделие — эффект масштаба производства!

Методы проверочного списка и суммарной оценки. Широко используемыми и весьма полезными инструментами стратегического планирования являются также метод проверочного списка и метод суммарной оценки (по системе баллов). Первый метод весьма прост. Выделяется некоторое количество «факторов успеха» и всем рассматриваемым проектам даются оценки (например, с помощью комиссии экспертов) по этим факторам. Например, в табл. 2.2 представлен бланк проверочного списка для проектов, состоящих в организации выпуска тех или иных товаров (стратегии типа «продукт—рынок»).

Обратите внимание, что оценки даются в качественном виде (измерены в порядковой шкале — см. о шкалах в главе 9). Любая количественная определенность была бы при подобных оценках лишь иллюзией.

Таблица 2.2

Пример проверочного списка

Факторы	Продукты		
	А	Б	В
Степень инноваций	«Хорошо»	«Средне»	«Плохо»
Число возможных покупателей	«Плохо»	«Хорошо»	«Средне»
Готовность к кооперации в торговле	«Средне»	«Хорошо»	«Хорошо»
Барьеры для вхождения новых продавцов	«Хорошо»	«Плохо»	«Плохо»
Обеспеченность сырьем	«Плохо»	«Средне»	«Хорошо»

Целесообразно разделить факторы на «обязательные», «необходимые» и «желательные», т.е. ввести веса факторов, выраженные в качественном виде. Правило принятия решения может иметь вид: «Форсируй планирование тех стратегий типа «продукт—рынок», при которых все обязательные факторы, и по меньшей мере два необходимых, соответствуют оценке «хорошо».

Методу проверочного списка, в котором как оценки отдельных факторов, так и веса факторов и способы принятия решений имеют качественный характер, соответствует количественный двойник — метод суммарной оценки.

Конечно, с числами оперировать гораздо легче, чем с качественными оценками. Недаром математики обычно рвутся «оцифровать» качественные факторы и веса. Но при этом, как мы знаем из теории измерений (см. главу 9), в окончательные выводы может быть внесен субъективизм, связанный с выбором способа «оцифровки» качественных оценок и весов. Обратите внимание в связи со сказанным на обсуждение методов принятия решений, основанных на применении оценок экспертов (часть III, в которой, в частности, даны рекомендации по снижению субъективизма в выборе весов факторов в единой суммарной оценке).

Рассмотрим условный пример по вычислению и использованию единой суммарной оценки. Пусть оценки факторов 1 и 2 для продуктов А и Б даны в табл. 2.3 (для простоты изложения мы опускаем способы получения численных значений в табл. 2.3 и не рассматриваем погрешности этих значений).

Для получения суммарной оценки необходимо знать веса факторов. Пусть фактор 1 оценивается экспертами как вдвое более важный, чем фактор 2. Поскольку сумма весов факторов должна составлять 1, то вес фактора 1 есть 0,67, а фактора 2 — 0,33.

Таблица 2.3

Метод суммарной балльной оценки

Факторы	Доля продуктов, %	
	А	Б
1	40	90
2	50	20

Суммарная оценка по продукту А равна

$$0{,}67 \times 40\% + 0{,}33 \times 50\% = 26{,}8\% + 16{,}5\% = 43{,}3\%,$$

а суммарная оценка по продукту Б —

$$0{,}67 \times 90\% + 0{,}33 \times 20\% = 60{,}3\% + 6{,}6\% = 66{,}9\%.$$

Однако получение суммарных оценок — только этап процесса принятия решений. Нужен еще один критерий отбора — какими продуктами заниматься, а какими нет. Простейшая формулировка состоит в задании границы. Если суммарная оценка продукта больше этой границы, то связанная с ним работа по планированию продолжается, если же нет — он исключается из рассмотрения как малоперспективный. Если в рассматриваемом случае такая граница выбрана на уровне 55%, то работа над продуктом А прекращается, а над продуктом Б — продолжается.

Принятие решения на основе границы несколько снижает влияние конкретных правил оцифровки. Например, если для продукта А оценки по факторам А и Б поднимутся на 10% и достигнут соответственно значений 50 и 60%, то суммарная оценка окажется равной

$$0{,}67 \times 50\% + 0{,}33 \times 60\% = 33{,}5\% + 19{,}8\% = 53{,}3\%,$$

т.е. общее решение не меняется, продукт А остается среди малоперспективных.

Менеджер — главное лицо в перспективном планировании. Если прогнозирование — научно-исследовательская работа, ее результаты можно сравнить с прожектором, освещающим основные черты грядущего, то планирование [130] — частный вид принятия решений. Для стратегического планирования и управления могут быть использованы не только те методы подготовки и принятия решений, о которых говорится в настоящей главе, но и весь арсенал современной теории принятия решений.

Однако все эти простые или хитроумные компьютерные приемы — лишь подспорье для менеджера. Именно он несет ответственность за судьбу фирмы, и именно на свое знание дела, на свою интуицию он должен полагаться при принятии решений в стратегическом менеджменте [50; 134].

2.2. ОПЕРАТИВНЫЕ ПРИЕМЫ ПРИНЯТИЯ РЕШЕНИЙ

При обсуждении проблем стратегического менеджмента был рассмотрен ряд оперативных приемов принятия решений — анализ «разрывов», анализ шансов и рисков (сильных и слабых сторон), анализ портфеля, метод проверочного списка, метод оценки по системе баллов и др. Такие методы хорошо применять при быстром сравнении вариантов, например, на совещании менеджеров высшего звена.

Рассмотрим в качестве примера матрицу портфеля Бостонской консалтинговой группы. Согласно этому методу подготовки управленческих решений товары, выпускаемые фирмой, распределяются по клеткам табл. 2.1. Однако совершенно ясно, что такое распределение может служить лишь отправной точкой для дальнейшего анализа.

Действительно, необходимо опираться на данные о прибыли и рентабельности тех или иных товаров. Ясно, например, что высокий рост спроса на товар типа «Знак вопроса» может быть обеспечен демпинговой ценой ниже себестоимости.

Необходимо оценить динамику смены марок товаров, понять, насколько долго смогут удержаться на рынке «Дойные коровы», насколько высоко смогут взлететь «Звезды».

Специального рассмотрения заслуживают «Собаки». Возможно, они вытесняются другими товарами. Но возможно и иное — их покупатели представляют собой отдельный рынок, лишь из-за недостатков предварительного анализа присоединенный к общему рынку. Тогда постановка задачи меняется. Руководство фирмы не должно сравнивать «Собак» с другими товарами. Ему следует решить совсем иной вопрос — обслуживать ли сравнительно небольшой рынок покупателей «Собак» или же отдать его конкурентам.

Бесспорно, что обоснованные решения не могут приниматься на основе только анализа матрицы портфеля Бостонской консалтинговой группы. Впрочем, это верно и для любого иного метода подготовки решений. Только всесторонний анализ с использованием многих методов может дать руководству организации необходимые аргументы для принятия обоснованного решения. Но и в этом случае ответственность лежит на ЛПР — на менеджерах.

Оперативных приемов принятия решений, или, в другой терминологии, простых методов принятия решений, существует весьма много. Один из них — изложить ситуацию в письменном виде. Эта простая рекомендация часто оказывается весьма полезной. Дело в том, что при составлении описания приходится уточнять многие факты и оценки, которые обычно не удается сопоставить при размышлениях. Далее,

письменное описание подсказывает различные альтернативы действий, а также оценки последствий этих альтернатив. Изложение ситуации в письменном виде во многом снимает эмоциональную составляющую при принятии решения, а также дает исходные данные и варианты действий для аналитического разбора.

Иногда рекомендуют проводить первичную формализацию описания ситуации, например, в виде ответов на следующие вопросы:

1. Совместим ли рассматриваемый вариант решения с моими нравственными принципами?
2. Что я выиграю при этом варианте решения?
а) деньги;
б) время;
в) известность;
г) уверенность;
д) удовольствие и т.д.
3. Что я потеряю при таком решении?
а) деньги;
б) время и т.д. (см. вопрос 2).
4. Какие новые возможности у меня появятся?
5. Какие новые задачи встанут передо мной?
6. Какие обязанности у меня появятся?
7. Какая новая ситуация для меня возникнет?
8. Каких побочных действий я должен ожидать?
а) положительных;
б) отрицательных.
9. Принесу ли я вред обществу или другим людям?
10. Принесу ли я пользу обществу или другим людям?
11. Возникнут ли в результате моего решения новые проблемы?
12. Потребуются ли новые решения? И т.д.

Можно выделить следующие этапы анализа ситуации, подготовки и принятия решения, анализа последствий [59]:

1) уяснение ситуации;
2) установление наличия проблемы, подлежащей решению;
3) формирование возможных решений;
4) описание последствий решений;
5) выбор решения;
6) обобщение накопленного опыта принятия решений.

Целесообразно уточнить содержание каждого из перечисленных этапов. Например, для уяснения ситуации целесообразно ответить на пять вопросов:

1. КТО должен или обязан (или хочет) принять решение?

2. ГДЕ (в каком месте, в каком окружении, в какой среде, при каких обстоятельствах) предстоит принимать решение?

3. КОГДА (до какого срока или насколько часто, с какой периодичностью) необходимо принимать решение?

4. КАК (каким образом, в какой форме, каким документом) должно быть выражено решение?

5. ЧТО обусловливает решение? (Зачем оно нужно? В чем его цель? Какой замысел лежит в его основе? Для чего оно служит? Зачем его надо принимать?)

После того как ситуация обдумана, необходимо рассмотреть варианты решений. Рассмотрим примеры.

Пример 2.1. На столе у секретаря начальника звонит телефон. Звонящий задает вопрос по делам фирмы, но такой, на который не могут ответить ни секретарь, ни ее начальник. Как должна реагировать секретарь? И какой реакции звонящего следует ожидать?

Реакция секретаря № 1. Она объясняет звонящему, что не может сообщить необходимые сведения, и соединяет его с нужным сотрудником.

Реакция звонящего № 1. Он будет признателен секретарю за то, что его быстро соединили с человеком, который может проинформировать его компетентно и с достаточной полнотой.

Реакция секретаря № 2. Она просит звонящего подождать у аппарата и бежит через все здание, чтобы получить нужную ему информацию.

Реакция звонящего № 2. Он будет раздражен, поскольку вынужден бессмысленно прождать длительное время у телефона, чтобы в конце концов узнать, что информации, которую ему здесь сообщили, недостаточно.

Побочный результат. В течение длительного времени телефон руководства фирмы будет занят.

Реакция секретаря № 3. Она адресует звонящего к начальнику, который, естественно, также не сможет ему помочь.

Реакция звонящего № 3. Он будет раздражен, поскольку вынужден провести телефонные разговоры с двумя сотрудниками фирмы и не получить нужной ему информации.

Побочный результат. Тот же, что и в предыдущем случае.

Реакция секретаря № 4. Она возвращает звонящего к коммутатору фирмы, так как не может быть ему полезной.

Реакция звонящего № 4. Он и на этот раз будет раздражен, так как только потерял время.

Очевидно, только первый вариант решения можно признать правильным. Отметим, однако, что для его реализации в распоряжении секретаря должны быть соответствующие технические средства, позволяющие перевести телефонный вызов на номер нужного сотрудника.

В рассмотренном примере сравнение вариантов решения нетрудно провести непосредственно. Однако в большинстве задач принятия решений целесообразно выделить перечень факторов, на основе значений которых и целесообразно сравнивать варианты решений.

Пример 2.2. Петя Иванов оканчивает МГТУ им. Н.Э. Баумана и выбирает место работы. У него есть четыре варианта.

А. Поступить в аспирантуру МГТУ им. Н.Э. Баумана. Стипендия ничтожна, но есть возможности для подработки. Лет через пять можно стать доцентом всемирно известного вуза, работать по совместительству преподавателем, консультантом, сотрудником той или иной фирмы.

Б. Пойти инженером на крупное предприятие, ранее входившее в военно-промышленный комплекс (ВПК), а ныне имеющее постоянный пакет заказов, в том числе зарубежных.

В. Стать сотрудником малого предприятия, выполняющего конкретные заказы, и получать оплату с каждого выполненного заказа.

Г. Пойти компьютерщиком в филиал зарубежной экспертно-импортной фирмы.

Как сравнивать эти варианты? Рассмотрим естественные факторы.

Оплата труда. На настоящий момент нарастает от А до Г.

Перспективы роста (*в том числе оплаты*). Наиболее велики в А, имеются в Б, практически отсутствуют в В и Г.

Устойчивость рабочего места. Наибольшая в А, значительная в Б, малая в В и Г.

Начальство. Знакомое и уважаемое в А, солидное и хмурое в Б, несерьезное, но активное в В, строгое и малопонятное в Г.

Коллектив. Знакомый и приемлемый в А, понятный и благожелательный в Б, конкурентный (борьба за заказы и тем самым за доходы) в В, пропитанный стукачеством в Г.

Криминальность. Отсутствует в А и Б, постоянна (хотя и сравнительно мелкая) в В, возможна в Г (причем в крупный размерах).

Режим. Весьма свободный в А, жесткий (вход и выход по пропускам в заданное время) в Б, «полосатый» в В (вообще-то свободный, но если начальство прикажет...), тюремного типа в Г (фиксированы двери, через которые можно проходить, за питье чая на рабочем месте — штраф в размере 10% месячной оплаты и т.п.)

Время на дорогу до места работы. Ближе всего В, затем Г, А и Б.

Ограничимся этими восемью факторами. Для принятия решения целесообразно составить таблицу, в которой строки соответствуют факторам, столбцы — возможным вариантам решения, а в клетках таблицы стоят оценки факторов для соответствующих вариантов таблицы. Пусть для определенности в качестве возможных оценок используются числа 1, 2, ..., 10, причем наихудшее значение — это 1, а наилучшее — это 10. Пусть мнение Пети Иванова выражено в табл. 2.4.

Таблица 2.4

Оценки фактов при выборе места работы

№ п/п	Фактор	Место работы			
		МГТУ им. Н.Э. Баумана	крупное предприятие	малое предприятие	зарубежная фирма
1	Оплата труда	1	5	10	9
2	Перспективы роста	10	7	1	2
3	Устойчивость	10	9	3	4
4	Начальство	8	6	4	2
5	Коллектив	9	7	2	1
6	Криминал	10	8	1	2
7	Режим	10	4	7	1
8	Время на дорогу	5	3	10	7
9	Сумма баллов	63	49	37	28

Непосредственный анализ данных табл. 2.4 не позволяет Пете Иванову сделать однозначный вывод. По одним показателям лучше один вариант, по другим — другой. Надо как-то соизмерить факторы. Проще всего приписать им веса, а затем сложить веса для каждого из вариантов (с точки зрения теории измерений такой подход имеет недостатки, которые обсуждаются в главе 9). А какие веса взять? Проще всего считать все факторы равноценными, т.е. взять их с одинаковыми весами — единичными. Тогда следует сложить баллы, приписанные факторам. Результаты приведены в последней строке. По сумме баллов на первом месте — МГТУ им. Н.Э. Баумана, на втором — крупное предприятие, на третьем — малое предприятие, на последнем — филиал зарубежной фирмы.

Аналогичным образом проводится технико-экономический анализ в некоторых реальных ситуациях. Например, в табл. 2.5 дается сравнительная характеристика по факторам конкурентоспособности главных производителей изделий из стекловаты. Помимо непосредственного сравнения производителей, подобная таблица дает возможность подготовить решения по мерам повышения конкурентоспособности, а также указать возможные пределы продвижения. Так, согласно данным табл. 2.5 ОАО «Мостермостекло» по конкурентоспособности находится на уровне одного из своих основных конкурентов и проигрывает второму 4 балла. Однако, повысив удобство монтирования на 1 балл (и дойдя до уровня худшего из своих конкурентов по этому фактору), перейдя к более привлекательной системе скидок (набрав при этом 2 балла), а также усилив рекламные мероприятия на 2 балла (и дойдя до уровня худшего из своих конкурентов по этому фактору), оно увеличит сумму баллов на 5 и станет лучшим.

Таблица 2.5

Сравнительная характеристика главных производителей изделий из стекловаты по факторам конкурентоспособности

№ п/п	Факторы конкурентоспособности	ОАО «Мостермостекло»	Главные конкуренты	
			URSA	ISOVER
1	Товар			
1.1	Качество	5	5	5
1.2	ТЭП*	5	4	4
1.3	Престиж торговой марки	3	4	5
1.4	Кашировка**	5	5	5
1.5	Удобство монтирования	3	4	5
1.6	Наличие сертификатов	5	5	5
2	Цена			
2.1	Продажная цена	5	3	2
2.2	Скидки с цены	2	4	0
3	Продвижение товаров на рынках			
3.1	Реклама, участие в выставках и т.д.	2	5	4
	Общее количество баллов	35	39	35

* ТЭП — технико-экономическое планирование.

** Кашировка — дополнительное покрытие.

В практике разработки управленческих решений приходится иногда вводить веса факторов. Так, при подготовке организационно-экономического обеспечения реализации проекта установки газоочистного оборудования на ОАО «Магнитогорский металлургический комбинат» сравнивались четыре проекта (табл. 2.6).

Таблица 2.6

Балльная оценка проектов

№ п/п	Приведенные показатели качества	Проект				
		Россия-1	Россия-2	Украина	Швеция	Вес
1	Наработка на отказ	0,9125	0,975	0,9	1	7
2	Назначенный срок службы до списания	0,72	1	0,8	1	6
3	Назначенный срок службы до капитального ремонта	0,9	1	0,8	1	6
4	Среднее время восстановления	0,897	0,959	0,886	1	5
5	Установленный срок сохраняемости	1	1	0,667	0,667	4
6	Энергетические затраты на очистку 1 000 $м^3$ газа	0,852	0,958	0,852	1	9
7	Масса	0,886	0,972	0,875	1	8
8	Степень очистки	1	1	0,999	1	10
9	Полная стоимость проекта	0,877	1	0,860	0,662	9
10	Срок исполнения	0,8	1	0,667	1	7
Интегральный итоговый показатель качества проекта		***56,46***	***63,20***	***53,76***	***59,62***	

Проекты оценивались по «интегральному итоговому показателю качества проекта», равному сумме (по всем факторам) произведений значения фактора на вес этого фактора. Для таблиц 2.4 и 2.5 все веса были единичными, для табл. 2.6 веса приведены в правом столбце. (Значения весов обычно определяют с помощью экспертов.) В соответствии с «интегральным итоговым показателем качества проекта» наилучшим является проект «Россия-2», далее следует проект «Швеция», затем — проект «Россия-1», и замыкает четверку проект «Украина». В соответствии с рассматриваемым подходом надо рекомендовать принять к исполнению проект «Россия-2».

Много ценных рекомендаций по разработке управленческих решений содержится в книгах Б.Г. Литвака [42].

2.3. ДЕКОМПОЗИЦИЯ ЗАДАЧ ПРИНЯТИЯ РЕШЕНИЯ

Решать задачи по очереди. Естественным является желание разбить сложную задачу принятия решения на несколько, чтобы воспользоваться возможностью решать их по очереди.

Пример 2.3. Простейшим вариантом является дихотомическая схема для наглядного представления возможных решений [59]. Например, необходимо решить задачу: «Как встречать Новый год?» На первом шаге надо выбрать одно из двух возможных решений:

1) остаться дома;
2) уехать.

В каждом из двух случаев возникает необходимость принять решения второго уровня. Так, в первом случае:

1.1) пригласить гостей;
1.2) не звать гостей.

Во втором случае:

2.1) уехать к родственникам или знакомым;
2.2) уехать в общедоступные места (отправиться в путешествие, пойти в клуб или ресторан и т.п.).

После двух шагов получили четыре возможных решения. Каждое из них, вообще говоря, предполагает дальнейшее деление. Так, например, вариант «пригласить гостей» приводит к дальнейшему обсуждению их списка. При этом могут сопоставляться различные варианты. Например, что предпочесть — гастрономические утехи за телевизором в хорошо знакомой компании либо бурное обсуждение злободневных проблем или нравов далеких стран с интересными людьми, с которыми давно не встречались?

Вариант «остаться дома и не звать гостей» также имеет свои варианты. Можно проводить новогоднюю ночь в семейном кругу, и одна из решаемых при этом задач — какую программу телевидения смотреть. А можно лечь спать вскоре после полуночи, например в случае болезни или после долгой тяжелой работы.

Вариант «уехать к родственникам или знакомым» также требует дальнейших решений. Поездка связана прежде всего с поддержанием родственных отношений или с желанием получить удовольствие? Какую пищу вы предпочитаете — физическую или духовную (гастрономические утехи или интересную беседу)?

Оставшийся четвертый вариант «уехать в общедоступные места» предполагает еще больше возможностей выбора. Можно остаться в своем городе, отправиться в другой город (например, из Москвы в Смоленск), выехать на природу (на горнолыжную базу, на курорт), пересечь границу. А тут возможностей масса — все страны, все континенты; можно покататься на слоне в Таиланде, искупаться в Атлантическом океане или побродить по Парижу.

Итак, рядовая задача принятия решения «Как встречать Новый год?» при проработке превращается в выбор из невообразимого количества вариантов. При этом нет необходимости доходить до перечня конкретных вариантов (выехать 28 декабря таким-то поездом туда-то), поскольку решения, очевидно, принимаются последовательно, и решение «остаться дома» делает ненужным рассмотрение всех туристических маршрутов.

Что дает нам декомпозиция решений? Пример 2.3 демонстрирует, как несколько принятых друг за другом решений позволяют справиться с многообразием вариантов. При принятии решений может использоваться весь арсенал теории принятия решений — такие понятия, как цели, критерии, ресурсы, риски и др. (см. главу 1), однако довольно часто решения принимаются на интуитивном уровне, без введения в обсуждение перечисленных понятий.

Дерево решений. Довольно часто удобно представить варианты графически. Обычно возможные решения представляют в виде одного из видов графов — дерева (рис. 2.1). Строго говоря, это перевернутое дерево. Корнем является исходная задача — «Как встречать Новый год?». От него идут две ветви к вариантам «остаться дома» и «уехать». От этих вариантов, в свою очередь являющихся задачами принятия решений («Что делать, оставшись дома?» и «Куда уехать?»), ветви ведут к вариантам задач принятия решений следующего порядка.

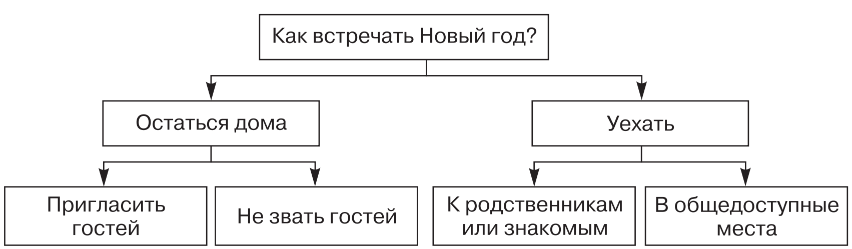


Рис. 2.1. Дерево решений — дихотомическая схема для наглядного представления возможных решений

Пример 2.4. Приведем начало (корень) «Дерева решений проекта», использованного в практической работе.

Задача предприятия — производить качественные изделия из стекловолокна, так как растет потребность в утеплителях и расширяется рынок. Необходимо сделать выбор из двух вариантов:

1) работать на существующем оборудовании;
2) провести реконструкцию цеха.

При выборе первого варианта следует иметь в виду, что мощности оборудования не столь большие, чтобы обеспечить возросшую потребность (из-за физического износа линии), а качество производимой продукции не соответствует международным требованиям (т.е. необходимо учитывать моральный износ линии). Поэтому следует ожидать, что даже в условиях предполагаемого повышенного спроса выпущенные на существующем оборудовании материалы не будут востребованы (реализация будет падать), соответственно мощность производства не будет расти.

При выборе второго варианта решения после реконструкции производительность увеличивается в 2 раза по сравнению с существующей технологической линией, качество выпускаемой предприятием продукции будет соответствовать международным требованиям, она сможет конкурировать с главными производителями стекловолокна. Повысятся основные технико-экономические показатели. Однако существует определенный риск проекта, поскольку необходимы большие капитальные вложения (бóльшая часть которых — из заемных источников).

Дальнейшее построение дерева решений здесь достаточно очевидно. От варианта «работать на существующем оборудовании» пойдут линии к решениям, связанным с упрощением ассортимента выпускаемой продукции, поиском ниши рынка, готовой принимать продукцию более низкого качества, и т.д. Это линия на выживание в условиях отставания от научно-технического прогресса вплоть до ликвидации предприятия. В некоторых условиях ликвидация предприятия — это оптимальный выход.

От варианта «провести реконструкцию цеха» пойдут линии двух типов — сначала «технологические», а затем «финансовые». Сначала надо выбрать конкретный вариант реконструкции и подготовить бизнес-план соответствующего инвестиционного проекта. Затем необходимо обеспечить финансовые поступления для выполнения этого инвестиционного проекта, обеспечив минимальный риск для предприятия. Здесь проблема — выбор кредиторов и заемщиков, заключение с ними договоров на приемлемых условиях.

Кроме последовательного принятия решений, декомпозиция задач принятия решений используется для «разделения проблем на части». При этом результатом декомпозиции является не выбор одного из большого числа вариантов, как при последовательном принятии решений, а представление решаемой задачи в виде совокупности более мелких задач, в пределе — таких задач, методы решения которых известны.

Пример 2.5. Рассмотрим проблему борьбы с транспортным шумом [59]. Целесообразно выделить следующие типы мероприятий:

1) связанные с источником шума;
2) проводимые на месте проявления шума;
3) проводимые на пути распространения шума;
4) относящиеся ко всей системе транспортных средств;
5) связанные с реконструкцией транспортной системы и разработкой способов ее технико-экономической оценки.

В отличие от примера 2.3, здесь не идет речь о том, чтобы выбрать один из вариантов решения. Наоборот, для эффективной борьбы с транспортным шумом необходимо использовать все ветви, все пять типов мероприятий.

Источник шума — это автомашина. Поэтому сразу выделяются три направления воздействия на ситуацию:

1.1) конструкция автомашины (включая регулировку ее узлов);
1.2) топливо;
1.3) дорога.

Непосредственная защита от шума может быть индивидуальная — шлемы, наушники, вставки в уши — «беруши» (сокращение от «берегите уши»), а может быть и коллективная (звуконепроницаемые оконные рамы, стены со звукоизоляцией). Поэтому мероприятия на месте проявления шума естественным образом делятся на два класса:

2.1) индивидуальная защита от шума;
2.2) подавление шума в зданиях.

Можно ослабить шум «по дороге». Хорошо известны различные способы, используемые для этого:

3.1) сооружение звукозащитных стен и экранов, отражающих звуковые волны в безопасных направлениях;
3.2) создание звукозащитных полос из деревьев и кустарников;
3.3) противошумное расположение зданий на местности (как по расстоянию от источника шума, так и по ориентации зданий относительно него и друг друга).

Снижение шума возможно также с помощью различных мероприятий, относящихся ко всей системе транспортных средств. Речь идет о рациональной организации движения в рамках действующей транспортной системы. Эта рациональная организация осуществляется региональными властями административными и частично организационно-экономическими методами. Примерами подобных мероприятий являются:

4.1) направление транзитного транспорта в объезд крупных городов;
4.2) ограничение движения транспорта в определенные часы или по определенным улицам;
4.3) планирование движения транспорта — по времени, по скорости, по маршрутам.

Наконец, необходимо обсудить мероприятия, нацеленные на будущее. Они связаны с реконструкцией транспортных систем и разработкой способов ее технико-экономической оценки. Каков должен быть транспорт будущего? Ясно, что в нем должны быть предусмотрены меры, направленные на снижение шумовой нагрузки. Технико-экономическая оценка транспортных систем будущего должна определяться с учетом шумовой нагрузки. Выразим это как

5.1) шумоподавление в проектируемых и реконструируемых транспортных системах.

Таким образом, одна исходная задача породила 12 новых. Надо не выбирать одну из них, а решать все 12. Однако каждая из 12 является более конкретной, чем исходная. Ее легче решить (после дальнейшей декомпозиции), чем исходную.

Декомпозиция задач принятия решений «от ветвей к корню». До сих пор мы разбирали ситуации, когда задача принятия решения разбивалась на составляющие (с целью уточнения постановки и выбора одной из конкретных формулировок либо с целью разделить одну большую задачу на ряд более мелких). Рассмотрим теперь противоположный процесс, когда конкретные потребности бизнес-процессов организации порождают единый комплекс задач принятия решений.

Рассмотрим процесс декомпозиции задач принятия решений «от ветвей к корню» на примере формирования задач службы контроллинга организации. Для многих организаций актуальны следующие проблемы:

1) отсутствие оперативной информации о производственных процессах, требующее внедрения на предприятии системы производственного учета;
2) высокий уровень накладных расходов в общей сумме затрат, заставляющий заниматься выявлением мест возникновения «ненужных» затрат;
3) излишне большая величина незавершенного производства, влекущая необходимость разработки системы управления заказами;
4) отсутствие эффективного механизма контроля над деятельностью службы закупок. Имеется лишь эпизодический контроль со стороны руководства организации. Это обусловливает необходимость разработки организационно-экономического механизма, позволяющего контролировать уровень цен на закупаемые материалы;
5) планирование накладных расходов на предприятии по факту предыдущего периода. Это требует внедрения процесса бюджетирования;
6) недостаточность используемой системы показателей для управления предприятием. Следовательно, необходима разработка системы показателей финансово-хозяйственной, производственной и социальной деятельности предприятия;
7) отсутствие у руководства предприятия системного представления о деятельности предприятия. Для принятия обоснованных решений по управлению предприятием необходимо создание аналитической службы поддержки принятия таких решений.

Для решения перечисленных актуальных проблем принятия решений при управлении предприятием вытекает необходимость специальной интегрирующей — службы контроллинга. Вполне очевидно, что все «ветви» в рассматриваемой задаче декомпозиции направлены к одному «корню», и этот «корень» описывает задачи принятия решений, поддерживаемые службой контроллинга [32; 130].

До сих пор в процессе декомпозиции все задачи одного уровня считались равнозначными, весовые коэффициенты не вводились. Однако иногда оказывается полезным различные варианты рассматривать с теми или иными коэффициентами.

Пример 2.6. Необходимо разработать процедуру принятия решений, связанных с оценкой эффективности разрабатываемого

медицинского прибора (магнитного сепаратора). Для вычисления обобщенного показателя качества и технического уровня подобных приборов естественно провести декомпозицию на три задачи принятия решений соответственно трем группам показателей:

1) основные показатели назначения;
2) экономические условия потребления;
3) условия обслуживания.

Пусть X — оценка по первой группе показателей, Y — по второй, Z — по третьей. Первая оценка учитывается с весовым коэффициентом 0,6, вторая — 0,2, третья — также 0,2 (сумма трех весовых коэффициентов равна 1). Таким образом, обобщенный показатель качества и технического уровня медицинского прибора оценивается как

$$W = 0{,}6X + 0{,}2Y + 0{,}2Z.$$

На следующем шаге декомпозиции в каждой из трех групп выделяются единичные показатели качества и технического уровня. Так, для группы «основные показатели назначения» выделяют:

1.1) степень очистки $X(1)$;
1.2) время очистки $X(2)$;
1.3) масса субстрата $X(3)$;
1.4) вероятность повреждения здоровых клеток $X(4)$.

Данным показателям также приписывают весовые коэффициенты 0,44, 0,09, 0,18, 0,29 соответственно (сумма весовых коэффициентов равна 1). Поэтому оценка по основным показателям назначения вычисляется как

$$X = 0{,}44X(1) + 0{,}09X(2) + 0{,}18X(3) + 0{,}29X(4).$$

Для группы «экономические условия потребления» выделяют два единичных показателя:

2.1) методы сепарации $Y(1)$;
2.2) патентную чистоту $Y(2)$

и приписывают весовые коэффициенты 0,74 и 0,26 соответственно (сумма весовых коэффициентов равна 1). Поэтому оценка по экономическим условиям потребления вычисляется как

$$Y = 0{,}74Y(1) + 0{,}26Y(2).$$

Для группы «условия обслуживания» выделяют три единичных показателя:

3.1) режим работы $Z(1)$;
3.2.) эргономику $Z(2)$;
3.3) надежность $Z(3)$

и приписывают весовые коэффициенты 0,55, 0,14 и 0,31 соответственно (сумма весовых коэффициентов равна 1). Поэтому оценка по блоку «условия обслуживания» вычисляется как

$$Z = 0{,}55Z(1) + 0{,}14Z(2) + 0{,}31Z(3).$$

Таким образом описан алгоритм декомпозиции в задаче принятия решения относительно оценки эффективности медицинского прибора. Для вычисления обобщенного показателя качества и технического уровня необходимо получить оценки девяти единичных показателей. Обычно это делают с привлечением экспертов, сопоставляющих разрабатываемый прибор с отечественными и зарубежными аналогами. Применение подобных показателей продемонстрировано в подразделе 2.2 на примерах сумм баллов и взвешенных сумм баллов. Однако только в этой главе показано, как могут обоснованно строиться системы факторов на основе идеи декомпозиции.

Для нахождения весовых коэффициентов обычно используют оценки экспертов (см. далее часть III). При этом для каждой группы показателей, а также при присвоении весов группам на верхнем уровне декомпозиции могут применяться свои экспертные процедуры и опрашиваться свои эксперты. Это важное преимущество рассматриваемой процедуры обеспечивается тем, что сумма весовых коэффициентов каждый раз равняется 1.

Дело в том, что из приведенных соотношений следует, что для вычисления обобщенного показателя качества и технического уровня можно использовать непосредственно оценки единичных показателей:

$$W = 0{,}6X + 0{,}2Y + 0{,}2Z = 0{,}6\,(0{,}44\,X(1) + 0{,}09\,X(2) +$$
$$+\,0{,}18\,X(3) + 0{,}29\,X(4)) + 0{,}2\,(0{,}74Y(1) + 0{,}26Y(2)) +$$
$$+\,0{,}2\,(0{,}55Z(1) + 0{,}14Z(2) + 0{,}31Z(3)) = 0{,}264\,X(1) + 0{,}054X(2) +$$
$$+\,0{,}108\,X(3) + 0{,}174X(4) + 0{,}148Y(1) + 0{,}052Y(2) + 0{,}11Z(1) +$$
$$+\,0{,}028Z(2) + 0{,}062Z(3).$$

Сумма итоговых девяти весовых коэффициентов, естественно, равна 1, поскольку так построена схема декомпозиции.

С первого взгляда может показаться рациональной оценка этих девяти коэффициентов непосредственно (с помощью экспертов). В главе 9 критикуется такое предложение. Из сказанного также ясно, что пошаговый метод декомпозиции дает возможность более точно сопоставить весовые коэффициенты (отдельно внутри групп, отдельно группы между собой), чем это можно сделать при объединении всех единичных показателей вместе.

Рассмотренные способы усреднения значений единичных показателей — это фактически применение средних взвешенных арифметических для значений единичных показателей. Целесообразно обратить внимание на возможность применения иных видов средних величин, а также на подходы и результаты теории измерений, позволяющие выбирать наиболее адекватные виды средних величин в соответствии с используемыми шкалами измерения (см. далее главу 9).

В теории и практике принятия решений накоплено большое число различных методов подготовки и принятия решений, как относительно простых, так и основанных на изощренной математической технике. В дальнейших главах мы подробно рассмотрим методы принятия решений, основанных на оптимизационных, вероятностно-статистических и экспертных методах, а в следующей части познакомимся с методом моделирования и различными видами моделей, используемых в теории и практике принятия решений.

Контрольные вопросы

1. Что содержит и как используется матрица портфеля Бостонской консалтинговой группы?
2. Чем отличаются методы проверочного списка и суммарной оценки?
3. Как проводят первичную формализацию описания ситуации при гипотетическом переходе на новую работу?
4. Как бы вы расставили баллы на месте Пети Иванова при принятии решения о выборе места работы?
5. Как проводят декомпозицию задачи принятия решения при гипотетическом переходе на новую работу?
6. Почему метод декомпозиции является весьма полезным при решении многих задач принятия решений?

Темы докладов и рефератов

1. Роль матрицы портфеля Бостонской консалтинговой группы при разработке и принятии управленческих решений.
2. Проблема устойчивости выводов (по отношению к малым отклонениям исходных данных и субъективным «оцифровкам» качественных оценок) при решении проблем стратегического менеджмента.
3. Методы построения суммарной оценки проекта по оценкам отдельных факторов.
4. Проблема агрегирования значений единичных показателей при принятии решений.
5. Использование весовых коэффициентов в задачах принятия решений.
6. Способы выбора весовых коэффициентов в задачах стратегического менеджмента.
7. Классификация постановок задач декомпозиции в теории и практике принятия решений.

ГЛАВА 3
ОСНОВЫ ТЕОРИИ УПРАВЛЕНИЯ

3.1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ УПРАВЛЕНИЯ

Что делает управляющий, т.е. менеджер? Управляет организацией или подразделением, т.е. людьми. Значит, управление людьми — это деятельность менеджера. Познакомимся с рядом понятий, используемых при обсуждении проблем управления.

Субъекты и объекты управления. В каждой стране как хозяйственная, так и нехозяйственная (военная, религиозная, спортивная и др.) деятельность общества сознательно управляется. Что имеется в виду под термином «управление»? *Управление — процесс воздействия субъекта на объект в целях перевода его в новое качественное состояние или поддержания в установленном режиме*.

Субъект управления — это тот, кто управляет. *Объект управления* — это тот, кем управляют.

Обсудим понятия субъектов и объектов управления применительно к экологической безопасности и природоохранной деятельности [136]. Субъектами управления природопользованием, в том числе и природоохранной деятельностью, выступают государственные органы общей компетенции, кроме того, специально уполномоченные органы по охране окружающей природной среды, а также органы местного самоуправления. На уровне предприятий субъектами управления являются подразделения и службы природопользования (цехи, отделы) или отдельные работники.

К государственным органам *общей компетенции* относятся Президент, Федеральное Собрание, Правительство, представительные и исполнительные органы власти субъектов Российской Федерации. Государственные и муниципальные органы общей компетенции ведают вопросами охраны природной среды наряду с множеством других направлений работы.

К государственным органам *специальной компетенции* относятся те, которые соответствующими правительственными актами уполномочены выполнять природоохранные функции. Органы специальной компетенции подразделяются на три вида: комплексные, отраслевые и функциональные. Комплексные природоохранные органы выполняют все экологические задачи или отдельные блоки задач, отраслевые занимаются своей отраслью (например, лесным хозяйством), функциональные отвечают за отдельные функции (например мониторинг состояния окружающей природной среды).

К охране природной среды имеют отношение многие министерства и ведомства. Например, Министерство экономического развития Российской Федерации, формируя стратегию социально-экономического развития, обязано учитывать и экологические интересы общественного развития. Федеральное агентство по техническому регулированию и метрологии осуществляет надзор за соблюдением государственных стандартов, и в том числе за выполнением зафиксированных в них экологических требований. Государственный комитет Российской Федерации по строительной, архитектурной и жилищной политике разрабатывает экологические требования при капитальном строительстве и т.п.

Компетенция органов местного самоуправления по охране окружающей среды отражена в уставах. Их роль напоминает предназначение государственных органов общей компетенции, только на гораздо более узком пространстве, относящемся к ведению соответствующего органа местного самоуправления, — на уровне района или города.

Объектами управления являются все природопользователи, как юридические, так и физические лица независимо от характера и направлений их деятельности. Поскольку все организации и предприятия, все жители городов и деревень находятся и действуют в природной среде, то объектами управления являются все юридические и физические лица на территории нашей страны. Связи и отношения между субъектами и объектами управления в процессе природопользования и охраны природной среды строятся двумя способами:

- на основе правил и процедур, зафиксированных в действующих законах и других нормативно-правовых актах;
- на основе договоров между конкретными субъектами и объектами управления.

Методы и механизмы управления. *Метод управления* — это набор способов, приемов, средств воздействия на управляемый объект. По содержанию воздействия на объект управления методы обычно делятся на организационно-административные, экономические, социально-психологические и др.

Так, *организационно-административные методы* основаны на приказах, распоряжениях, законах и других нормативно-правовых документах и опираются на возможность применения силы государственными органами, в том числе непосредственно на силовые структуры.

Внутри организации взаимоотношения менеджеров их подчиненных регулируются Трудовым кодексом Российской Федерации.

Экономические методы воздействия основаны на использовании материальных (экономических, денежных) интересов. Конкретный экономический метод включает как отдельные приемы воздействия, так и их совокупности. Комплекс взаимосвязанных экономических мер, направленных на достижение конкретного результата, образует экономический механизм управления.

Социально-психологические методы управления опираются на убеждение, моральное стимулирование, сознательность; держатся на обычаях и традиционных ценностях общества.

Под ***механизмом управления*** понимают совокупность тех или иных методов управления. Организационно-административные, экономические и социально-психологические методы управления применяются совместно. Ясно, что сама возможность использования экономических и социально-психологических методов опирается на существующую административную структуру предприятия. С другой стороны, чисто административными (командными) методами, без материального и морального стимулирования нельзя добиться существенного повышения эффективности работы предприятия.

Организационно-административный, экономический, социально-психологический механизмы являются частями системы управления в целом. На различных уровнях управления эта система имеет свои особенности. Можно выделить *макроуровень*, т.е. управление в рамках всей страны, и *мезоуровень*, касающийся отдельных секторов и отраслей народного хозяйства, например управления добычей нефти и газа. На уровне конкретных предприятий системы управления носят более специальный характер, приспособленный к особенностям этих предприятий и их подразделений. Большое практическое значение имеет и самый нижний уровень управления — управление собой. Можно сказать, что каждый является менеджером, поскольку он управляет, по крайней мере, одним человеком — самим собой.

Цели управления. Определение целей, к которым следует стремиться, или, как говорят, *целеполагание*, — самая трудная и ответственная часть работы менеджера. Выбор целей зависит от конкретной ситуации. Рассмотрим это на примере агропромышленного комплекса (АПК) [136].

Пусть установлено, что производство в АПК, несмотря на резкое снижение его объема за последние 15 лет, излишне велико для данной области, а дефицит продовольствия объясняется не недостаточным объемом производства, а отсталостью в сфере хранения и переработки. Тогда целью управления природопользованием в данной области должно стать сокращение природного базиса сельского хозяйства, т.е. сокращение объема используемых в сельском хозяйстве природных ресурсов.

Меры экономического воздействия будут включать, например, установление высокой арендной платы за земли сельскохозяйственного назначения. Это позволит затормозить вовлечение новых земель в хозяйственный оборот. Следует повысить налоги на дополнительное освоение земель, увеличить штрафы за нерациональное использование земель, стимулировать различными способами консервацию деградированных участков и т.д. Все эти меры направлены на сокращение сельскохозяйственного производства и снятие сельскохозяйственной нагрузки с окружающей природной среды. Одновременно следует бороться с отсталостью в сфере хранения произведенного продовольствия и в сфере перерабатывающей промышленности. Необходимо создание благоприятных экономических условий для совершенствования технологий хранения и переработки сельскохозяйственной продукции, развития соответствующей отрасли народного хозяйства.

Если же целью развития АПК на определенный период считать всемерное увеличение производства сельскохозяйственной продукции, то меры экономического воздействия, наоборот, не только не должны препятствовать вовлечению новых земельных и водных ресурсов, химических средств защиты растений, минеральных удобрений, а всемерно стимулировать их. Подавляющее число специалистов считает, что проводимые в 1990-х гг. в России аграрная и земельная реформы были направлены на природоемкий вариант функционирования АПК. За эти годы производство сельскохозяйственной продукции упало в среднем на 30—40%.

Управляющие параметры. В математических моделях, применяемых при управлении, используются различные виды переменных. Одни из них описывают состояние системы, другие — выход системы, т.е. результаты ее работы, третьи — управляющие воздействия. Выделяют экзогенные переменные, значения которых определяются извне, и эндогенные переменные, используемые только для описания процессов внутри системы.

Управляющие параметры — часть экзогенных. Задавая их значения (или изменения этих переменных во времени), менеджер меняет выход системы в нужную для себя сторону.

Поскольку невозможно абсолютно точно предсказать поведение системы под влиянием тех или иных воздействий, приходится изучать устойчивость социально-экономических моделей, используемых при управлении [88]. Чаще всего вводят случайные воздействия (возмущения), что приводит к замене единственной траектории движения на пучок (трубку) и снижает эффективность управления. Соответствующая математическая теория хорошо разработана, но достаточно трудна для понимания и применения.

Различные варианты формулировок целей управления. В простейшем случае цель полностью описана. Например, необходимо попасть в определенное место за минимальное время, или построить дом по заранее выбранному проекту, руководствуясь утвержденной сметой. Этот вариант целеполагания назовем «попасть в точку».

Двойственным к нему является вариант «продвинуться дальше». Например, в течение заданного времени изготовить как можно больше деталей или при заданном рекламном бюджете организовать наиболее эффективную рекламную кампанию.

В этом варианте есть критерии, по которым надо оптимизировать системы. В первом случае — число деталей. Во втором — эффективность рекламной кампании, которую можно измерить по увеличению числа покупателей и объема продаж. Критерии — часть переменных, описывающих выход системы, т.е. те результаты ее работы, которые представляют интерес для менеджера.

Частный случай варианта «продвинуться дальше» — как можно ближе подойти к заранее определенному идеальному состоянию. Например, сконструировать двигатель, коэффициент полезного действия которого возможно ближе к идеалу — к 100%. Назовем этот вариант целеполагания «приблизиться к идеалу». Для его реального использования необходимо иметь измерять степень близости к идеалу.

От варианта «попасть в точку» естественно перейти к варианту «попасть в область». Например, человек может определить для себя желательный уровень заработной платы и считать цель достигнутой, как только его заработная плата превысит заданный порог. Менеджеру, отвечающему за отопительную систему, необходимо обеспечить температуру в помещениях в заданных пределах — от и до. Руководителям предприятия необходимо обеспечить попадание показателей финансово-хозяйственной деятельности предприятия в заданные интервалы.

Одна из наиболее важных целей как для отдельного человека, так и для организации — самосохранение. Стремление сохраниться как самостоятельное целое, обеспечить равновесие с окружающей средой, стабильность и целостность — этот вариант целеполагания назовем «Приказано выжить». Именно самосохранение, а не максимизация

прибыли — основная цель предприятия (подробнее эта проблема обсуждается в следующей главе).

Другой подход к проблеме управления связан с идеей обратной связи. Управление не выбирается заранее, а корректируется в каждый текущий момент времени на основании информации о состоянии или о выходе системы.

Для подразделения внутри организации обычно имеется конфликт между внешними и внутренними целями. Например, центральное руководство имеет целью получить от цеха возможно больше продукции при возможно меньшей оплате труда, в то время как руководство цеха желает прямо противоположного — несколько сократить выпуск продукции, но увеличить фонд оплаты труда. Аналогична ситуация и для организации во взаимоотношениях с внешним миром. Потребители хотят получить товары возможно более высокого качества при возможно более низкой оплате труда, а поставщики, наоборот, предпочли бы не заботиться о качестве, но зато увеличить цены.

Неизбежен конфликт между внутренними целями субъекта экономической жизни (от отдельного человека до групп стран) и внешними целями его окружения. Разрешение подобных конфликтов — одна из основных задач менеджера.

3.2. МНОГОКРИТЕРИАЛЬНОСТЬ РЕАЛЬНЫХ ЗАДАЧ УПРАВЛЕНИЯ

Обычно субъект экономической жизни стремится достичь сразу многих целей, например, одновременно следовать и внутренним и внешним целям, тем самым снимая конфликт между ними. Но этих целей у него может быть много.

Например, простейший набор целей на день для студента может выглядеть так:

- воспринять очередную порцию знаний, посетив занятия;
- продвинуться в научно-исследовательской работе;
- пообщаться с товарищами по учебе;
- провести спортивную тренировку;
- несколько часов провести на работе, выполняя производственные задания;
- повстречаться с друзьями вне вуза;
- помочь родителям по дому;
- выполнить домашние задания, подготовиться к следующему учебному дню;
- отдохнуть вечером.

Совершенно ясно, что эти цели конкурируют между собой, борясь за время и силы студента. Трудно совместить даже две из этих целей, связанные с учебой и работой.

Как же быть? Многокритериальность реальных задач управления состоит в том, что менеджеру необходимо оптимизировать управляемую им систему сразу по нескольким критериям, например, добиться максимизации прибыли при минимуме затрат. Ясно, что этого невозможно достичь. Минимум затрат равен 0, он достигается при прекращении выпуска продукции (оказания услуг) и ликвидации предприятия. Но при этом прибыль тоже равна 0. Если же добиться максимально возможной прибыли, то затраты при этом также будут достаточно большими, отнюдь не минимальными.

Теория управления предлагает два основных способа борьбы с многокритериальностью. Один из них — превратить все критерии, кроме одного, в ограничения и решать задачу оптимизации по оставшемуся критерию (о задачах оптимизации рассказывается в главе 4). Например, можно потребовать, чтобы затраты не превосходили заданной величины, и при этом условии максимизировать прибыль. Второй вариант — принять, что прибыль должна быть не меньше заданной величины (например, если выполняется определенный заказ), а затраты при этом условии минимизировать.

Другой подход в борьбе с многокритериальностью — на основе исходных критериев сконструировать один новый и его оптимизировать. В рассматриваемом случае можно использовать рентабельность (по затратам), т.е. частное от деления прибыли на затраты. При максимизации рентабельности находится наилучшее (в определенном смысле) соотношение между затратами и прибылью.

Есть и другие методы борьбы с многокритериальностью. Например, можно выделить все варианты решений менеджера, при которых прибыль мало отличается от максимально возможной, а затем в этой области минимизировать затраты [88]. Или же сначала выделить все Парето-оптимальные варианты решений менеджера, т.е. все те решения, которые не хуже любого возможного решения хотя бы по одному критерию, а затем анализировать множество Парето-оптимальных решений [100].

Аналогична ситуация и с лозунгом «Максимум прибыли при минимуме риска». Здесь, как и в ранее разобранном случае, надо либо максимизировать прибыль при задании верхней границы для риска, либо минимизировать риск при заданной прибыли, либо конструировать из двух критериев один. Дополнительная сложность состоит в необходимости численно оценивать риск.

3.3. ОБ ОПТИМАЛЬНОМ УПРАВЛЕНИИ ЭКОНОМИЧЕСКИМИ СИСТЕМАМИ

Рассмотрим две проблемы сравнительной оценки эффективности различных подходов к оптимизации управления экономическими системами.

Сравнение по эффективности государственных и частных предприятий. К настоящему времени имеется огромное количество книг, статей, ресурсов Интернета, связанных с поисками оптимального управления теми или иными экономическими системами. Среди обсуждаемых вопросов, например, такой: «Какие предприятия работают более эффективно — государственные или частные?»

Частные предприятия хвалят за возможность развертывания личной инициативы, быстроту реакции на изменение рыночной среды. Критикуют их за высокую вероятность различных противоправных и противообщественных действий. Государственные предприятия критикуют за неповоротливость, вызванную бюрократизмом управления, а хвалят за законопослушность. На основе словесного обсуждения нельзя сделать однозначного вывода. Обратимся к статистике.

В таблице 3.1 приведены статистические данные Европейского союза (ЕС) на конец 1995 г. [1; 107]. Для каждой страны и для ЕС в целом указаны доля занятых в государственном секторе (в процентах от всех занятых в стране) и доля государственного сектора в создании добавленной стоимости (в процентах от валового внутреннего продукта — ВВП). В последнем столбце указана эффективность государственного сектора в странах ЕС относительно предприятий и организаций других форм собственности. Она получена делением второй из этих долей на первую. Страны упорядочены в порядке убывания ВВП.

Таблица 3.1.

Эффективность государственного сектора в странах Европейского союза

Страна	Доля занятых в государственном секторе, %	Доля государственного сектора в добавленной стоимости, %	Относительная эффективность государственного сектора
Германия	8,0	10,0	1,25
Франция	11,9	14,2	1,19
Италия	11,5	13,0	1,13
Великобритания	2,3	2,6	1,13
Испания	6,0	7,2	1,20
Швеция	11,4	13,3	1,17
Австрия	10,0	14,0	1,40

Окончание

Страна	Доля занятых в государственном секторе, %	Доля государственного сектора в добавленной стоимости, %	Относительная эффективность государственного сектора
Бельгия	9,7	8,6	0,89
Греция	12,0	14,0	1,17
Финляндия	14,7	19,0	1,29
Португалия	6,0	13,1	2,18
Нидерланды	3,4	6,0	1,76
Дания	7,8	8,0	1,03
Ирландия	9,3	11,0	1,18
Люксембург	5,6	6,3	1,13
Европейский союз в целом	8,0	9,7	1,21

Таким образом, производительность труда (объем созданной добавленной стоимости на одного работающего) в государственном секторе выше, чем на частных предприятиях. Это утверждение справедливо для ЕС в целом и для всех его стран по отдельности, кроме одной — Бельгии.

Итак, согласно мировому опыту экономическая эффективность государственных предприятий выше, чем у частных. Одна из причин состоит в том, что зарубежные государства предпочитают оставлять у себя (или национализировать) высокоэффективные предприятия и избавляться от малоэффективных и убыточных, приватизируя их.

Догма эффективности конкуренции. Она состоит в том, что экономические структуры, основанные на конкуренции, более эффективны (производят больше при тех же затратах), чем плановые. Западная экономическая теория пришла к противоположному выводу: эффективность конкурентной системы может в отдельных случаях достигать эффективности плановой системы, но никогда не в состоянии превзойти ее. Это и понятно — оптимальная программа действий, учитывающая всю информацию, всегда эффективнее, чем стихийно складывающийся поток несогласованных действий отдельных субъектов рыночной конкуренции.

Однако у каждого типа системы есть свои причины недостижения максимума. В конкурентной системе — стихийные отклонения от оптимальной траектории, вызванные несогласованными решениями экономических субъектов. Для рыночной экономики характерна нерациональная трата материальных, финансовых, кадровых ресурсов.

Недаром в период войны, когда необходимо сосредоточить ресурсы для обеспечения вооруженных сил, все государства переходят к плановой системе управления экономикой.

В плановой экономике основная причина недостижения максимума — нерациональность процедур принятия решений, принципиальная невозможность точного решения оптимизационных задач, а главное — неустранимая неопределенность в их постановке. Существуют подходы, позволяющие модернизировать плановую экономику с целью повышения эффективности, вводя в нее элементы рыночных отношений. Утверждать, что «рынок» эффективнее «плана», нельзя.

Термин «конкуренция» означает «соревнование». В СССР конкурировали между собой разработчики военной техники — из нескольких образцов, разработанных различными организациями по одним и тем же тактико-техническим требованиям, комиссия экспертов выбирала один для запуска в серию. В настоящее время подобная процедура именуется тендером (конкурсом). Конкуренция процветала и на рынке потребительских товаров — население «голосовало рублем» за те товары, которые предпочитало, остальные оставались на полках магазинов и со временем списывались.

Должен ли менеджер специально создавать конкуренцию между работниками? Практический опыт показывает, что вводить элементы соревнования следует весьма осторожно, упирая на моральное стимулирование (как в Японии). Работники должны образовывать сплоченный коллектив (команду), а не стаю особей, готовых перегрызть глотку друг другу.

Контрольные вопросы

1. Что называют субъектом и объектом управления?
2. Какие вы знаете виды методов управления?
3. Как можно бороться с многокритериальностью задач управления?
4. Сравните по эффективности государственные и частные предприятия в странах Европейского союза.
5. Всегда ли полезна конкуренция?

Темы докладов и рефератов

1. Особенности системы управления высшим учебным заведением.
2. Особенности системы управления космической отраслью (на основе публикаций [1; 107]).
3. Многокритериальность в поведении студентов.
4. Примеры Парето-оптимальных решений многокритериальных задач (с использованием [100]).
5. Управление экономикой при рыночном социализме.

Часть II

МАТЕМАТИЧЕСКИЕ МЕТОДЫ РАЗРАБОТКИ И ПРИНЯТИЯ РЕШЕНИЙ

ГЛАВА 4
МЕТОДЫ ОПТИМИЗАЦИИ ПРИ ПРИНЯТИИ РЕШЕНИЙ

4.1. ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

Менеджер может использовать при принятии решения различные компьютерные и математические средства. В памяти компьютеров хранят много информации, организованной с помощью баз данных и других программных продуктов, позволяющих оперативно ею пользоваться. Экономико-математические и эконометрические модели позволяют просчитывать последствия тех или иных решений, прогнозировать развитие событий. Методы экспертных оценок, о которых пойдет речь далее, также весьма математизированы и используют компьютеры.

Наиболее часто используются оптимизационные модели принятия решений. Их общий вид таков:

$$F(X) \to \max;$$

$$X \in A,$$

где X — параметр, который менеджер может выбирать (управляющий параметр). Он может иметь различную природу — число, вектор, множество и т.п.

Цель менеджера — максимизировать целевую функцию $F(X)$, выбрав соответствующий X. При этом он должен учитывать ограничения $X \in A$ на возможные значения управляющего параметра X — он должен лежать в множестве A. Рассмотрим примеры оптимизационных задач менеджмента.

Среди оптимизационных задач менеджмента наиболее известны задачи линейного программирования, в которых максимизируемая функция $F(X)$ линейная, а ограничения A задаются линейными неравенствами.

Производственная задача. Цех может производить стулья и столы. На производство стула идет 5 единиц материала, на производство стола — 20 единиц (футов красного дерева). Стул требует 10 человеко-часов, стол — 15. Имеется 400 единиц материала и 450 человеко-часов. Прибыль при производстве стула — 45 дол. США, при производстве стола — 80 дол. Сколько надо сделать стульев и столов, чтобы получить максимальную прибыль?

Обозначим X_1 число изготовленных стульев, X_2 — число столов. Задача оптимизации имеет вид:

$$45X_1 + 80X_2 \to \max;$$

$$5X_1 + 20X_2 \le 400;$$

$$10X_1 + 15X_2 \le 450;$$

$$X_1 \ge 0;$$

$$X_2 \ge 0.$$

В первой строке выписана целевая функция — прибыль при выпуске X_1 стульев и X_2 столов. Ее требуется максимизировать, выбирая оптимальные значения переменных X_1 и X_2. При этом должны быть выполнены ограничения по материалу (вторая строчка) — истрачено не более 400 футов красного дерева. А также и ограничения по труду (третья строчка) — затрачено не более 450 ч. Кроме того, нельзя забывать, что число столов и число стульев неотрицательны. Если $X_1 = 0$, то это значит, что стулья не выпускаются. Если же хоть один стул сделан, то X_1 положительно. Но невозможно представить себе отрицательный выпуск — X_1 не может быть отрицательным с экономической точки зрения, хотя с математической точки зрения такого ограничения усмотреть нельзя. В четвертой и пятой строчках задачи и констатируется, что переменные неотрицательны.

Условия производственной задачи можно изобразить на координатной плоскости. Будем по оси абсцисс откладывать значения X_1, а по оси ординат — значения X_2. Тогда ограничения по материалу и последние две строчки оптимизационной задачи выделяют возможные значения (X_1, X_2) объемов выпуска в виде треугольника (рис. 4.1).

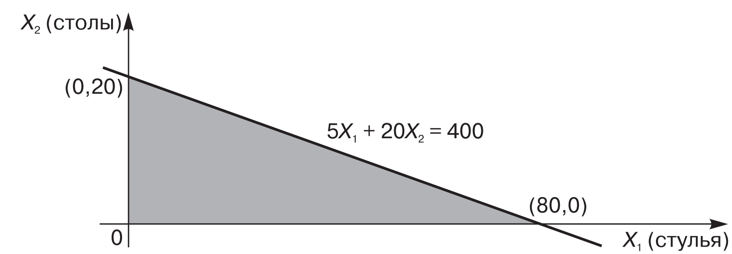


Рис. 4.1. Ограничения по материалу

Таким образом, ограничения по материалу изображаются в виде выпуклого многоугольника, в данном случае — треугольника. Этот треугольник получается путем отсечения от первого квадранта примыкающей к началу координат зоны. Отсечение проводится прямой, соответствующей второй строке исходной задачи, с заменой неравенства на равенство. Прямая пересекает ось X_1, соответствующую стульям, в точке (80,0). Это означает, что если весь материал пустить на изготовление стульев, то будет изготовлено 80 стульев. Та же прямая пересекает ось X_2, соответствующую столам, в точке (0,20). Это означает, что если весь материал пустить на изготовление столов, то будет изготовлено 20 столов. Для всех точек внутри треугольника выполнено неравенство, что означает — материал останется.

Аналогичным образом можно изобразить и ограничения по труду (рис. 4.2).

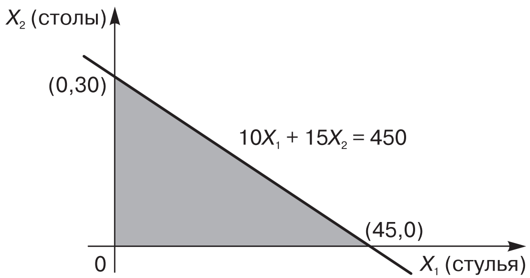


Рис. 4.2. Ограничения по труду

Ограничения по труду, как и ограничения по материалу, изображаются в виде треугольника, который получается аналогично — путем отсечения от первого квадранта примыкающей к началу координат зоны. Отсечение проводится прямой, соответствующей третьей строке исходной задачи, с заменой неравенства на равенство. Прямая пересекает ось X_1, соответствующую стульям, в точке (45,0). Это означает, что если все трудовые ресурсы пустить на изготовление стульев, то будет сделано 45 стульев. Та же прямая пересекает ось X_2, соответствующую столам, в точке (0,30). Это означает, что если всех рабочих поставить на изготовление столов, то будет сделано 30 столов. Для всех точек внутри треугольника выполнено неравенство, что означает — часть рабочих будет простаивать.

Мы видим, что очевидного решения нет — для изготовления 80 стульев есть материал, но не хватает рабочих рук, а для производства 30 столов есть рабочая сила, но нет материала, Значит, надо изготавливать и то и другое. Но в каком соотношении?

Чтобы ответить на этот вопрос, надо «совместить» рис. 4.1 и рис. 4.2, получив область возможных решений, а затем проследить, какие значения принимает целевая функция на этом множестве (рис. 4.3).

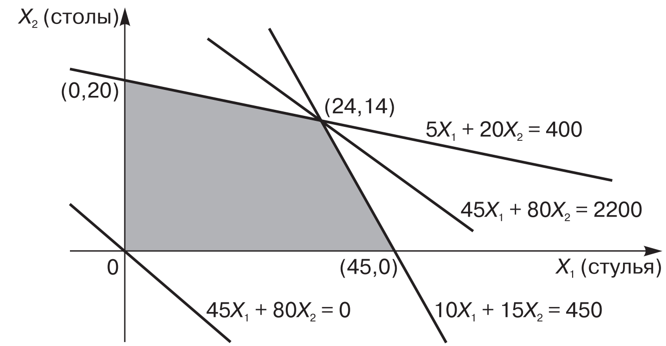


Рис. 4.3. Основная идея линейного программирования

Таким образом, множество возможных значений объемов выпуска стульев и столов (X_1, X_2), или, в других терминах, множество A, задающее ограничения на параметр управления в общей оптимизационной задаче, представляет собой пересечение двух треугольников, т.е. выпуклый четырехугольник, показанный на рис. 4.3. Три его вершины очевидны — это (0,0), (45,0) и (0,20). Четвертая — это пересечение двух прямых — границ треугольников на рис. 4.1 и рис. 4.2, т.е. решение системы уравнений

$$5X_1 + 20X_2 = 400;$$

$$10X_1 + 15X_2 = 450.$$

Из первого уравнения: $5X_1 = 400 - 20\,X_2$, $X_1 = 80 - 4X_2$. Подставляем значение X_1, выраженное через X_2, во второе уравнение:

$$10(80 - 4X_2) + 15X_2 = 800 - 40X_2 + 15X_2 = 800 - 25X_2 = 450,$$

следовательно, $25X_2 = 350$, $X_2 = 14$, откуда $X_1 = 80 - 4 \times 14 = 80 - 56 = 24$. Итак, четвертая вершина четырехугольника — это (24, 14).

Надо найти максимум линейной функции на выпуклом многоугольнике. (В общем случае линейного программирования — максимум линейной функции на выпуклом многограннике, лежащем в конечномерном линейном пространстве.) Основная идея линейного программирования состоит в том, что максимум достигается в вершинах многоугольника. В общем случае — в одной вершине, и это — единственная точка максимума. В частном — в двух, и тогда отрезок, их соединяющий, тоже состоит из точек максимума.

Целевая функция $45X_1 + 80X_2$ принимает минимальное значение, равное 0, в вершине (0,0). При увеличении аргументов эта функция увеличивается. В вершине (24, 14) она принимает значение 2200. При этом прямая $45X_1 + 80X_2 = 2200$ проходит между прямыми ограничений $5X_1 + 20X_2 = 400$ и $10X_1 + 15X_2 = 450$, пересекающимися в той же точке. Отсюда, как и из непосредственной проверки двух оставшихся вершин, вытекает, что максимум целевой функции, равный 2200, достигается в вершине (24, 14).

Таким образом, оптимальный выпуск таков: 24 стула и 14 столов. При этом используется весь материал и все трудовые ресурсы, а прибыль равна 2200 дол.

Двойственная задача. Каждой задаче линейного программирования соответствует так называемая двойственная задача. В ней по сравнению с исходной задачей строки переходят в столбцы, неравенства меняют знак, вместо максимума ищется минимум (или, наоборот, вместо минимума — максимум). Задача, двойственная к двойственной — эта сама исходная задача. Сравним исходную задачу (слева) и двойственную к ней (справа):

$$\begin{array}{cc}
45X_1 + 80X_2 \to \max; & 400W_1 + 450W_2 \to \min; \\
5X_1 + 20X_2 \le 400; & 5W_1 + 10W_2 \ge 45; \\
10X_1 + 15X_2 \le 450; & 20W_1 + 15W_2 \ge 80; \\
X_1 \ge 0; & W_1 \ge 0; \\
X_2 \ge 0. & W_2 \ge 0.
\end{array}$$

Почему двойственная задача столь важна? Можно доказать, что оптимальные значения целевых функций в исходной и двойственной задачах совпадают (т.е. максимум в исходной задаче совпадает с минимумом в двойственной). При этом оптимальные значения W_1 и W_2 показывают стоимость материала и труда соответственно, если их оценивать по вкладу в целевую функцию. Чтобы не путать с рыночными ценами этих факторов производства, W_1 и W_2 называют «объективно обусловленными оценками» сырья и рабочей силы.

Линейное программирование как научно-практическая дисциплина. Из всех задач оптимизации задачи линейного программирования выделяются тем, что в них имеются ограничения — системы линейных неравенств или равенств. Ограничения задают выпуклые линейные многогранники в конечном линейном пространстве. Целевые функции также линейны.

Впервые такие задачи решались советским математиком Л.В. Канторовичем (1912—1986) в 1930-х гг. как задачи производственного менеджмента с целью оптимизации организации производства и производственных процессов, например процессов загрузки станков и раскройки листов материалов. После Второй мировой войны аналогичными задачами занялись в США. В 1975 г. Т. Купманс (1910—1985, родился в Нидерландах, работал в основном в США) и академик АН СССР Л.В. Канторович награждены Нобелевскими премиями по экономике.

Рассмотрим несколько типовых задач линейного программирования.

Задача о диете (упрощенный вариант). Предположим для определенности, что необходимо составить самый дешевый рацион питания цыплят, содержащий необходимое количество определенных питательных веществ (для простоты, тиамина Т и ниацина Н).

Пищевая ценность рациона (в калориях) должна быть не менее заданной. Пусть для простоты смесь для цыплят изготавливается из двух продуктов — *К* и *С*. Известно содержание тиамина и ниацина в этих продуктах, а также их пищевая ценность. Сколько продуктов *К* и *С* надо взять для одной порции куриного корма, чтобы цыплята получили необходимую им дозу веществ Н и Т и калорий (или больше), а стоимость порции была минимальна? Исходные данные для расчетов приведены в табл. 4.1.

Таблица 4.1

Исходные данные к задаче об оптимизации смеси

Вещество, параметр	Содержание в 1 унции продукта *К*	Содержание в 1 унции продукта *С*	Потребность
Вещество Т	0,1 мг	0,25 мг	1 мг
Вещество Н	1,0 мг	0,25 мг	5 мг
Пищевая ценность смеси, кал	110,0	120,00	400
Стоимость 1 унции, центов	3,8	4,2	—

Задача линейного программирования имеет вид:

$$3{,}8K + 4{,}2\,C \to \min;$$

$$0{,}10K + 0{,}25C \ge 1{,}00;$$

$$1{,}00K + 0{,}25\,C \ge 5{,}00;$$

$$110{,}00K + 120{,}00C \ge 400{,}00;$$

$$K \ge 0;$$

$$C \ge 0.$$

Ее графическое решение представлено на рис. 4.4. Ради облегчения восприятия четыре прямые обозначены номерами *1—4*. Прямая *1* — это прямая $1{,}00K + 0{,}25C = 5{,}00$ (ограничение по веществу Н). Она проходит, как и показано на рис. 4.1, через точки (5,00) на оси абсцисс и (0,20) на оси ординат. Обратите внимание, что допустимые значения параметров (*К*, *С*) лежат выше прямой *1* или на ней, в отличие от ранее рассмотренных случаев в предыдущей производственной задаче линейного программирования.

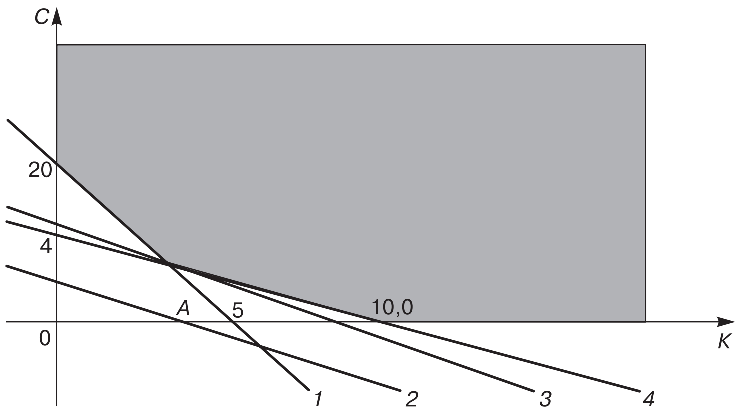


Рис. 4.4. Графическое решение задачи об оптимизации смеси

Прямая *2* — это прямая $110{,}00К + 120{,}00С = 400{,}00$ (ограничение по пищевой ценности). Обратим внимание, что в области неотрицательных *С* она расположена всюду ниже прямой *1*. Действительно, это верно при $К = 0$, прямая *1* проходит через точку (0; 20), а прямая *2* — через расположенную ниже точку $(0; {}^{400}/_{120})$. Точка пересечения двух прямых находится при решении системы уравнений

$$1{,}00К + 0{,}25С = 5{,}00;$$

$$110{,}00К + 120{,}00С = 400{,}00.$$

Из первого уравнения $К = 5 - 0{,}25\ С$ подставим во второе:

$$110(5 - 0{,}25С) + 120С = 400,$$

откуда $550 - 27{,}5С + 120С = 400$.

Следовательно, $150 = -92{,}5С$, т.е. решение достигается при отрицательном *С*. Это и означает, что при всех положительных *С* прямая *2* лежит ниже прямой *1*. Значит, если выполнено ограничение по веществу, то обязательно выполнено и ограничение по пищевой ценности. Мы столкнулись с новым явлением — некоторые ограничения с математической точки зрения могут оказаться лишними. С точки зрения менеджера, они необходимы, отражают существенные черты постановки задачи, но в данном случае внутренняя структура задачи оказалась такова, что ограничение по пищевой ценности не участвует в формировании допустимой области параметров и нахождении решения.

Прямая *4* — это прямая $0{,}1К + 0{,}25С = 1$ (ограничение по веществу Т). Она проходит, как и показано на рис. 4.4, через точки (10; 0) на оси абсцисс и (0; 4) на оси ординат. Обратите внимание, что допустимые значения параметров (*К*, *С*) лежат выше прямой *4* или на ней, как и для прямой *1*.

Следовательно, область допустимых значений параметров (*К*, *С*) является неограниченной сверху. Из всей плоскости она выделяется осями координат (лежит в первом квадранте) и прямыми *1* и *4* (лежит выше этих прямых, а также включает граничные отрезки). Область допустимых значений параметров, т.е. точек (*К*, *С*), можно назвать «неограниченным многоугольником». Минимум целевой функции $3{,}8К + 4{,}2С$ может достигаться только в вершинах этого «многоугольника». Вершин всего три. Это пересечения с осями абсцисс (10; 0) и ординат (0; 20) прямых *1* и *4* (в каждом случае из двух пересечений берется то, которое удовлетворяет обоим ограничениям). Третья вершина — это точка *А* пересечения прямых *1* и *4*, координаты которой находятся при решении системы уравнений

$$0{,}10К + 0{,}25С = 1{,}00;$$

$$1{,}00К + 0{,}25С = 5{,}00.$$

Из второго уравнения $К = 5 - 0{,}25С$, из первого $0{,}10(5 - 0{,}25С) + 0{,}25С = 0{,}5 - 0{,}025С + 0{,}25С = 0{,}5 + 0{,}225С = 1$, откуда $С = {}^{0{,}5}/_{0{,}225} = {}^{20}/_{9}$ и $К = 5 - {}^{5}/_{9} = {}^{40}/_{9}$. Итак, $А = ({}^{40}/_{9}; {}^{20}/_{9})$.

Прямая *3* на рис. 4.4 — это прямая, соответствующая целевой функции $3{,}8К + 4{,}2С$. Она проходит между прямыми *1* и *4*, задающими ограничения, и минимум достигается в точке *А*, через которую и проходит прямая *3*. Следовательно, минимум равен $3{,}8 \times {}^{40}/_{9} + 4{,}2 \times {}^{20}/_{9} = {}^{236}/_{9}$. Задача об оптимизации смеси полностью решена.

Двойственная задача, построенная по описанным правилам, имеет следующий вид (мы повторяем здесь и исходную задачу об оптимизации смеси, чтобы наглядно продемонстрировать технологию построения двойственной задачи):

$3{,}8К + 4{,}2С \to \min;$	$W_1 + 5\ W_2 + 400\ W_3 \to \max;$
$0{,}10К + 0{,}25С \geq 1{,}00;$	$0{,}1W_1 + 1{,}10W_2 + 110W_3 \leq 3{,}8;$
$1{,}00К + 0{,}25С \geq 5{,}00;$	$0{,}25W_1 + 0{,}25W_2 + 120W_3 \leq 4{,}2;$
$110{,}00К + 120{,}00С \geq 400{,}00;$	$W_1 \geq 04$
$К \geq 0;$	$W_2 \geq 0;$
$С \geq 0.$	$W_3 \geq 0.$

Минимальное значение в прямой задаче, как и должно быть, равно максимальному значению в двойственной задаче, т.е. оба числа равны ${}^{236}/_{9}$. Интерпретация двойственных переменных: W_1 — «стоимость» единицы вещества Т, а W_2 — «стоимость» единицы вещества Н, измеренные «по их вкладу» в целевую функцию. При этом $W_3 = 0$, поскольку ограничение на число калорий никак не участвует в формировании оптимального решения. Итак, W_1, W_2, W_3 — это так называемые «объективно обусловленные оценки» (по Л.В. Канторовичу) ресурсов (массы веществ Т и Н, пищевой ценности).

Планирование номенклатуры и объемов выпуска. Вернемся к организации производства. Предприятие может выпускать автоматические кухни (вид кастрюль), кофеварки и самовары. В таблице 4.2 приведены данные о производственных мощностях, имеющихся на предприятии (в штуках изделий).

Таблица 4.2

Производственные мощности, шт.

Операция, показатель	Продукция		
	Кухни	Кофеварки	Самовары
Штамповка	20 000	30 000	12 000
Отделка	30 000	10 000	10 000
Сборка	20 000	12 000	8 000
Объем выпуска	X_1	X_2	X_3
Удельная прибыль (на одно изделие)	15	12	14

При этом штамповка и отделка проводятся на одном и том же оборудовании. Оно позволяет штамповать в течение заданного времени либо 20 000 кухонь, либо 30 000 кофеварок, либо и то и другое, но в меньшем количестве. А вот сборка проводится на отдельных участках.

Задача линейного программирования имеет вид:

$$X_1 \ge 0\ , X_2 \ge 0\ , X_3 \ge 0; \qquad (0)$$

$$X_1 / 200 + X_2 / 300 + X_3 / 120 \le 100; \qquad (1)$$

$$X_1 / 300 + X_2 / 100 + X_3 / 100 \le 100; \qquad (2)$$

$$X_1 / 200 \le 100; \qquad (3)$$

$$X_2 / 120 \le 100; \qquad (4)$$

$$X_3 / 80 \le 100; \qquad (5)$$

$$F = 15\ X_1 + 12\ X_2 + 14\ X_3 \to \max,$$

где (0) — обычное в экономике условие неотрицательности переменных; (1) — ограничение по возможностям штамповки (выраженное для облегчения восприятия в процентах); (2) — ограничение по возможностям отделки; (3) — ограничение по сборке для кухонь; (4) — то же для кофемолок; (5) — то же для самоваров (как уже говорилось, все три вида изделий собираются на отдельных линиях).

Наконец, целевая функция F — общая прибыль предприятия.

Заметим, что неравенство (3) вытекает из неравенства (1), а неравенство (4) — из (2). Поэтому неравенства (3) и (4) можно из формулировки задачи линейного программирования исключить.

Отметим сразу любопытный факт. Как будет установлено, в оптимальном плане $X_3 = 0$, т.е. самовары выпускать невыгодно.

Методы решения задач линейного программирования. Методы решения задач линейного программирования относятся к вычислительной математике, а не к экономике и менеджменту. Однако инженеру, менеджеру и экономисту полезно знать о свойствах интеллектуального инструмента, которым он пользуется.

С ростом мощности компьютеров необходимость применения изощренных математических методов снижается, поскольку во многих случаях время счета перестает быть лимитирующим фактором, оно весьма мало (доли секунд). Поэтому разберем лишь три метода.

Простой перебор. Возьмем некоторый многомерный параллелепипед, в котором лежит многогранник, задаваемый ограничениями. Как его построить? Например, если имеется ограничение типа $2X_1 + 5X_2 \le 10$, то, очевидно,

$$0 \le X_1 \le {}^{10}/_{2} = 5 \text{ и } 0 \le X_2 \le {}^{10}/_{5} = 2.$$

Аналогичным образом от линейных ограничений общего вида можно перейти к ограничениям на отдельные переменные. Остается взять максимальные границы по каждой переменной. Если многогранник, задаваемый ограничениями, неограничен, как было в задаче о диете, можно похожим, но несколько более сложным образом выделить его «обращенную» к началу координат часть, содержащую решение, и заключить ее в многомерный параллелепипед.

Проведем перебор точек параллелепипеда с шагом $1/10^n$ последовательно при $n = 2, 3, ...$, вычисляя значения целевой функции и проверяя выполнение ограничений. Из всех точек, удовлетворяющих ограничениям, возьмем ту, в которой целевая функция максимальна. Решение найдено! (Более строго выражаясь, найдено с точностью до $1/10^n$.)

Направленный перебор. Начнем с точки, удовлетворяющей ограничениям (ее можно найти простым перебором). Будем последовательно (или случайно — с помощью так называемого метода случайного поиска) менять ее координаты на определенную величину Δ, каждый раз в точку с более высоким значением целевой функции. Если выйдем на плоскость ограничения, будем двигаться по ней (находя одну из координат по уравнению ограничения). Затем движение по ребру (когда два ограничения-неравенства переходят в равенства)... Остановка — в вершине линейного многогранника. Решение найдено! (Более строго выражаясь, найдено с точностью до Δ; если необходимо, в окрестности найденного решения проводим направленный перебор с шагом $\Delta/2$, $\Delta/4$ и т.д.)

Симплекс-метод. Этот один из первых специализированных методов оптимизации, нацеленный на решение задач линейного программирования, в то время как методы простого и направленного перебора могут быть применены для решения практически любой задачи оптимизации. Симплекс-метод предложен американцем Г. Данцигом в 1951 г. Основная его идея состоит в продвижении по выпуклому многограннику ограничений от вершины к вершине, при котором на каждом шаге значение целевой функции улучшается до тех пор, пока не будет достигнут оптимум. Разберем пример на основе данных табл. 4.2.

Рассмотрим задачу линейного программирования, сформулированную ранее при рассмотрении оптимизации номенклатуры и объемов выпуска:

$$F = 15X_1 + 12X_2 + 14X_3 \to \max;$$

$$X_1 / 200 + X_2 / 300 + X_3 / 120 \le 100;$$

$$X_1 / 300 + X_2 / 100 + X_3 / 100 \le 100;$$

$$X_3 / 80 \le 100.$$

Неотрицательность переменных не будем специально указывать, поскольку в задачах линейного программирования это предположение всегда принимается.

В соответствии с симплекс-методом введем так называемые «свободные переменные» X_4, X_5, X_6, соответствующие недоиспользованным мощностям, т.е. от системы неравенств перейдем к системе уравнений:

$$X_1 / 200 + X_2 / 300 + X_3 / 120 + X_4 = 100;$$

$$X_1 / 300 + X_2 / 100 + X_3 / 100 + X_5 = 100;$$

$$X_3 / 80 + X_6 = 100;$$

$$15X_1 + 12X_2 + 14X_3 = F.$$

У этой системы имеется очевидное решение, соответствующее одной из вершин многогранника допустимых значений переменных:

$$X_1 = X_2 = X_3 = 0; X_4 = X_5 = X_6 = 100; F = 0.$$

В терминах исходной задачи это означает, что ничего не надо выпускать. Такое решение приемлемо только на период летних отпусков.

В соответствии с симплекс-методом выбираем переменную, которая входит в целевую функцию F с самым большим положительным коэффициентом. Это X_1.

Сравниваем частные от деления свободных членов в первых трех уравнениях на коэффициенты при только что выбранной переменной X_1:

$$100 / (^1/_{200}) = 20\,000; 100 / (^1/_{300}) = 30\,000; 100/0 = +\infty.$$

Выбираем строку из системы уравнений, которой соответствует минимальное из всех положительных отношений. В рассматриваемом примере — это первая строка, которой соответствует отношение 20 000.

Умножим первую строку на 200, чтобы получить X_1 с единичным коэффициентом:

$$X_1 + 2/3X_2 + 2/1{,}2X_3 + 200X_4 = 20\,000.$$

Затем умножим вновь полученную строку на (–1/300) и сложим со второй строкой, чтобы исключить член с X_1, получим:

$$7/900X_2 + 4/900X_3 - 2/3X_4 + X_5 = {}^{100}/_3.$$

Ту же преобразованную первую строку умножим на (–15) и сложим со строкой, в правой части которой стоит F, получим:

$$2X_2 - 11X_3 - 3000X_4 = F - 300\,000.$$

В результате система уравнений преобразуется к виду, в котором переменная X_1 входит только в первое уравнение:

$$X_1 + 2/3X_2 + 2/1{,}2X_3 + 200X_4 = 20\,000;$$

$$7/900X_2 + 4/900X_3 - 2/3X_4 + X_5 = {}^{100}/_3;$$

$$X_3 / 80 + X_6 = 100;$$

$$2X_2 - 11\,X_3 - 3000\,X_4 = F - 300\,000.$$

Очевидно, у новой системы имеется улучшенное по сравнению с исходным решение, соответствующее другой вершине выпуклого многогранника в шестимерном пространстве:

$$X_1 = 20\,000; X_2 = X_3 = X_4 = 0; X_5 = {}^{100}/_3; X_6 = 100; F = 300\,000.$$

В терминах исходной задачи это решение означает, что надо выпускать только кухни. Такое решение приемлемо, если допустимо выпускать только один вид продукции.

Повторим описанную операцию. В строке с F имеется еще один положительный коэффициент — при X_2 (если бы положительных коэффициентов было несколько, мы взяли бы максимальный из них). На основе коэффициентов при X_2 (а не при X_1, как в первый раз) образуем частные от деления соответствующих свободных членов на эти коэффициенты:

$$20\,000 / (^2/_3) = 30\,000;\ (^{100}/_3) / (^7/_{900}) = 30\,000/7;\ 100/0 = +\infty.$$

Таким образом, нужно выбрать вторую строку, для которой имеем наименьшее положительное отношение 30 000/7. Вторую строку умножим на $^{900}/_7$ (чтобы коэффициент при X_2 равнялся 1). Затем добавим обновленную строку ко всем строкам, содержащим X_2, предварительно умножив их на подходящие числа, т.е. такие, чтобы все коэффициенты при X_2 стали бы после сложения равны 0, за исключением коэффициента второй строки, который уже стал равняться 1. Получим систему уравнений:

$$X_1 + 9/7X_3 + 1800/7X_4 - 600/7X_5 = 120\,000/7;$$

$$X_2 + 4/7X_3 - 600/7X_4 + 900/7X_5 = 30\,000/7;$$

$$X_3 / 80 + X_6 = 100;$$

$$-85/7X_3 - 19\,800/7X_4 - 1800/7X_5 = F - 308\,571.$$

Поскольку все переменные неотрицательны, то из последнего уравнения следует, что прибыль F достигает своего максимального значения, равного 308 571, при $X_3 = X_4 = X_5 = 0$.

Из остальных уравнений следует, что при этом $X_1 = 120\,000/7 = 17\,143$; $X_2 = 30\,000/7 = 4286$; $X_6 = 100$.

Поскольку в строке с F не осталось ни одного положительного коэффициента при переменных, то алгоритм симплекс-метода закончил свою работу, оптимальное решение найдено.

Практические рекомендации таковы: надо выпустить 17 143 кухни, вчетверо меньше, т.е. 4286, кофемолок, самоваров не выпускать вообще. При этом прибыль будет максимальной и равной 308 571. Все производственное оборудование будет полностью загружено за исключением линии по сборке самоваров.

Транспортная задача. Различные технико-экономические и экономические задачи менеджмента, от оптимальной загрузки станка и раскройки стального листа или полотна ткани до анализа межотраслевого баланса и оценки темпов роста экономики страны в целом, приводят к необходимости решения тех или иных задач линейного программирования. Имеется обширный перечень публикаций, посвященных многочисленным применениям линейного программирования в металлургии, угольной, химической, нефтяной, бумажной и прочих отраслях промышленности, в проблемах транспорта и связи, планирования производства, конструирования и хранения продукции, сельском хозяйстве, в научных исследованиях, в том числе экономических, и даже при регулировании уличного движения.

В качестве очередного примера рассмотрим так называемую «транспортную задачу». Имеются склады, запасы на которых известны. Известны потребители и объемы их потребностей. Необходимо доставить товар со складов потребителям. Можно по-разному организовать «прикрепление» потребителей к складам, т.е. установить, с какого склада какому потребителю и сколько везти. Кроме того, известна стоимость доставки единицы товара с определенного склада определенному потребителю. Требуется минимизировать издержки по перевозке.

Например, может идти речь о перевозке песка — сырья для производства силикатных кирпичей. В Москву песок обычно доставляется самым дешевым транспортом — водным. Поэтому в качестве складов можно рассматривать порты, а в качестве запасов — их суточную пропускную способность. Потребителями являются кирпичные заводы, а их потребности определяются суточным производством (в соответствии с имеющимися заказами). Для доставки необходимо загрузить автотранспорт, проехать по определенному маршруту и разгрузить его. Стоимость этих операций рассчитывается по известным правилам, на которых не имеет смысла останавливаться. Поэтому затраты на доставку товара с определенного склада тому или иному потребителю можно считать известными.

Рассмотрим пример транспортной задачи, исходные данные к которой представлены в табл. 4.3. В ней, кроме объемов потребностей и величин запасов, приведены стоимости доставки единицы товара со склада i ($i = 1, 2, 3$) потребителю j ($j = 1, 2, 3, 4$). Например, самая дешевая доставка — со склада 2 потребителям 1 и 3, а также со склада 3 потребителю 2. Однако на складе 2 имеется 80 единиц товара, а потребителям 1 и 3 требуется 50 + 70 = 120 единиц, поэтому к ним придется везти товар и с других складов. Обратите внимание, что в табл. 4.3 запасы на складах равны суммарным потребностям. Для примера с доставкой песка кирпичным заводам это вполне естественное ограничение — при невыполнении такого ограничения либо порты будут засыпаны горами песка, либо кирпичные заводы не выполнят заказы.

Таблица 4.3

Исходные данные к транспортной задаче

Склад, потребности	Потребитель				Запасы на складах
	1	2	3	4	
Склад 1	2	5	5	5	60
Склад 2	1	2	1	4	80
Склад 3	3	1	5	2	60
Потребности	50	40	70	40	200

Надо спланировать перевозки, т.е. выбрать объемы X_{ij} поставок товара со склада i потребителю j, где $i = 1, 2, 3$; $j = 1, 2, 3, 4$. Таким образом, всего в задаче имеется 12 переменных. Они удовлетворяют двум группам ограничений. Во-первых, заданы запасы на складах:

$$X_{11} + X_{12} + X_{13} + X_{14} = 60;$$

$$X_{21} + X_{22} + X_{23} + X_{24} = 80;$$

$$X_{31} + X_{32} + X_{33} + X_{34} = 60.$$

Во-вторых, известны потребности клиентов:

$$X_{11} + X_{21} + X_{31} = 50;$$

$$X_{12} + X_{22} + X_{32} = 40;$$

$$X_{13} + X_{23} + X_{33} = 70;$$

$$X_{14} + X_{24} + X_{34} = 40.$$

Итак, всего 7 ограничений типа равенств. Кроме того, все переменные неотрицательны — еще 12 ограничений.

Целевая функция — издержки по перевозке, которые необходимо минимизировать:

$$F = 2X_{11} + 5X_{12} + 4X_{13} + 5X_{14} + X_{21} + 2X_{22} + X_{23} + 4X_{24} + \\ + 3X_{31} + X_{32} + 5X_{33} + 2X_{34} \to \min.$$

Кроме обсуждаемой, рассматриваются также различные иные варианты транспортной задачи. Например, если доставка производится вагонами, то объемы поставок должны быть кратны вместимости вагона.

Количество переменных и ограничений в транспортной задаче таково, что для ее решения не обойтись без компьютера и соответствующего программного продукта.

Линейное программирование имеет дело с числовыми переменными. Если вспомнить общую постановку оптимизационной задачи, приведенную в начале главы, то X — вектор в конечномерном линейном пространстве, A — многогранник в таком пространстве. Рассмотрим несколько задач оптимизации, в которых X и A имеют иную математическую природу.

4.2. ЦЕЛОЧИСЛЕННОЕ ПРОГРАММИРОВАНИЕ

Задачи оптимизации, в которых переменные принимают целочисленные значения, относятся к целочисленному программированию. Рассмотрим несколько таких задач.

Задача о выборе оборудования. На приобретение оборудования для нового участка цеха выделено 20 000 дол. При этом можно занять

площадь не более 38 м^2. Имеется возможность приобрести станки типа А и станки типа Б. При этом станки типа А стоят 5000 дол., занимают площадь 8 м^2 (включая необходимые технологические проходы) и имеют производительность 7 тыс. единиц продукции за смену. Станки типа Б стоят 2000 дол., занимают площадь 4 м^2 и имеют производительность 3 тыс. единиц продукции за смену. Необходимо рассчитать оптимальный вариант приобретения оборудования, обеспечивающий при заданных ограничениях максимум общей производительности участка.

Пусть X — число станков типа А, а Y — число станков типа Б, входящих в комплект оборудования. Требуется выбрать комплект оборудования так, чтобы максимизировать производительность C участка (в тысячах единиц за смену):

$$C = 7X + 3Y \to \max.$$

При этом должны быть выполнены следующие ограничения:
по стоимости (в тысячах долларов):

$$5X + 2Y \le 20;$$

по занимаемой площади (в квадратных метрах):

$$8X + 4Y \le 38,$$

а также вновь появляющиеся специфические ограничения по целочисленности, а именно:

$$X \ge 0;\ Y \ge 0, \text{ где } X \text{ и } Y \text{ — целые числа.}$$

Сформулированная математическая задача отличается от задачи линейного программирования только последним условием целочисленности. Однако наличие этого условия позволяет (в данном конкретном случае) легко решить задачу перебором. Действительно, как ограничение по стоимости, так и ограничение по площади дают, что $X \le 4$. Значит, X может принимать лишь одно из пяти значений: 0, 1, 2, 3, 4.

Если $X = 4$, то из ограничения по стоимости следует, что $Y = 0$, а потому $C = 7X = 28$.

Если $X = 3$, то из первого ограничения вытекает, что $Y \le 2$, из второго $Y \le 3$. Значит, максимальное C при условии выполнения ограничений достигается при $Y = 2$, а именно: $C = 21 + 6 = 27$.

Если $X = 2$, то из первого ограничения следует, что $Y \le 5$, из второго также $Y \le 5$. Значит, максимальное C при условии выполнения ограничений достигается при $Y = 5$, а именно: $C = 14 + 15 = 29$.

Если $X = 1$, то из первого ограничения имеем $Y \le 7$, из второго также $Y \le 7$. Значит, максимальное C при условии выполнения ограничений достигается при $Y = 7$, а именно: $C = 7 + 21 = 28$.

Если $X = 0$, то из первого ограничения вытекает $Y \leq 10$, из второго $Y \leq 9$. Значит, максимальное C при условии выполнения ограничений достигается при $Y = 9$, а именно $C = 27$.

Все возможные случаи рассмотрены. Максимальная производительность $C = 29$ тыс. единиц продукции за смену, достигается при $X = 2$, $Y = 5$. Следовательно, надо покупать 2 станка типа А и 5 станков типа Б.

Задача о ранце. Общий вес ранца заранее ограничен. Какие предметы положить в ранец, чтобы общая полезность отобранных предметов была максимальна? Вес каждого предмета известен.

Есть много эквивалентных формулировок. Например, можно вместо ранца рассматривать космический аппарат — спутник Земли, а в качестве предметов — научные приборы. Тогда задача интерпретируется как отбор приборов для запуска на орбиту. Правда, при этом предполагается решенной предварительная задача — оценка сравнительной ценности исследований, для которых нужны те или иные приборы.

С точки зрения экономики предприятия и организации производства более актуальна другая интерпретация задачи о ранце, в которой в качестве «предметов» рассматриваются заказы (или варианты выпуска партий тех или иных товаров), в качестве полезности — прибыль от выполнения того или иного заказа, а в качестве веса — себестоимость заказа.

Перейдем к математической постановке. Предполагается, что имеется n предметов, и для каждого из них необходимо решить, класть его в ранец или не класть. Для описания решения вводятся булевы переменные X_k, $k = 1, 2, ..., n$ (т.е. переменные, принимающие два значения, а именно 0 и 1). При этом $X_k = 1$, если предмет размещают в ранце, и $X_k = 0$, если нет, $k = 1, 2, ..., n$. Для каждого предмета известны две константы: A_k — масса k-го предмета и C_k — полезность k-го предмета, $k = 1, 2, ..., n$. Максимально возможную вместимость ранца обозначим B. Оптимизационная задача имеет вид:

$$C_1X_1 + C_2X_2 + C_3X_3 + ... + C_nX_n \to \max;$$

$$A_1X_1 + A_2X_2 + A_3X_3 + ... + A_nX_n \leq B.$$

В отличие от предыдущих задач управляющие параметры X_k, $k = 1, 2, ..., n$, принимают значения из множества, содержащего два элемента — 0 и 1.

К целочисленному программированию относятся задачи размещения (производственных объектов), теории расписаний, календарного и оперативного планирования, назначения персонала и т.д.

Укажем два распространенных метода решения задач целочисленного программирования.

Метод приближения непрерывными задачами. В соответствии с ним сначала решается задача линейного программирования без учета целочисленности, а затем в окрестности оптимального решения ищутся целочисленные точки.

Методы направленного перебора. Из них наиболее известен метод ветвей и границ. Суть метода такова. Каждому подмножеству X множества возможных решений X_0 ставится в соответствие число — «граница» $A(X)$. При решении задачи минимизации необходимо, чтобы $A(X_1) \geq A(X_2)$, если X_1 входит в X_2 или совпадает с X_2.

Каждый шаг метода ветвей и границ состоит в делении выбранного на предыдущем шаге множества X_C на два — X_{1C} и X_{2C}. При этом пересечение X_{1C} и X_{2C} пусто, а их объединение совпадает с X_C. Затем вычисляют границы $A(X_{1C})$ и $A(X_{2C})$ и выделяют «ветвь» X_{C+1} — из двух множеств X_{1C} и X_{2C}, для которого граница меньше. Алгоритм прекращает работу, когда диаметр вновь выделенной ветви оказывается меньше заранее заданного малого числа.

Для каждой конкретной задачи целочисленного программирования (другими словами, дискретной оптимизации) метод ветвей и границ реализуется по-своему. Есть много модификаций этого метода. Однако менеджеру нет необходимости вникать в подробности, относящиеся к вычислительной математике. Вместе с тем он должен знать о возможностях, предоставляемых ему теорией оптимизации.

4.3. ТЕОРИЯ ГРАФОВ И ОПТИМИЗАЦИЯ

Один из разделов дискретной математики, часто используемый при принятии решений, — теория графов [9]. *Граф* — это совокупность точек, называемых *вершинами графа*, некоторые из которых соединены дугами (дуги называют также *ребрами*). Примеры графов приведены на рис. 4.5.

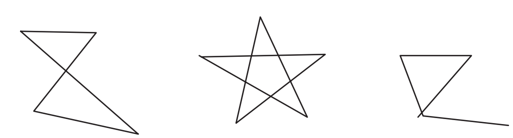


Рис. 4.5. Примеры графов

На только что введенное понятие графа «навешиваются» новые свойства. Исходному объекту приписывают новые качества. Например, вводится и используется понятие ориентированного графа. В таком графе дуги имеют стрелки, направленные от одной вершины к другой. Примеры ориентированных графов даны на рис. 4.6.

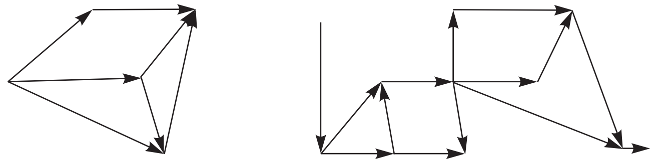


Рис. 4.6. Примеры ориентированных графов

Ориентированный граф был бы полезен, например, для иллюстрации организации перевозок в транспортной задаче. В экономике дугам ориентированного или обычного графа часто приписывают числа, например стоимость проезда или перевозки груза из пункта А (начальная вершина дуги) в пункт Б (конечная вершина дуги).

Рассмотрим несколько типичных задач принятия решений, связанных с оптимизацией на графах.

Задача коммивояжера. Требуется посетить все вершины графа и вернуться в исходную вершину, минимизировав затраты на проезд (или минимизировав время).

Исходные данные здесь — это граф, дугам которого приписаны положительные числа — затраты на проезд или время, необходимое для продвижения из одной вершины в другую. В общем случае граф является ориентированным, и каждые две вершины соединяют две дуги — туда и обратно. Действительно, если пункт А расположен на горе, а пункт Б — в низине, то время на проезд из А в Б, очевидно, меньше времени на обратный проезд из Б в А.

Многие постановки экономического содержания сводятся к задаче коммивояжера, например:

- составить наиболее выгодный маршрут обхода наладчика в цехе (контролера, охранника, милиционера), отвечающего за должное функционирование заданного множества объектов (каждый из этих объектов моделируется вершиной графа);
- составить наиболее выгодный маршрут доставки деталей рабочим или хлеба с хлебозавода по заданному числу булочных и других торговых точек (парковка у хлебозавода).

Задача о кратчайшем пути. Как кратчайшим путем попасть из одной вершины графа в другую? В терминах производственного менеджмента: как кратчайшим путем (и следовательно, с наименьшим расходом топлива и времени, наиболее дешево) попасть из пункта А в пункт Б? Для решения этой задачи каждой дуге ориентированного графа должно быть сопоставлено число — время движения по этой дуге от начальной вершины до конечной (рис. 4.7).

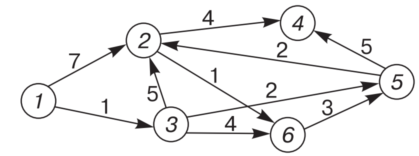


Рис. 4.7. Исходные данные к задаче о кратчайшем пути

Ситуацию можно описать не только ориентированным графом с весами, приписанными дугам, но и таблицей, в которой двум вершинам — началу и концу пути — ставится в соответствие время в пути. Пути без промежуточных остановок рассматриваются в табл. 4.4. Более сложные маршруты составляются из элементарных отрезков, перечисленных в табл. 4.4.

Таблица 4.4

Исходные данные к задаче о кратчайшем пути

Начало дуги	Конец дуги	Время в пути, ч
1	*2*	7
1	*3*	1
2	*4*	4
2	*6*	1
3	*2*	5
3	*5*	2
3	*6*	3
5	*2*	2
5	*4*	5
6	*5*	3

Спрашивается в задаче: как кратчайшим путем попасть из вершины 1 в вершину 4?

Решение. Введем обозначение: $C(T)$ — длина кратчайшего пути из вершины *1* в вершину T. (Поскольку любой путь, который надо рассмотреть, состоит из дуг, а дуг конечное число, и каждая входит не более одного раза, то претендентов на кратчайший путь конечное число, и минимум из конечного числа элементов всегда достигается.) Рассматриваемая задача состоит в вычислении $C(4)$ и указании пути, на котором этот минимум достигается.

Для исходных данных, представленных на рис. 4.7 и в табл. 4.4, в вершину *3* входит только одна стрелка, как раз из вершины *1*, и около этой стрелки стоит ее длина, равная 1, поэтому $C(3) = 1$. Кроме того, очевидно, что $C(1) = 0$.

В вершину *4* можно попасть либо из вершины *2*, пройдя путь за 4 ч, либо из вершины *5*, пройдя путь за 5 ч. Поэтому справедливо соотношение

$$C(4) = \min\{C(2) + 4; C(5) + 5\}.$$

Таким образом, проведена реструктуризация (упрощение) задачи — нахождение $C(4)$ сведено к нахождению $C(2)$ и $C(5)$.

В вершину *5* можно попасть либо из вершины *3*, пройдя путь за 2 ч, либо из вершины *6*, пройдя путь за 3 ч. Поэтому справедливо соотношение

$$C(5) = \min\{C(3) + 2; C(6) + 3\}.$$

Мы знаем, что $C(3) = 1$. Поэтому

$$C(5) = \min\{3; C(6) + 3\}.$$

Поскольку очевидно, что $C(6)$ — положительное число, то из последнего соотношения вытекает, что $C(5) = 3$.

В вершину *2* можно попасть либо из вершины *1*, пройдя путь за 7 ч, либо из вершины *3*, пройдя путь за 5 ч, либо из вершины *5*, пройдя путь за 2 ч. Поэтому справедливо соотношение

$$C(2) = \min\{C(1) + 7; C(3) + 5; C(5) + 2\}.$$

Нам известно, что $C(1) = 0$, $C(3) = 1$, $C(5) = 3$. Поэтому

$$C(2) = \min\{0 + 7; 1 + 5; 3 + 2\} = 5.$$

Теперь мы можем найти $C(4)$:

$$C(4) = \min\{C(2) + 4; C(5) + 5\} = \min\{5 + 4; 3 + 5\} = 8.$$

Таким образом, время на кратчайший путь равно 8. Из последнего соотношения ясно, что в вершину *4* надо идти через вершину *5*. Возвращаясь к вычислению $C(5)$, видим, что в вершину *5* надо идти через вершину *3*. А в вершину *3* можно попасть только из вершины *1*. Итак, кратчайший путь таков:

$$1 \to 3 \to 5 \to 4.$$

Задача о кратчайшем пути для конкретных исходных данных (рис. 4.7 и табл. 4.4) полностью решена.

Оптимизационные задачи на графах, возникающие при подготовке управленческих решений в производственном менеджменте, весьма многообразны. Рассмотрим в качестве примера еще одну задачу, связанную с перевозками.

Задача о максимальном потоке. По каким маршрутам послать максимально возможное количество грузов из начального в конечный пункт, если пропускная способность путей между пунктами ограничена?

Для решения этой задачи каждой дуге ориентированного графа, соответствующего транспортной системе, должно быть сопоставлено число — пропускная способность этой дуги. Рассмотрим пример (рис. 4.8).

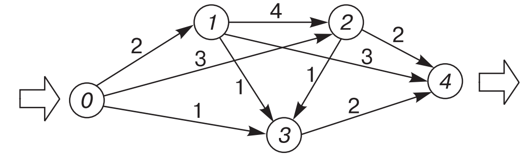


Рис. 4.8. Исходные данные к задаче о максимальном потоке

Исходные данные о транспортной системе, например внутризаводской, приведенные на рис. 4.8, можно также задать таблицей (табл. 4.5).

Таблица 4.5

Исходные данные к задаче о максимальном потоке

Пункт отправления	Пункт назначения	Пропускная способность, единиц
0	*1*	2
0	*2*	3
0	*3*	1
1	*2*	4
1	*3*	1
1	*4*	3
2	*3*	1
2	*4*	2
3	*4*	2

Решение задачи о максимальном потоке может быть получено из следующих соображений.

Очевидно, что максимальная пропускная способность транспортной системы не превышает шести единиц, поскольку не более шести единиц грузов можно направить из начального пункта *0*, а именно: две единицы в пункт *1*, три единицы в пункт *2* и одну единицу в пункт *3*.

Далее надо добиться, чтобы все шесть вышедших из пункта *0* единиц груза достигли конечного пункта *4*. Очевидно, две единицы груза, пришедшие в пункт *1*, можно непосредственно направить в пункт *4*. Пришедшие в пункт *2* грузы придется разделить: две единицы сразу направить в пункт *4*, а одну единицу — в промежуточный пункт *3* (из-за ограниченной пропускной способности участка между пунктами *2* и *4*). В пункт *3* доставлены грузы: одна единица из пункта *0* и одна единица из пункта *2*. Их направляем в пункт *4*.

Итак, максимальная пропускная способность рассматриваемой транспортной системы — шесть единиц груза. При этом не используются внутренние участки (ветки) между пунктами *1* и *2*, а также между пунктами *1* и *3*. Не догружена ветка между пунктами *1* и *4* — по ней направлены две единицы груза при пропускной способности в три единицы.

Решение можно представить в виде табл. 4.6.

Таблица 4.6

Решение задачи о максимальном потоке

Пункт отправления	Пункт назначения	План перевозок	Пропускная способность
0	*1*	2	2
0	*2*	3	3
0	*3*	1	1
1	*2*	0	4
1	*3*	0	1
1	*4*	2	3
2	*3*	1	1
2	*4*	2	2
3	*4*	2	2

Задача линейного программирования при максимизации потока. Пусть X_{KM} — объем перевозок из пункта K в пункт M. Согласно рис. 4.8 $K = 0, 1, 2, 3$, $M = 1, 2, 3, 4$, причем перевозки возможны лишь в пункт с бóльшим номером. Значит, всего имеется 9 переменных X_{KM}, а именно: $X_{01}, X_{02}, X_{03}, X_{12}, X_{13}, X_{14}, X_{23}, X_{24}, X_{34}$. Задача линейного программирования, нацеленная на максимизацию потока, имеет вид:

$$F \to \max;$$

$$X_{01} + X_{02} + X_{03} = F; \qquad (0)$$

$$-X_{01} + X_{12} + X_{13} + X_{14} = 0; \qquad (1)$$

$$-X_{02} - X_{12} + X_{23} + X_{24} = 0; \qquad (2)$$

$$-X_{03} - X_{13} - X_{23} + X_{34} = 0; \qquad (3)$$

$$-X_{14} - X_{24} - X_{34} = -F; \qquad (4)$$

$$X_{01} \leq 2;$$

$$X_{02} \leq 3;$$

$$X_{03} \leq 1;$$

$$X_{12} \leq 4;$$

$$X_{13} \leq 1;$$

$$X_{14} \leq 3;$$

$$X_{23} \leq 1;$$

$$X_{24} \leq 2;$$

$$X_{34} \leq 2;$$

$$X_{KM} \geq 0;\ K, M = 0, 1, 2, 3, 4;\ F \geq 0.$$

Здесь F — целевая функция, условие (0) описывает вхождение грузов в транспортную систему. Условия (1)—(3) задают балансовые соотношения для узлов 1—3 систем. Другими словами, для каждого из внутренних узлов входящий поток грузов равен выходящему потоку, грузы не скапливаются внутри системы и не «рождаются» в ней. Условие (4) — это условие «выхода» грузов из системы. Вместе с условием (0) оно составляет балансовое соотношение для системы в целом («вход» равен «выходу»).

Следующие девять неравенств задают ограничения на пропускную способность отдельных «веток» транспортной системы. Затем в системе ограничений задачи линейного программирования указана неотрицательность объемов перевозок и целевой функции. Ясно, что последнее неравенство вытекает из вида целевой функции (соотношения (0) или (4)) и неотрицательности объемов перевозок. Однако последнее неравенство несет некоторую общую информацию — через систему может быть пропущен либо положительный объем грузов, либо нулевой (например, если внутри системы происходит движение по кругу), но не отрицательный (он не имеет экономического смысла, но формальная математическая модель об этом «не знает»).

Многообразие оптимизационных задач. В различных проблемах принятия решений возникают самые разнообразные задачи оптимизации. Для их решения применяют те или иные методы, точные или приближенные. Задачи оптимизации часто используются в теоретико-экономических исследованиях. Достаточно вспомнить оптимизацию экономического роста страны с помощью матрицы межотраслевого баланса В. Леонтьева или микроэкономические задачи определения оптимального объема выпуска по функции издержек при фиксированной цене (или в условиях монополии) либо минимизации издержек при заданном объеме выпуска путем выбора оптимального соотношения факторов производства (с учетом платы за них).

Кроме затронутых ранее методов решения задач оптимизации, напомним о том, что гладкие функции оптимизируют, приравнивая нулю производную (для функций нескольких переменных — частные производные). При наличии ограничений используют множители Лагранжа. Эти методы обычно излагаются в курсах высшей математики и потому опущены здесь.

Представляют интерес задачи оптимизации с нечеткими переменными [70], а также задачи оптимизации, возникающие в эконометрике [89]. Например, метод наименьших квадратов, разобранный в следующей части, основан на решении задачи оптимизации. Итоговое мнение комиссии экспертов часто вычисляют как решение задачи оптимизации (часть III). Конкретные виды задач оптимизации и методы их решения рассматриваются в соответствующей литературе.

Контрольные вопросы

1. Как изображают на плоскости ограничения задачи линейного программирования? Покажите на примере решения (графически) задачи:

$$400W_1 + 450W_2 \to \min;$$
$$5W_1 + 10W_2 \ge 45;$$
$$20W_1 + 15W_2 \ge 80;$$
$$W_1 \ge 0,\ W_2 \ge 0.$$

2. Как решают задачи линейного программирования? Покажите на примере следующей задачи:

$$W_1 + 5W_2 \to \max;$$
$$0{,}1W_1 + W_2 \le 3{,}8;$$
$$0{,}25W_1 + 0{,}25W_2 \le 4{,}2;$$
$$W_1 \ge 0,\ W_2 \ge 0.$$

3. Как решают задачи целочисленного программирования? Покажите на примере такой задачи:

$$10X + 5Y \to \max;$$
$$8X + 3Y \le 40;$$
$$3X + 10Y \le 30;$$
$$X \ge 0,\ Y \ge 0;\ X \text{ и } Y \text{ — целые числа.}$$

4. Как решают задачи о ранце? Покажите на примере следующей задачи:

$$X_1 + X_2 + 2X_3 + 2X_4 + X_5 + X_6 \to \max;$$
$$0{,}5X_1 + X_2 + 1{,}5X_3 + 2X_4 + 2{,}5X_5 + 3X_6 \le 3.$$

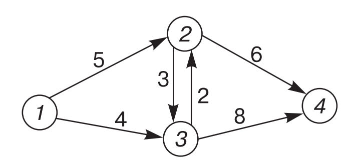
Управляющие параметры X_k, $k = 1, 2, ..., 6$, принимают значения из множества, содержащего два элемента — 0 и 1.

5. Как решают задачи коммивояжера для четырех городов? (Основное условие — маршрут должен быть замкнутым и не содержать повторных посещений.) Найдите решение при условии, что затраты на проезд будут равны приведенным в таблице.

Исходные данные к задаче коммивояжера

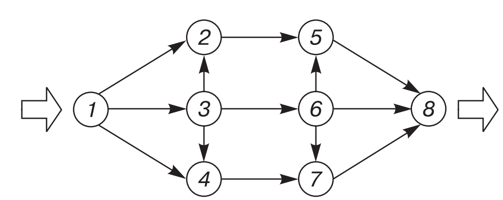
Город отправления	Город назначения	Затраты на проезд
А	Б	2
А	В	1
А	Д	5
Б	А	3
Б	В	2
Б	Д	1
В	А	4
В	Б	1
В	Д	2
Д	А	5
Д	Б	3
Д	В	3

6. Как найти кратчайший путь из пункта *1* в пункт *4*? Транспортная сеть (с указанием расстояний) приведена на рисунке.



Исходные данные к задаче о кратчайшем пути

7. Как послать максимальное количество грузов из начального пункта *1* в конечный пункт *8*, если пропускная способность путей между пунктами транспортной сети ограничена (см. рисунок и таблицу)?



Транспортная сеть к задаче о максимальном потоке

Исходные данные к задаче о максимальном потоке

Пункт отправления	Пункт назначения	Пропускная способность
1	*2*	1
1	*3*	2
1	*4*	3
2	*5*	2
3	*2*	2
3	*4*	2
3	*6*	1
4	*7*	4
5	*8*	3
6	*5*	2
6	*7*	1
6	*8*	1
7	*8*	3

Темы докладов и рефератов

1. Классификация оптимизационных задач.
2. Решения, оптимальные по Парето.
3. Многокритериальные задачи оптимизации: различные методы свертки критериев.
4. Задачи оптимизации и нечеткие переменные.
5. Место метода множителей Лагранжа в теории оптимизации.

ГЛАВА 5
РЕГРЕССИЯ, КОРРЕЛЯЦИЯ И ПРОГНОЗИРОВАНИЕ

5.1. ВОССТАНОВЛЕНИЕ ЛИНЕЙНОЙ ЗАВИСИМОСТИ МЕЖДУ ДВУМЯ ПЕРЕМЕННЫМИ

Одним из наиболее важных этапов эконометрического моделирования является восстановление (выявление) зависимости между переменными на основе статистических данных, которая затем используется при организационно-экономическом моделировании, в частности для прогнозирования, оптимизации принятия решений.

Начнем с задачи точечного и доверительного оценивания линейной функции одной переменной $x(t) = at + b$.

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 линейной зависимостью

$$x_k = at_k + b + e_k, k = 1, 2, ..., n,$$

где a и b — параметры, неизвестные статистику и подлежащие оцениванию, а e_k — погрешности, искажающие зависимость.

Обычно оценивают параметры a и b линейной зависимости методом наименьших квадратов. Затем восстановленную зависимость используют, например, для точечного и интервального прогнозирования.

Метод наименьших квадратов. Немецкий математик Ф. Клейн, тщательно изучавший научное наследство своего великого соотечественника К. Гаусса, писал, что метод наименьших квадратов был разработан К. Гауссом в 1795 г. [29, с. 37]. (Как много позже сказано в [46, с. 181], «Гаусс указывает две даты: 1794 г. и 1795 г. Современные исследователи склонны считать, что верная дата — это 1794 г.».) Согласно этому методу для расчета наилучшей функции, приближающей линейным образом зависимость x от t, следует рассмотреть функцию двух переменных

$$f(a,b) = \sum_{k=1}^{n} (x_k - at_k - b)^2.$$

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:

$$\frac{\partial f(a,b)}{\partial a} = \sum_{k=1}^{n} 2(x_k - at_k - b)(-t_k);$$

$$\frac{\partial f(a,b)}{\partial b} = \sum_{k=1}^{n} 2(x_k - at_k - b)(-1).$$

Преобразуем правые части полученных соотношений. Вынесем за знак суммы общие множители 2 и (–1). Затем рассмотрим слагаемые. Раскрыв 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:

$$\frac{\partial f(a,b)}{\partial a} = (-2)\left(\sum_{i=1}^{n} x_i t_i - a\sum_{i=1}^{n} t_i^2 - b\sum_{i=1}^{n} t_i\right);$$

$$\frac{\partial f(a,b)}{\partial b} = (-2)\left(\sum_{i=1}^{n} x_i - a\sum_{i=1}^{n} t_i - bn\right).$$

Приравняем частные производные 0. Тогда в полученных уравнениях можно сократить множитель (–2). Оценки метода наименьших квадратов находим из системы двух линейных уравнений с двумя неизвестными a и b:

$$\sum_{i=1}^{n} x_i t_i - a\sum_{i=1}^{n} t_i^2 - b\sum_{i=1}^{n} t_i = 0;$$

$$\sum_{i=1}^{n} x_i - a\sum_{i=1}^{n} t_i - bn = 0.$$

Запись решения этой системы будет более компактной, если использовать не суммы, а средние арифметические:

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i,\ \ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,\ \ \overline{xt} = \frac{1}{n}\sum_{i=1}^{n} x_i t_i,\ \ \overline{t^2} = \frac{1}{n}\sum_{i=1}^{n} t_i^2.$$

Разделив обе части уравнений на n, перейдем к системе:

$$\overline{xt} - a\overline{t^2} - b\bar{t} = 0;$$

$$\bar{x} - a\bar{t} - b = 0.$$

Из второго уравнения получаем, что

$$b = \bar{x} - a\bar{t}.$$

Подставим в первое уравнение:

$$\overline{xt} - a\overline{t^2} - (\bar{x} - a\bar{t})\bar{t} = 0.$$

Раскрыв скобки, решаем это линейное уравнение с одной переменной a. Получаем оценки метода наименьших квадратов:

$$a^* = \frac{\overline{xt} - \bar{x}\bar{t}}{\overline{t^2} - (\bar{t})^2}, \quad b^* = \bar{x} - a^*\bar{t}. \qquad (5.1)$$

Следовательно, восстановленная функция имеет вид

$$x^*(t) = a^*t + b^*.$$

С точки зрения теории оптимизации равенство частных производных нулю — необходимое условие экстремума, но недостаточное. Однако в силу единственности решения системы линейных уравнений равенства (5.1) дают точку минимума, поскольку существование минимума вытекает из того, что в качестве области определения непрерывной функции $f(a, b)$ может рассматриваться некоторое ограниченное замкнутое множество. Если же имеются априорные ограничения на параметры, то формулы (5.1) не всегда применимы. Например, пусть $x(t)$ — издержки производства при выпуске продукции объема t. Тогда параметры линейной зависимости имеют экономическую интерпретацию: b — постоянные издержки (вне зависимости от объема производства), a — переменные издержки (на одну единицу выпущенной продукции). Очевидно, постоянные издержки неотрицательны: $b > 0$. Однако при расчетах по формулам (5.1) при «неудачных» исходных данных может быть получено значение $b^* < 0$. Очевидно, отрицательным значением постоянных издержек пользоваться нельзя. В таком случае можно порекомендовать принять в исходной модели $b = 0$ и методом наименьших квадратов найти наилучшую оценку единственного параметра — переменных издержек a.

Пример 5.1. Пусть даны $n = 6$ пар чисел (t_k, x_k), $k = 1, 2, ..., 6$, представленных во втором и третьем столбцах табл. 5.1 (строки 1—6). Расчеты по методу наименьших квадратов удобно проводить с помощью таблицы, подобной табл. 5.1, последовательно заполняя ее столбцы либо вручную, либо на компьютере с помощью электронной таблицы EXCEL или иного программного продукта.

В соответствии с формулами (5.1) для вычисления оценок метода наименьших квадратов необходимо рассчитать четыре величины: $\bar{t}; \bar{x}; \overline{t^2}; \overline{xt}$. Для получения первых двух из них достаточно найти суммы чисел, представленных во втором и третьем столбцах табл. 5.1. Соответствующие суммы записаны в седьмой строке (обозначена символом Σ), а средние арифметические — в восьмой строке (Σ/n).

Для расчета двух оставшихся средних заполнены клетки столбцов (4) и (5), а затем проведено суммирование по этим столбцам. Все необходимые операции — поэлементное возведение в квадрат, перемножение столбцов, суммирование по столбцам — легко осуществить с помощью электронной таблицы EXCEL.

Остальные столбцы табл. 5.1 используются ниже при дальнейшем развертывании алгоритмов метода наименьших квадратов.

Таблица 5.1

Расчет по методу наименьших квадратов при восстановлении линейной функции одной переменной

Номер строки i	Расчетная величина							
	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)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
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	12	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	1 111	—	—	0,06	13,64
$\frac{\Sigma}{n}$	$\bar{t} = 5{,}67$	$\bar{x} = 26{,}83$	$\overline{t^2} = 42{,}67$	$\overline{xt} = 185{,}17$	—	—	—	$(\sigma^2)^* = 2{,}27$

В соответствии с формулами (5.1) оценки метода наименьших квадратов для приведенных в табл. 5.1 данных таковы:

$$a^* = \frac{185{,}17 - 26{,}83 \times 5{,}67}{42{,}67 - (5{,}67)^2} = \frac{33{,}04}{10{,}52} = 3{,}14;$$

$$b^* = 26{,}83 - 3{,}14 \times 5{,}67 = 9{,}03,$$

а восстановленная зависимость имеет вид

$$x^*(t) = 3{,}14\, t + 9{,}03.$$

Варианты оценок метода наименьших квадратов. Метод наименьших квадратов рассматривается во многих литературных источниках,

и формулы для оценок параметров зачастую имеют различный вид. Однако все они переходят друг в друга в результате тождественных преобразований.

Для развертывания вероятностно-статистической теории нам понадобится другая параметризация линейной зависимости, а именно

$$x_k = a(t_k - \bar{t}) + d + e_k, \quad k = 1, 2, \ldots, n,$$

где a и d — параметры, неизвестные статистику и подлежащие оцениванию, а e_k — погрешности, искажающие зависимость; среднее арифметическое моментов времени $\bar{t}$ введено в модель для облегчения математико-статистических выкладок.

Для оценивания параметров a и d необходимо, согласно методу наименьших квадратов, минимизировать функцию

$$F(a, d) = \sum_{k=1}^{n} \left[x_i - a(t_i - \bar{t}) - d \right]^2.$$

Как и раньше, вычисляем частные производные и приравниваем их 0. Поскольку

$$\sum_{k=1}^{n} (t_k - \bar{t}) = 0, \qquad (5.2)$$

то уравнения приобретают вид

$$\sum_{k=1}^{n} x_i (t_i - \bar{t}) - a \sum_{k=1}^{n} (t_i - \bar{t})^2 = 0;$$

$$\sum_{k=1}^{n} x_k - dn = 0.$$

Следовательно, оценки метода наименьших квадратов имеют вид

$$a^* = \frac{\sum_{k=1}^{n} x_k (t_k - \bar{t})}{\sum_{k=1}^{n} (t_k - \bar{t})^2}; \quad d^* = \bar{x} = \frac{1}{n} \sum_{k=1}^{n} x_k. \qquad (5.3)$$

В силу соотношения (5.2) оценку a^* можно записать в более симметричном виде:

$$a^* = \frac{\sum_{k=1}^{n} (x_k - \bar{x})(t_k - \bar{t})}{\sum_{k=1}^{n} (t_i - \bar{t})^2}.$$

Эту оценку нетрудно преобразовать и к виду

$$a^* = \frac{\sum_{i=1}^{n} x_i t_i - \frac{1}{n}\sum_{i=1}^{n} x_{ii} \sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} t_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} t_i\right)^2}. \qquad (5.4)$$

Следовательно, восстановленная функция, с помощью которой можно прогнозировать и интерполировать, имеет вид

$$x^*(t) = a^*(t - \bar{t}) + d^*.$$

Обратим внимание на то, что использование $\bar{t}$ в последней формуле ничуть не ограничивает ее общность. Сравним с ранее рассмотренной моделью вида

$$x_k = ct_k + b + e_k,\ k = 1, 2, ..., n.$$

Ясно, что

$$c = a,\ \ b = d - a\bar{t}.$$

Аналогичным образом связаны оценки параметров:

$$c^* = a^*,\ \ b^* = d^* - a^*\bar{t}.$$

Для данных табл. 5.1 в соответствии с формулой (5.3) $d^* = 26{,}83$, а согласно формуле (5.4)

$$a^* = \frac{1111 - \frac{1}{6}161 \times 34}{256 - \frac{1}{6}(34)^2} = \frac{1111 - 912{,}33}{256 - 192{,}67} = \frac{198{,}67}{63{,}33} = 3{,}14.$$

Следовательно, прогностическая функция (т.е. восстановленная зависимость) имеет вид:

$$x^*(t) = 3{,}14(t - 5{,}67) + 26{,}83 = 3{,}14t - 3{,}14 \times 5{,}67 + 26{,}83 =$$
$$= 3{,}14t - 17{,}80 + 26{,}83 = 3{,}14t + 9{,}03.$$

Естественно, результат тот же, что и при использовании первоначальной параметризации (формы линейной зависимости).

Восстановленные значения и оценка точности приближения. Следующий (второй) этап анализа данных — оценка точности восстановления (приближения) зависимости (функции) методом наименьших квадратов. Сначала рассматриваются так называемые «восстановленные значения»:

$$\hat{x}_i = x^*(t_i),\ \ i = 1, 2, \ldots, n.$$

Восстановленные значения — это те значения, которые полученная в результате расчетов прогностическая функция принимает в тех точках, в которых известны истинные значения зависимой переменной x_i.

Вполне естественно сравнить восстановленные и истинные значения. Это и сделано в столбцах (6)—(8) табл. 5.1. Для простоты расчетов в столбце (6) представлены произведения a^*t_i, столбец (7) отличается от (6) добавлением константы $b^* = 9{,}03$ и содержит восстановленные значения. Столбец (8) — это разность столбцов (3) и (7).

Непосредственный анализ столбца (8) табл. 5.1 показывает, что содержащиеся в нем числа сравнительно невелики по величине по сравнению со столбцом (3) — 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 (неправильно оценен коэффициент a).

Верно следующее утверждение.

Теорема 5.1. Справедливо тождество

$$\sum_{i=1}^{n}(x_i - \hat{x}_i) = 0.$$

Доказательство этой теоремы оставляем читателю в качестве упражнения.

Однако сумма по столбцу (8) 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 соответствии с ним сумма элементов столбца (8) должна быть мала по сравнению с элементами этого столбца.

Представляет интерес относительная погрешность восстановления. В точке t_k — это величина

$$\left|\frac{x_i - \hat{x}_i}{x_i}\right|.$$

Для данных табл. 5.1 — это величины

$$\frac{0{,}17}{12}, \frac{1{,}55}{20}, \frac{1{,}59}{20}, \frac{0{,}99}{32}, \frac{2{,}29}{35}, \frac{1{,}57}{42}.$$

Максимальной из них является 1,59/20 = 0,08. Точность восстановления естественно выразить в процентах:

$$\max_{1\le k\le n}\left|\frac{x_i-\hat{x}_i}{x_i}\right|\times 100=8\%.$$

В социально-экономических исследованиях точность восстановления 10—15% признается хорошей. Конечно, в астрономических вычислениях, например при восстановлении орбиты астероида Церера (именно для решения этой задачи К. Гаусс разработал метод наименьших квадратов), точность должна быть гораздо выше (т.е рассматриваемый показатель — максимум модуля относительных погрешностей — должен быть гораздо меньше).

Непараметрическая вероятностная модель. Для получения оценок параметров и прогностической функции 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.

Формулировка модели такова. *Пусть значения независимой переменной t детерминированы, а погрешности e_k, $k = 1, 2, ..., n$, — независимые одинаково распределенные случайные величины с нулевым математическим ожиданием и дисперсией σ^2, неизвестной статистику. В остальном функция распределения погрешностей e_k произвольна.*

Поскольку не предполагается, что эта функция входит в то или иное параметрическое семейство, то рассматриваемая модель является непараметрической.

В дальнейшем неоднократно будем использовать центральную предельную теорему (ЦПТ) теории вероятностей (для разнораспределенных слагаемых) для величин e_k, $k = 1, 2, ..., n$ (с весами), поэтому для выполнения ее условий необходимо предположить, например, что погрешности e_k, $k = 1, 2, ..., n$ финитны или имеют конечный третий абсолютный момент. Каждое конкретное слагаемое (с учетом веса) должно быть бесконечно малым относительно всей суммы. Однако заострять здесь внимание на этих внутриматематических «условиях регулярности» нет необходимости [77].

Асимптотические распределения оценок параметров. Из формулы (5.3) следует, что

$$d^*=\frac{1}{n}\sum_{k=1}^{n}x_k=\frac{a}{n}\sum_{k=1}^{n}(t_k-\bar{t})+d+\frac{1}{n}\sum_{k=1}^{n}e_k=d+\frac{1}{n}\sum_{k=1}^{n}e_k. \qquad (5.5)$$

Согласно ЦПТ оценка d^* имеет асимптотически нормальное распределение с математическим ожиданием d и дисперсией σ^2/n, оценка которой приводится ниже. С точки зрения математической статистики вектор оценок (a^*, d^*) обладает более простыми свойствами и легче поддается изучению, чем вектор оценок (a^*, b^*), что и является причиной введения в рассмотрение модели вида $x_k=a(t_k-\bar{t})+d+e_k$.

Из формул (5.3) и (5.5) вытекает, что

$$x_k-\bar{x}=a(t_k-\bar{t})+d+e_k-d-\frac{1}{n}\sum_{k=1}^{n}e_k=a(t_k-\bar{t})+e_k-\frac{1}{n}\sum_{k=1}^{n}e_k;$$

$$(x_k-\bar{x})(t_k-\bar{t})=a(t_k-\bar{t})^2+e_k(t_k-\bar{t})-\frac{(t_k-\bar{t})}{n}\sum_{k=1}^{n}e_k.$$

Последнее слагаемое во втором соотношении при суммировании по k обращается в 0, поэтому из приведенных формул для a^* следует, что

$$a^*=a+\sum_{k=1}^{n}c_ke_k,\quad c_k=\frac{(t_k-\bar{t})}{\sum_{k=1}^{n}(t_k-\bar{t})^2}. \qquad (5.6)$$

Формула (5.6) показывает, что оценка a^* является асимптотически нормальной с математическим ожиданием a и дисперсией

$$D(a^*)=\sum_{k=1}^{n}c_k^2D(e_k)=\frac{\sigma^2}{\sum_{k=1}^{n}(t_k-\bar{t})^2}.$$

Отметим, что многомерная нормальность имеет место, когда каждое слагаемое в формуле (5.6) мало сравнимо со всей суммой, т.е. когда справедливо предельное соотношение

$$\lim_{n\to\infty}\frac{\max_{1\le k\le n}|t_k-\bar{t}\,|}{\sqrt{\sum_{k=1}^{n}(t_k-\bar{t})^2}}=0.$$

Из формул (5.5) и (5.6) и исходных предположений о погрешностях вытекает также то, что математические ожидания оценок равны оцениваемым параметрам (в терминах математической статистики — оценки параметров являются несмещенными).

Несмещенность и асимптотическая нормальность оценок метода наименьших квадратов позволяют легко указывать для них асимптотические доверительные границы 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. Поскольку

$$x^*(t) = \left(a + \sum_{k=1}^{n} c_k e_k\right)(t - \bar{t}) + d + \frac{1}{n}\sum_{k=1}^{n} e_k = a(t - \bar{t}) + d + \sum_{k=1}^{n}\left(c_k + \frac{1}{n}\right)e_k,$$

то в силу ЦПТ случайная величина $x^*(t)$ имеет асимптотически нормальное распределение. Из формул (5.5) и (5.6) следует, что

$$M\left[x^*(t)\right] = M\left[a^*(t-\bar{t}) + d^*\right] = M(a^*)(t-\bar{t}) + M(d^*) =$$

$$= a(t-\bar{t}) + d = x(t),$$

т.е. рассматриваемая оценка прогностической функции является несмещенной. Поэтому

$$x^*(t) - M\left[x^*(t)\right] = (a^* - a)(t - \bar{t}) + d^* - d.$$

Следовательно, дисперсия оценки имеет вид

$$D(x^*(t)) = D(a^*)(t-\bar{t})^2 + 2M\{(a^*-a)(d^*-d)(t-\bar{t})\} + D(d^*).$$

При этом поскольку погрешности e_k независимы в совокупности и имеют нулевое математическое ожидание, то $M(e_k e_j) = 0$, $k \neq j$ и

$$M\left[(a^*-a)(d^*-d)(t-\bar{t})\right] = M\left[\left(\sum_{k=1}^{n} c_k e_k\right)\left(\frac{1}{n}\sum_{k=1}^{n} e_k\right)(t-\bar{t})\right] =$$

$$= (t-\bar{t})\left[\frac{1}{n}\sum_{r=1}^{n} c_k\left(\sum_{j=1}^{n} M(e_k e_j)\right)\right] =$$

$$= \frac{1}{n}(t-\bar{t})\sum_{k=1}^{n} c_k M(e_k^2) = \frac{1}{n}(t-\bar{t})\sigma^2\sum_{k=1}^{n} c_k = 0.$$

Таким образом, с учетом найденных ранее выражений для дисперсий параметров получаем, что

$$D\left[x^*(t)\right] = \frac{\sigma^2}{\sum_{k=1}^{n}(t_k - \bar{t})^2}(t-\bar{t})^2 + 0 + \frac{\sigma^2}{n} = \sigma^2\left\{\frac{1}{n} + \frac{(t-\bar{t})^2}{\sum_{k=1}^{n}(t_k-\bar{t})^2}\right\}. \qquad (5.7)$$

Итак, оценка $x^*(t)$ прогностической функции является несмещенной и асимптотически нормальной. Для практического использования ее асимптотического распределения с целью построения доверительных прогностических интервалов 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 естественную характеристику расхождения между исходными и восстановленными значениями:

$$SS = \sum_{k=1}^{n}(x^*(t_k) - x_k)^2 = \sum_{k=1}^{n}\left[(a^*-a)(t_k-\bar{t}) + (d^*-d) - e_k\right]^2.$$

Величина SS называется *остаточной суммой квадратов*.

В соответствии с формулами (5.5) и (5.6)

$$SS = \sum_{k=1}^{n}\left[(t_k-\bar{t})\sum_{j=1}^{n} c_j e_j + \frac{1}{n}\sum_{j=1}^{n} e_j - e_k\right]^2 =$$

$$= \sum_{k=1}^{n}\left\{\sum_{j=1}^{n}\left[c_j(t_k-\bar{t}) + \frac{1}{n}\right]e_j - e_k\right\}^2 = \sum_{k=1}^{n} SS_k,$$

где

$$SS_k = \left\{\sum_{j=1}^{n}\left[c_j(t_k-\bar{t}) + \frac{1}{n}\right]e_j - e_k\right\}^2, \quad k = 1, 2, \dots, n.$$

Найдем математическое ожидание каждого из слагаемых:

$$M(SS_k) = M\left\{\sum_{j=1}^{n}\left[c_j(t_k-\bar{t}) + \frac{1}{n}\right]e_j - e_k\right\}^2 = M\left\{\sum_{j=1}^{n}\left[c_j(t_k-\bar{t}) + \frac{1}{n}\right]e_j\right\}^2 -$$

$$-2M\left(\left\{\sum_{j=1}^{n}\left[c_j(t_k-\bar{t}) + \frac{1}{n}\right]e_j\right\}e_k\right) + M\left(e_k^2\right) = \sum_{j=1}^{n}\left\{c_j(t_k-\bar{t}) + \frac{1}{n}\right\}^2\sigma^2 -$$

$$-2\left\{c_k(t_k-\bar{t}) + \frac{1}{n}\right\}\sigma^2 + \sigma^2.$$

Из сделанных ранее предположений вытекает, что при $n \to \infty$ имеем $M(SS_i) \to \sigma^2$, $i = 1, 2, ..., n$. Оценив дисперсию случайной величины SS/n, можно показать, что статистика SS/n является состоятельной оценкой дисперсии σ^2.

В столбце (9) табл. 5.1 приведены квадраты значений из столбца (8). Их сумма — это остаточная сумма квадратов SS = 13,64. В соответствии со сказанным ранее рассчитанными по исходным данным табл. 5.1 значениями состоятельных (в смысле математической статистики) оценок дисперсии погрешностей и их среднего квадратичного отклонения являются

$$(\sigma^2)^* = \frac{SS}{n} = \frac{13,64}{6} = 2,27; \quad \sigma^* = \sqrt{\frac{SS}{n}} = \sqrt{\frac{13,64}{6}} = 1,51.$$

Доверительные границы для прогностической функции. Получение состоятельной оценки дисперсии погрешностей дает возможность завершить последовательность задач, связанных с рассматриваемым простейшим вариантом метода наименьших квадратов. Исходим из установленной ранее асимптотической нормальности точечного прогноза $x^*(t)$. Не представляет труда выписывание верхней $x_в(t)$ и нижней $x_н(t)$ границ для прогностической функции:

$$x_в(t) = x^*(t) + \delta(t), \;\; x_н(t) = x^*(t) - \delta(t),$$

где погрешность прогноза $\delta(t)$ имеет вид

$$\delta(t) = U(p)\sqrt{D[x^*(t)]}.$$

Оценив дисперсию $x^*(t)$ с помощью формулы (5.7), в которую вместо неизвестной исследователю дисперсии погрешностей σ^2 подставлена ее оценка, получаем окончательный вид формулы для расчета полуширины доверительного интервала:

$$\delta(t) = U(p)\sigma^* \sqrt{\frac{1}{n} + \frac{(t-\bar{t})^2}{\sum\limits_{k=1}^{n}\left(t_k - \bar{t}\right)^2}}, \;\; \sigma^* = \sqrt{\frac{SS}{n}},$$

где p — доверительная вероятность, $U(p)$ — квантиль нормального распределения порядка $(1 + p)/2$, т.е.

$$\Phi[U(p)] = \frac{1+p}{2}, \quad U(p) = \Phi^{-1}\left(\frac{1+p}{2}\right).$$

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, [7]).

Пример 5.2. Продолжим анализ данных табл. 5.1, исходя из описанной ранее вероятностно-статистической модели.

Рассмотрим распределения оценок параметров. Оценка d^* имеет асимптотически нормальное распределение с математическим ожиданием d и дисперсией, которая оценивается как $(\sigma^2)^*/n = 2,27/6 = 0,38$ (здесь считаем, что 6 — «достаточно большое» число, что, конечно, можно обосновать, изучив точность

приближения методом статистических испытаний или иным). Оценкой среднего квадратичного отклонения является 0,615. Следовательно, при доверительной вероятности 0,95 доверительный интервал для параметра d имеет вид $(26,83 - 1,96 \times 0,615; 26,83 + 1,96 \times 0,615) = (25,625; 28,035)$.

В формулах для дисперсий участвует величина, которую можно представить двумя способами:

$$\sum_{i=1}^{n}(t_i - t_{ср})^2 = \sum_{i=1}^{n}(t_i^2 - 2t_i t_{ср} + t_{ср}^2) =$$

$$= \sum_{i=1}^{n} t_i^2 - 2t_{ср}\sum_{i=1}^{n} t_i + nt_{ср}^2 = \sum_{i=1}^{n} t_i^2 - nt_{ср}^2.$$

Подставив численные значения (табл. 5.1), получаем, что

$$\sum_{i=1}^{n} t_i^2 - nt_{ср}^2 = 256 - 6(5,67)^2 = 63,1.$$

Дисперсия оценки a^* коэффициента при линейном члене прогностической функции оценивается как $2,27/63,1 = 0,036$, а среднее квадратичное отклонение — как 0,19. Следовательно, при доверительной вероятности 0,95 доверительный интервал для параметра a имеет вид $(3,14 - 1,96 \times 0,19; 3,14 + 1,96 \times 0,19) = (2,77; 3,51)$.

Прогностическая функция с учетом погрешности имеет вид (при доверительной вероятности 0,95)

$$x^*(t) = 3,14t + 9,03 \pm 1,96 \times 1,51\sqrt{\frac{1}{6} + \frac{(t-5,67)^2}{63,1}}.$$

В этой записи сохранено происхождение различных составляющих. Упростим:

$$x^*(t) = 3,14t + 9,03 \pm 2,96\sqrt{\frac{1}{6} + \frac{(t-5,67)^2}{63,1}}. \tag{5.8}$$

Например, при $t = 12$ эта формула дает

$$x^*(12) = 46,71 \pm 2,65.$$

Следовательно, нижняя доверительная граница — это 44,06, а верхняя доверительная граница — это 49,36.

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 К. 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 окрестностях Солнца сильно меняется масса кометы. В социально-экономической области горизонты надежного прогнозирования еще менее определены. В частности, они сильно зависят от решений государственных органов.

Чтобы выявить роль погрешностей в прогностической формуле, рассмотрим формальный предельный переход $t \to \infty$. Тогда входящие в формулу (5.8) величины 9,03; $^1/_6$; 5,67 становятся бесконечно малыми и

$$x^*(t) \approx 3{,}14t \pm \frac{2{,}96}{\sqrt{63{,}1}} t = (3{,}14 \pm 0{,}37)t.$$

Таким образом, погрешности составляют около

$$\frac{100 \times 0{,}37}{3{,}14} = 11{,}8\%$$

от тренда (математического ожидания) прогностической функции. В социально-экономических исследованиях подобные погрешности считаются вполне приемлемыми.

Краткое 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 при оценивании дисперсии погрешностей остаточная сумма квадратов SS делится не на n, а на $(n - 2)$. Ясно, что при росте объема данных различия стираются.

Рассмотренный ранее непараметрический подход не использует нереалистическое предположение о нормальности погрешностей. Платой за это является асимптотический характер результатов. В случае простейшей модели метода наименьших квадратов оба подхода дают практически совпадающие рекомендации (поскольку квантили распределения Стьюдента при росте объема выборки приближаются к квантилям нормального распределения). Это не всегда так, не всегда два подхода дают близкие результаты. Например, известно, что в задаче

обнаружения выбросов методы, опирающиеся на нормальное распределение, нельзя считать обоснованными, и это их неприятное свойство было обнаружено с помощью непараметрического подхода [77, подраздел 7.2], [89, подраздел 4.2].

Об общих принципах организационно-экономического моделирования. Ход проведенного исследования позволяет кратко сформулировать несколько общих принципов построения, описания и использования организационно-экономических методов, в частности, в области 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 было сделано ранее. 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.

5.2. ОСНОВЫ ЛИНЕЙНОГО РЕГРЕССИОННОГО АНАЛИЗА

Метод наименьших квадратов, рассмотренный ранее 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), но восстанавливать надо не линейную зависимость, а квадратичную:

$$x(t) = at^2 + bt + c.$$

В этом случае следует рассмотреть функцию трех переменных

$$f(a,b,c) = \sum_{k=1}^{n} \left(x_k - at_k^2 - bt_k - c\right)^2.$$

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:

$$\frac{\partial f(a,b,c)}{\partial a} = \sum_{k=1}^{n} \frac{\partial}{\partial a}\left(x_k - at_k^2 - bt_k - c\right)^2 =$$

$$= \sum_{k=1}^{n} 2\left(-t_k^2\right)\left(x_k - at_k^2 - bt_k - c\right).$$

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 ранее (доверительные границы для параметров и прогностической функции и др.), также могут быть решены. Соответствующие алгоритмы более громоздки. Для их записи полезен аппарат матричной алгебры [111]. Для реальных расчетов используют соответствующие компьютерные программы.

Линейный регрессионный анализ. 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:

$$f\left(a_0, a_1, a_2, a_3, \ldots, a_m\right) = \sum_{k=1}^{n}\left(x_k - a_0 - a_1 t_k - a_2 t_k^2 - a_3 t_k^3 - \ldots - a_m t_k^m\right)^2.$$

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: в среднем цены быстрее всего растут зимой, в декабре—январе, а медленнее всего (иногда в среднем даже падают) — летом, в июле—августе (подробнее проблемы измерения инфляции рассмотрены в главе 6). 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:

$$f\left(a_0, a_1, a_2, a_3, \ldots, a_m, A, B\right) = \sum_{k=1}^{n}\begin{pmatrix} x_k - a_0 - a_1 t_k - a_2 t_k^2 - a_3 t_k^3 - \\ - \ldots - a_m t_k^m - A\sin Bt_k \end{pmatrix}^2.$$

Если время измеряется в годах, то в качестве периодической составляющей (с периодом год) естественно взять $A\cos(2\pi t)$.

Преобразования переменных. 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:

$$\ln I(t) = \ln A + Bt,$$

и ввести новую переменную $x = \ln I(t)$, то получим линейную зависимость x от t, процедура (алгоритм) оценивания параметров которой рассмотрена ранее.

В работе [89] был описан метод оценивания функции спроса по экспериментальным данным. Естественно восстановить зависимость по наборам пар чисел p_i, $D_i = D(p_i)$, $i = 1, 2, \ldots, n$. На центральной части диапазона изменения цены вполне естественно использовать линейную зависимость (поскольку, как отметили еще Ньютон и Лейбниц, любую гладкую функцию на небольшом интервале можно с достаточной точностью приблизить линейной зависимостью). Однако по краям диапазона (т.е. при малых и при больших ценах) линейную зависимость применять нельзя, поскольку ее график выходит за пределы первого квадранта. Естественно использовать степенную зависимость

$$D = D(p) = Cp^{-\alpha},$$

где $C > 0$; $\alpha > 0$.

Эта зависимость нелинейна, метод наименьших квадратов непосредственно применить нельзя. Но можно преобразовать переменные с целью приведения зависимости к линейной. Логарифмируя, получим:

$$\ln D = \ln C - \alpha \ln p.$$

Следовательно, если ввести переменные $x = \ln D = \ln C - \alpha \ln p$, то между этими переменными имеется линейная зависимость, параметры которой оценивать умеем.

Итак, во многих случаях путем преобразования переменных удается перейти от нелинейных зависимостей к линейным в целях обеспечения корректного применения метода наименьших квадратов. Например, если

$$y = \sqrt{at + b},$$

то такой переход осуществляется путем введения переменной $x = y^2$. А если

$$y = \frac{1}{at + b},$$

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 $z = a + bx$. Для каждого конкретного случая подбирают свое преобразование.

Многомерная регрессия. Независимых переменных может быть несколько. Пусть, например, по исходным данным (x_k, y_k, z_k), $k = 1, 2, ..., n$, требуется оценить неизвестные параметры a и b в зависимости

$$z = ax + by + \varepsilon,$$

где ε — погрешность.

Это можно сделать, минимизируя функцию:

$$f(a, b) = \sum_{k=1}^{n} \left(z_k - ax_k - by_k\right)^2.$$

Зависимость от x и y не обязательно должна быть линейной. Предположим, что из каких-то соображений известно, что зависимость должна иметь вид

$$z = ax + by + cx^2y + dxy + ey^3 + \varepsilon,$$

тогда для оценки пяти неизвестных параметров a, b, c, d, e необходимо минимизировать функцию:

$$f(a, b, c, d, e) = \sum_{k=1}^{n} \left(z_k - ax_k - by_k - cx_k^2 y_k - dx_k y_k - ey_k^3\right)^2.$$

Более подробно рассмотрим важный пример из микроэкономики. В одной из оптимизационных моделей поведения фирмы используется так называемая «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 всех предприятий отрасли. Поэтому целесообразно собрать информацию в виде трехмерных векторов $(f_k, K_{k, k})$, $k = 1, 2, ..., n$, где f_k — объем выпуска на k-м предприятии, K_k — объем затрат капитала на k-м 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:*

$$x_k = \ln K_k,\ y_k = \ln L_k,\ z_k = \ln f_k,\ k = 1, 2, ..., n,$$

а затем находить оценки параметров α и β, минимизируя функцию:

$$g(\alpha, \beta) = \sum_{k=1}^{n} \left(z_k - \alpha x_k - \beta y_k\right)^2.$$

Найдем частные производные:

$$\frac{\partial g(\alpha, \beta)}{\partial \alpha} = \sum_{k=1}^{n} 2\left(z_k - \alpha x_k - \beta y_k\right)\left(-x_k\right);$$

$$\frac{\partial g(\alpha, \beta)}{\partial \beta} = \sum_{k=1}^{n} 2\left(z_k - \alpha x_k - \beta y_k\right)\left(-y_k\right).$$

Приравняем частные производные к 0, сократим на 2, раскроем скобки, перенесем свободные члены вправо. Получим систему двух линейных уравнений с двумя неизвестными:

$$\alpha\sum_{k=1}^{n} x_k^2 + \beta\sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} x_k z_k;$$

$$\alpha\sum_{k=1}^{n} x_k y_k + \beta\sum_{k=1}^{n} y_k^2 = \sum_{k=1}^{n} y_k z_k.$$

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 табл. 5.1, что применялась ранее. Отметим, что рассмотренная в предыдущем подразделе постановка переходит в разбираемую сейчас при $y_k = 1$, $k = 1, 2, ..., n$.

Решая систему линейных уравнений, получаем оценки метода наименьших квадратов α^* и β^*.

Сравним исходные f_k с восстановленными:

$$\hat{f}_k = f^*(K_k, L_k) = K_k^{\alpha^*} L_k^{\beta^*}.$$

Если $f_k > \hat{f}_k$, то это означает, что предприятие работает лучше, чем в среднем по отрасли (из-за различной исходной фондовооруженности нельзя сравнивать предприятия непосредственно, необходимо сначала восстановить зависимость выпуска от факторов производства с помощью модели Кобба — Дугласа). Это заслуга руководства предприятия. Следовательно, вышестоящие структуры имеют основания продвигать и награждать топ-менеджеров этого предприятия. Деловые партнеры имеют основания для налаживания долговременных отношений. Банки имеют основания давать льготные кредиты.

Если $f_k < \hat{f}_k$, то ситуация прямо противоположная. Предприятие работает хуже, чем в среднем по отрасли, т.е. хуже, чем можно было бы ожидать при имеющихся у него ресурсах. Это вина руководства предприятия. Следовательно, вышестоящим структурам целесообразно наказать директора и его заместителей, например «укрепить руководство предприятия», заменив часть или всех топ-менеджеров. Деловым партнерам надо учесть, что выполнение договорных отношений находится под угрозой срыва. Из-за повышенного риска банки повысят процент платы за кредит, ужесточат требования к обеспечению возврата кредитов, в частности беря имущество предприятия в залог.

Случай $f_k = \hat{f}_k$ — промежуточный между двумя описанными. Предприятие работает на уровне, среднем для отрасли. Нет оснований ни хвалить его руководство, ни применять санкции.

Как можно использовать восстановленную зависимость? Например, проектируем новое предприятие. Определяем для него величины факторов производства K_0 и L_0. Тогда можем рассчитывать на выпуск продукции в объеме $f^*(K_0, L_0)$.

Пример 5.3. Руководитель маркетинговой службы завода ГАРО (г. Великий Новгород) А.А. Пивень применил метод линейного регрессивного анализа для построения математической модели рынка легковых подъемников. Требовалось 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 (X_1);
- цена (X_2);
- наличие напольной рамы (X_3);
- наличие синхронизации (X_4);
- число двигателей (X_5);
- суммарная мощность двигателей (X_6);
- высота подхвата в нижнем положении (X_7);
- максимальная высота подъема (X_8);
- скорость подъема (X_9);
- гарантийный срок (X_{10});
- срок службы (X_{11});
- время на рынке (X_{12});
- внешний вид (X_{13});
- срок поставки (X_{14});
- уровень сервисного обслуживания (X_{15});
- наличие системы смазки (X_{16});
- масса (X_{17}).

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 от 0 в сравнении с их дисперсией. Для этого использовался пакет STATISTICA 6.0, конкретно модуль «Множественная регрессия» (*Multiple regression*).

В результате расчетов получена зависимость объема продаж подъемника ПЗ-Т от 12 факторов:

$$V = -1769{,}77 - 65{,}09X_1 - 0{,}03X_2 + 68{,}79X_3 + 147{,}54X_4 + 156{,}28X_5 + 2{,}53X_7 + 1{,}06X_8 + 25{,}75X_{12} - 132{,}26X_{13} - 12{,}41X_{14} + 107{,}78X_{15} + 397X_{16}.$$

Влияние остальных пяти факторов оказалось незначимым.

Исходя из расчетов, прогнозное значение продаж подъемников на второй год продаж составило ориентировочно 1010 шт. С вероятностью 95% можно утверждать, что объем продаж будет лежать в границах 695—1332 шт.

5.3. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

Прежде чем восстанавливать зависимость, целесообразно убедиться, что переменные связаны между собой. Термин «корреляция» и 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, т.е. выборочным линейным парным коэффициентом корреляции К. Пирсона, как известно, называется число

$$r_n = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}.$$

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:

$$r_n \to \rho = \frac{M\left[x_1 - M(x_1)\right]\left[y_1 - M(y_1)\right]}{\sqrt{D(x_1)}\sqrt{D(y_1)}}$$

(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; $D_0(r_n)$ — асимптотическая дисперсия выборочного коэффициента корреляции.

Она имеет довольно сложное выражение, приведенное в монографии [35, 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:

$$\mu_{km} = M\left[x_1 - M(x_1)\right]^k\left[y_1 - M(y_1)\right]^m.$$

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 [77; 89]. 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 $x_1, x_2, ..., x_n$. Для каждого y_i рассчитать его ранг q_i в вариационном ряду, построенном по выборке $y_1, y_2, ..., y_n$. Для набора из n пар (r_i, q_i), $i = 1, 2, ..., n$, вычислить линейный коэффициент корреляции. Он называется коэффициентом ранговой корреляции, поскольку определяется через ранги. В качестве примера рассмотрим данные из табл. 5.2.

Таблица 5.2

Данные для расчета коэффициентов корреляции

Переменная	Число пар				
	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. 5.2 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 9), как и другие ранговые статистики, например статистики Вилкоксона, Смирнова, типа омега-квадрат для проверки однородности независимых выборок [77].

Широко используется также коэффициент ранговой корреляции τ Кендалла[1], коэффициент ранговой конкордации Кендалла и Б. Смита

[1] Известный английский статистик M.G. Kendall известен в нашей стране как Кендалл (в книгах, выпущенных издательствами «Наука» и «Мир») и Кендэл (в книгах

(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 [28], необходимые для практических расчетов таблицы имеются в справочнике [7]. Дискуссия о наиболее адекватном выборе вида коэффициентов корреляции продолжается и в настоящее время.

5.4. ПРОГНОЗИРОВАНИЕ В ОТРАСЛИ ЛОМА ЧЕРНЫХ МЕТАЛЛОВ[1]

Среди статистических методов прогнозирования [84] основной — метод наименьших квадратов. Продемонстрируем его практическую пользу на примере исследования [36], посвященного решению задач прогнозирования цены лома черных металлов — сырья для Магнитогорского металлургического комбината.

Как показано в [36], в настоящее время для мировой и российской металлургии одна из наиболее актуальных проблем — повышение эффективности использования ресурсов лома черных металлов. Каждая промышленно развитая страна стремится максимально увеличить долю использования металлолома в сталеплавильном производстве, решая таким образом не только вопросы экологии, но и проблему дефицита рудного сырья, коксующихся углей, которые постоянно дорожают.

В России производство стали ежегодно возрастает. За 2001—2006 гг. суммарный прирост выплавки, согласно [36], составил свыше 28% (рис. 5.1). По мере увеличения выплавки стали возрастает масса необходимой для ее производства металлической шихты. Одним из главных составляющих шихты выступает металлолом. Его потребление определяется структурой сталеплавильного производства, развитием новых технологий выплавки, разливки и прокатки стали.

На смену мартеновскому производству стали в мировой металлургии пришли кислородно-конвертерный и электросталеплавильный процессы. Если первый имеет ограничение по применению металлолома, то второй базируется преимущественно на потреблении стального лома [36].

Современные конвертерный и электросталеплавильный процессы ориентированы на 100%-ную непрерывную разливку стали, что существенно сокращает образование оборотного лома в процессе производства. В настоящее время потребность черной металлургии в ресурсах лома практически в 2 раза превышает объем внутризаводского оборота

издательства «Финансы и статистика»). Мы придерживаемся первого написания.

[1] Подраздел 5.4 составлен по материалам статьи [36] ведущего менеджера отдела маркетинга ЗАО «Профит» Е.М. Крюковой, подготовленной в 2007 г.

металлолома. Все это ведет к значительному росту потребности металлургической промышленности в ломе черных металлов, поставляемом специализированными ломоперерабатывающими предприятиями, усилению конкуренции за лом на российском рынке.

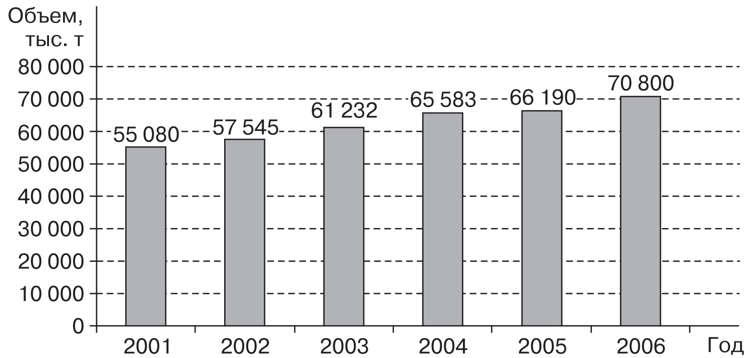


Рис. 5.1. Выплавка стали в России

Увеличение спроса, естественно, сопровождается ростом цен на лом черных металлов. За 2002—2006 гг. средневзвешенные цены на металлолом на российском рынке увеличились более чем в 2 раза (табл. 5.3). Это увеличивает привлекательность ломоперерабатывающего бизнеса.

Таблица 5.3

Изменение среднегодовой цены на внутреннем рынке в 2002—2006 гг.

Период	Среднегодовая цена с НДС и ЖДТ*, руб./т	Изменение по отношению к предыдущему периоду		Индекс инфляции**, ед.	Среднегодовая цена с НДС и ЖДТ, руб./т (сопоставимые цены на 2002 г. — без учета влияния инфляции)	Изменение, %
		в руб./т	в %			
2002	1 812	—	—	—	1 812	—
2003	3 039	1 226	167,7	1,12	2 713	149,7
2004	4 169	1 130	137,2	1,117	3 332	122,8
2005	4 349	181	104,3	1,109	3 135	94,1
2006	5 559	1 210	127,8	1,09	3 676	117,3

* Средневзвешенная цена на лом черных металлов (по крупнейшим металлургическим комбинатам России), НДС — налог на добавленную стоимость, ЖДТ — железнодорожный тариф.

** По данным Росстата.

Структуру рынка лома можно представить в виде взаимосвязи трех важнейших участников рынка: ломосдатчиков, ломопереработчиков и потребителей лома (рис. 5.2). *Ломосдатчики* — это юридические и физические лица, которые либо списывают машины, оборудование, либо находят их «на земле», проводят демонтаж и отгружают ломопереработчикам. *Ломопереработчики* в свою очередь подразделяются на предприятия, имеющие собственные ломозаготовительные площадки и осуществляющие на них сбор и переработку лома, и на транзитные организации, которые не имеют собственных площадок, а просто осуществляют транспортировку и реализацию металлолома. Замыкают эту цепочку *потребители лома* — металлургические комбинаты и заводы, которые осуществляют приемку, складирование лома в копровых цехах и подготовку для сталеплавильного процесса.

Рис. 5.2. Структура участников рынка, каналы сбыта

Система ценообразования на рынке лома выглядит следующим образом. Потребители лома устанавливают цены на материал, исходя из потребности производства и наличия остатков на складах. В свою очередь ломоперерабатывающие предприятия устанавливают цены «на земле», т.е. цены закупки лома у ломосдатчиков. Задача ломопереработчика — устанавливать в разные периоды такие цены, которые обеспечат им максимальную рентабельность. Таким образом, эффективность деятельности ломопереработчика зависит от оценки рыночной конъюнктуры, характеристики влияния рыночных факторов и прогноза дальнейшей динамики закупочных цен сталепроизводителей на лом черных металлов.

Динамика цен на лом черных металлов имеет свои особенности.

Во-первых, это изменение носит сезонный характер. Рост цен ежегодно начинается в марте—апреле, это связано с «подъеданием» зимних запасов на складах металлургических предприятий и ростом потребности сталепроизводителей в привозном металлоломе. К маю, в связи с окончанием зимы, увеличивается заготовка лома и подход его на склады, что ведет к снижению закупочных цен на лом. В конце лета вновь начинается рост: потребность предприятий в ломе возрастает в связи с началом накопления зимних запасов материала на складах. После достижения максимума цен в сентябре—октябре следует понижение. Данная тенденция с небольшими сдвигами повторяется из года в год.

Во-вторых, в последние годы очевидна тенденция сглаживания цен в течение года, цена изменяется равномерно без резких скачков и «провалов» (рис. 5.3).

Наконец, множество факторов определяют цену на лом:

- потребность металлургического предприятий в ломе черных металлов (месячные планы поставки) и остатки лома на складах копровых цехов комбинатов. Чем выше потребность сталепроизводителя в ломе на текущий месяц, тем большую цену он готов заплатить за сырье. С помощью цен потребители регулируют потоки материала в сторону комбината: невыполнение запланированных показателей ведет к увеличению закупочных цен на металлолом; перевыполнение, напротив, как правило, приводит к понижению цен. Последнее связано, в первую очередь, с ограниченными производственными мощностями копровых цехов комбинатов (перегрузочная техника, железнодорожные тупики), которые не могут обеспечить своевременную выгрузку и складирование всего поступающего объема металлолома;

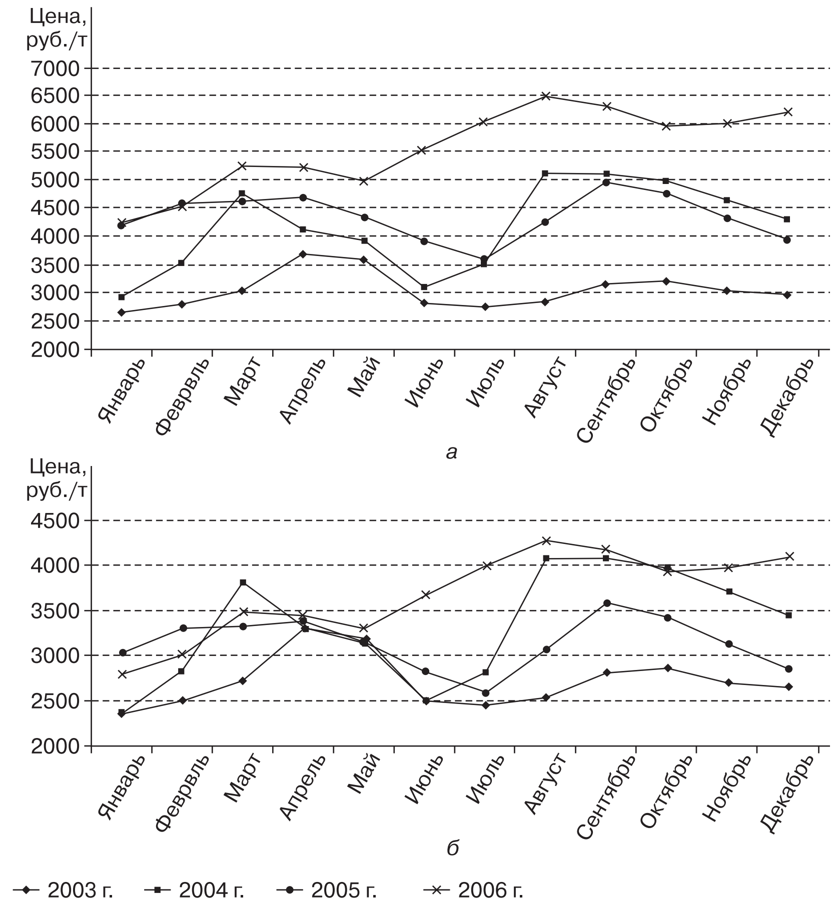


Рис. 5.3. Динамика средних цен на лом марки 3А без учета (*а*) и с учетом (*б*) инфляции

- суммарный объем потребления внутреннего рынка. Как уже было сказано ранее, потребление лома черных металлов носит сезонный характер: потребность сталепроизводителей существенно возрастает в периоды формирования остатков лома на складах. Это ведет к росту конкуренции за сырье на рынке и, как следствие, к росту среднерыночных цен;
- объемы поставки российского лома на экспорт и цены на мировых рынках. Значительные объемы металлолома по-прежнему отвлекаются на экспорт. Крупнейшим импортером российского

металлолома является Турция. Поэтому когда турецкие заводы дают высокие цены на металлолом, возрастают закупочные цены экспортеров в российских портах, что неизбежно ведет к ответному росту цен производителями внутреннего рынка;

- цены на готовую продукцию металлургических комбинатов так как значительные объемы металлопродукции российских сталезаводов направляются на экспорт; в качестве индикаторов в [36] выбраны цены на арматуру в портах Черного и Балтийского морей. Динамика цен на лом черных металлов в течение года повторяет динамику цен на арматуру (рис. 5.4).

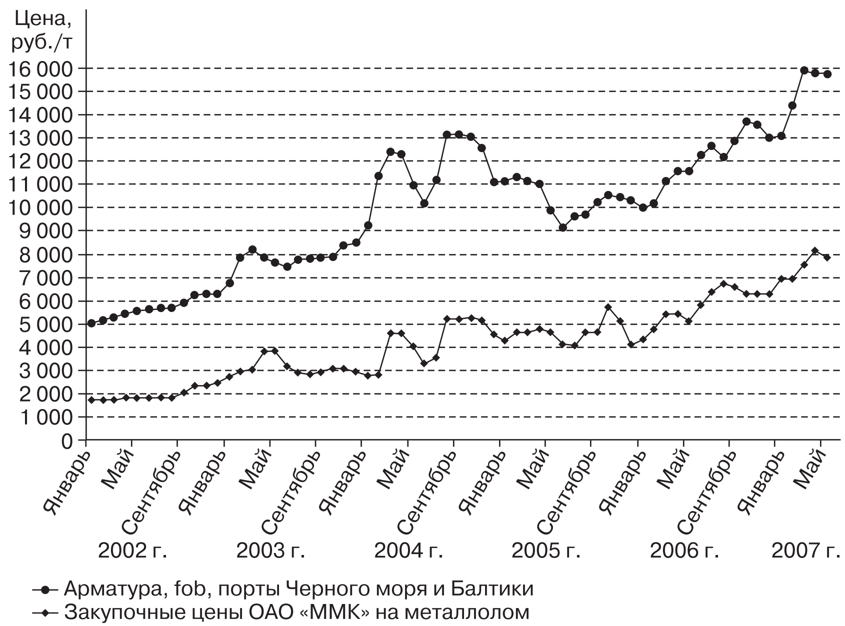


Рис. 5.4. Динамика цен на арматуру и металлолом

От оценки рыночной конъюнктуры, характеристики влияния рыночных факторов и прогноза дальнейшей динамики цен на лом черных металлов зависит эффективность деятельности ломопереработчика.

Рассмотрим эконометрическую модель прогноза закупочных цен на лом на примере ОАО «Магнитогорский металлургический комбинат» (ОАО «ММК»). Адекватность модели во многом зависит от выбора факторов, их точности и достоверности. В работе [36] отобраны следующие факторы для построения модели:

- поставка на внутренний рынок, тыс. т (X_1);

- поставка на экспорт, тыс. т (X_2);
- цены на лом, Турция, руб./т [36] (X_3);
- поставка на ОАО «ММК», тыс. т (X_4);
- остаток лома на ОАО «ММК», тыс. т (X_5);
- цены на арматуру, порты Черного моря и Балтики, руб./т [36] (X_6).

Исходя из понимания того, что реальные данные всегда имеют распределение, отличное от нормального, для восстановления зависимости целесообразно применять непараметрический подход, включая доверительное оценивание параметров вероятностной модели и прогностической функции [89].

Предполагается, что переменные связаны линейной регрессионной зависимостью вида:

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_nx_n,$$

где Y — зависимая переменная (цена на лом марки 3А); a — свободный член; x_i — независимые переменные (факторы); b_i — регрессионные коэффициенты; $i = 1, 2, \ldots, n$.

Исходные данные за 2003—2005 гг. приведены в табл. 5.4.

Таблица 5.4

Исходные данные для построения модели расчета цен ОАО ММК на лом марки 3А

Период	Поставка на внутренний рынок, тыс. т*	Поставка на экспорт, тыс. т*	Цены на лом, Турция, руб./т*	Поставка на ОАО «ММК», тыс. т**	Остаток лома на ОАО «ММК», тыс. т**	Цены на арматуру, порты Черного и Балтийского морей, руб./т*	Закупочные цены ОАО «ММК» на лом, руб./т**
	X_1	X_2	X_3	X_4	X_5	X_6	Y
			2003 г.				
Январь	662	256	4 464	65 810	77 849	6 745	2 750
Февраль	715	212	5 092	85 046	67 695	7 830	2 950
Март	842	416	5 033	104 212	33 533	8 178	3 050
Апрель	1 209	482	4 760	166 203	34 091	7 865	3 800
Май	1 492	542	4 267	247 723	109 589	7 637	3 800
Июнь	1 298	628	3 860	247 555	217 976	7 468	3 150
Июль	1 057	667	4 250	168 568	264 461	7 742	2 900
Август	1 126	725	4 460	125 541	281 170	7 798	2 825
Сентябрь	1 144	702	4 884	116 449	282 234	7 863	2 950
Октябрь	1 244	802	4 826	119 978	282 971	7 903	3 100
Ноябрь	1 123	717	4 923	123 509	296 580	8 347	3 100

Окончание

Декабрь	1 034	924	5 521	86 707	282 933	8 481	2 950
Период	Поставка на внутренний рынок, тыс. т*	Поставка на экспорт, тыс. т*	Цены на лом, Турция, руб./т*	Поставка на ОАО «ММК», тыс. т**	Остаток лома на ОАО «ММК», тыс. т**	Цены на арматуру, порты Черного и Балтийского морей, руб./т*	Закупочные цены ОАО «ММК» на лом, руб./т**
	X_1	X_2	X_3	X_4	X_5	X_6	Y
2004 г.							
Январь	719	869	6 547	40 027	224 376	9 234	2 761
Февраль	789	847	7 520	27 144	143 178	11 348	2 824
Март	1 210	969	7 562	207 303	97 722	12 413	4 602
Апрель	1 199	1 141	6 454	78 375	112 797	12 335	4 602
Май	1 554	1 256	4 584	137 042	197 946	10 987	4 050
Июнь	1 179	1 231	5 135	216 765	244 032	10 189	3 298
Июль	1 230	1 189	6 732	188 504	232 936	11 196	3 540
Август	1 447	1 340	6 975	120 708	222 165	13 146	5 201
Сентябрь	1 359	1 400	7 086	134 830	319 025	13 150	5 201
Октябрь	1 649	1 318	7 795	261 497	374 111	13 030	5 268
Ноябрь	1 423	1 394	7 579	191 016	398 175	12 584	5 141
Декабрь	1 247	1 405	6 557	152 518	409 001	11 106	4 543
2005 г.							
Январь	733	926	7 142	58 251	319 696	11 148	4 307
Февраль	957	816	6 771	56 763	219 397	11 338	4 661
Март	1 132	1 032	7 328	75 006	149 786	11 133	4 661
Апрель	1 542	1 429	6 578	123 000	100 000	11 014	4 779
Май	1 740	1 437	5 501	294 793	242 233	9 891	4 661
Июнь	1 599	1 252	4 952	295 934	397 205	9 148	4 130
Июль	1 307	968	6 026	161 804	353 555	9 642	4 071
Август	1 427	1 189	6 834	137 785	357 436	9 710	4 660
Сентябрь	1 942	1 466	7 095	207 267	404 595	10 217	4 660
Октябрь	2 005	1 214	6 055	272 533	529 378	10 539	5 723
Ноябрь	1 532	980	6 126	155 956	513 734	10 470	5 133
Декабрь	1 301	1 210	5 905	141 797	496 039	10 341	4 130

* По данным аналитического агентства «Металл-Курьер».

** Фактические данные ОАО «ММК».

С помощью модуля «Множественная регрессия» пакета STATISTICA были определены регрессионные коэффициенты (табл. 5.5) [116] и коэффициенты линейной корреляции Пирсона (табл. 5.5, 5.6).

Таблица 5.5

Характеристика связи цены и факторов

	Независимые переменные	Регрессионные коэффициенты	Коэффициенты линейной корреляции
X_1	Поставка на внутренний рынок, тыс. т	2,009	0,772
X_2	Поставка на экспорт, тыс. т	–1,076	0,806
X_3	Цены на лом, Турция, руб./т	0,136	0,768
X_4	Поставка на ОАО «ММК», тыс. т	–0,925	0,644
X_5	Остаток лома на ОАО «ММК», тыс. т	0,708	0,393
X_6	Цены на арматуру, порты Черного и Балтийского морей, руб./т	0,326	0,837

Наиболее сильные связи можно отметить между ценами на лом и такими факторами, как цены на арматуру в портах, цены на импортируемый лом в Турции и объемы поставки на внутренний рынок России и на экспорт.

Необходимо отметить существенную взаимозависимость между такими факторами, как поставка на внутренний рынок и поставка на ОАО «ММК» (табл. 5.6). Это связано с тем, что периоды роста объемов закупки лома и накопления зимних запасов по разным комбинатам (которые и задают совокупный рост внутреннего рынка) часто совпадают.

Таблица 5.6

Взаимозависимость факторов

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	—	–0,089	0,004	–0,452	–0,035	0,001
X_2	–0,089	—	0,001	0,056	–0,102	–0,021
X_3	0,004	0,001	—	0,038	–0,016	–0,008
X_4	–0,452	0,056	0,038	—	–0,028	–0,005
X_5	–0,035	–0,102	–0,016	–0,028	—	0,018
X_6	0,001	–0,021	–0,008	–0,005	0,018	—

Результаты прогнозирования приведены в табл. 5.7. Коэффициент детерминации для модели составляет R^2 = 0,91 (коэффициент детерминации — это квадрат коэффициента линейной корреляции между зависимой переменной и прогностической функцией). Это означает, что отобранные факторы на 91% объясняют дисперсию зависимой переменной — цены на лом черных металлов.

Таблица 5.7

Прогноз закупочных цен ОАО «ММК» на лом черных металлов на 2006 г.

Период	Поставка на внутренний рынок, тыс. т	Постав-ка на экспорт, тыс. т	Цены на лом, Турция, руб./т	Поставка на ОАО «ММК», тыс. т	Остаток лома на ОАО «ММК», тыс. т	Цены на арматуру, порты Черного и Балтийского морей, руб./т	Прогноз цен ОАО «ММК» на лом, руб./т	Закупочные цены ОАО «ММК» на лом, руб./т	Отклонения в руб./т	Отклонения в %
	Коэффициенты регрессии									
	X_1	X_2	X_3	X_4	X_5	X_6	Y			
	2,01	−1,08	0,14	−0,93	0,71	0,33	−1570,3			
Январь	532	576	5 937	51	354	10 037	3 164	4 336	1 172	27,03
Февраль	875	364	6 569	53	256	10 178	4 142	4 773	631	13,21
Март	1 346	573	6 829	91	174	11 150	5 122	5 430	308	5,68
Апрель	1 652	925	7 308	204	180	11 583	5 464	5 430	−34	−0,63
Май	1 776	1 199	7 036	300	292	11 583	5 371	5 133	−238	−4,64
Июнь	1 816	1 298	7 717	295	345	12 277	5 707	5 800	93	1,60
Июль	1 980	1 301	7 536	371	470	12 650	6 150	6 372	222	3,49
Август	2 128	1 588	6 771	410	609	12 177	5 942	6 726	784	11,65
Сен-тябрь	2 171	974	7 248	393	714	12 864	7 066	6 608	−458	−6,93
Октябрь	2 037	757	7 334	443	827	13 700	7 350	6 281	−1 069	−17,01
Ноябрь	1 815	901	7 386	390	881	13 575	6 804	6 281	−523	−8,32
Декабрь	1 692	1 008	7 525	275	814	13 013	6 336	6 281	−55	−0,87

Данная математическая модель достаточно точно показывает динамику цен в течение года, есть основания использовать ее и для долгосрочного прогнозирования (рис. 5.5). Прогнозирование на краткосрочные периоды требует ежемесячного пересчета для уточнения модели.

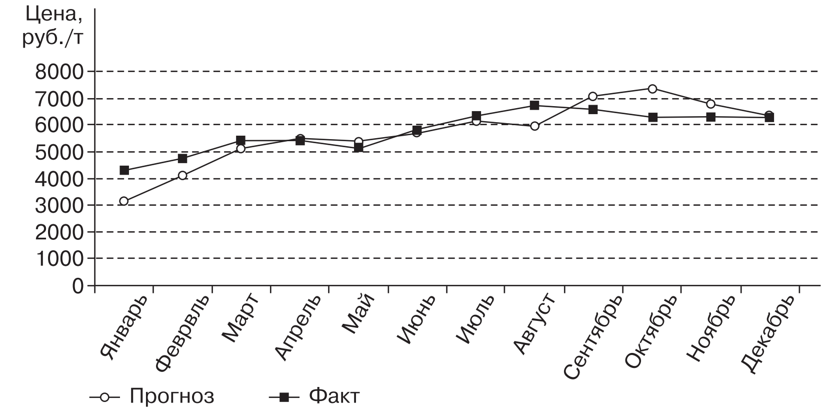


Рис. 5.5. Сопоставление прогнозных и фактических закупочных цен ОАО «ММК» на лом черных металлов в 2006 г. (с учетом НДС и железнодорожных тарифов)

Наиболее высокие ошибки прогноз показывает в январе—феврале (27%), августе (12%) и октябре (17%). Средняя абсолютная ошибка прогноза составляет 466 руб./т, средняя относительная ошибка — 8,4%, среднеквадратичная ошибка прогноза — 593 руб./т.

Ошибки (т.е. отклонения факта от прогноза) объясняются воздействием неучтенных в модели факторов. В частности, отклонение в январе—феврале связано с суровыми зимними условиями в 2006 г., которые привели к осложнению и практическому прекращению ломосбора на территории России. В результате первые месяцы года сталепроизводители работали в основном за счет накопленных накануне остатков на складах, поступление лома было минимальным, что привело к существенному росту цен.

По аналогии с прошлыми годами в июле—августе 2006 г. прогноз показывал снижение цен на рынке, а в сентябре—октябре, наоборот, — рост. Однако летом 2006 г. комбинаты повели себя иначе. Низкий уровень ломосбора на территории страны в первой половине года (в связи с затяжной зимой) и недостаточные объемы поступления металлолома потребителям явились причиной того, что к концу первого полугодия на складах комбинатов оставалось не более 40% необходимого уровня складских запасов. В результате уже в июле потребители приступили к формированию зимних запасов на складах, что повлекло за собой рост цен. Ценовой максимум в 2006 г. был достигнут не в октябре, как в предыдущие годы, а в августе, и в сентябре уже началось постепенно снижение уровня цен.

Еще одним фактором такой динамики послужил существенный рост потребления на рынке в связи с вводом новых производственных мощностей. В частности, ОАО «ММК» в 2006 г. запустило две электросталеплавильные печи, в результате потребность в привозном металлоломе выросла с 1,9 млн т в 2005 г. до 3,2 млн т в 2006 г., прирост составил 65,4%. Рост потребности ведет к обострению конкуренции на рынке, и так как цена является основным рычагом привлечения лома, комбинаты начинают вести активную ценовую политику, стараясь регулировать потоки лома.

Кроме того, на рыночную конъюнктуру оказывают влияние политические факторы, реализация различных государственных программ — лицензирование поставщиков металлолома, изменение экспортных пошлин, изменение правил налогообложения ломозаготовительных организаций (например, введение льготного режима по налогу на добавленную стоимость для ломоперерабатывающих организаций). Также значительное влияние на цены оказывает внутренняя политика организации — волевые решения менеджеров, наличие оборотных средств для осуществления расчетов с поставщиками и формирование запасов на складах.

Таким образом, одной из основных сложностей в получении точных прогнозов экономических показателей являются неожиданные и важные сдвиги в ключевых экономических факторах, а также влияние социально-экономических факторов, которые невозможно учесть с помощью математики. Очевидно, что для корректировки математического прогноза, для принятия обоснованных решений в области ценообразования важно также опираться на опыт, знания и интуицию специалистов [36]. Для этих целей используются методы экспертных оценок. В состав экспертной комиссии целесообразно включить, во-первых, высококвалифицированных опытных специалистов, непосредственно участвующих в процессе ценообразования и занимающихся реализацией лома черных металлов (руководителей отделов сбыта, маркетинга и сотрудников данных отделов), а также независимых аналитиков, оказывающих информационные и консультационные услуги на рынке (представителей аналитических агентств). Сбор экспертной информации для целей прогнозирования цен на лом предлагается проводить

ежемесячно. Прогнозные оценки экспертов могут быть получены на месяц вперед и на длительный период — до одного года. Полученная от экспертов информация позволяет скорректировать прогноз, полученный с помощью метода наименьших квадратов, и выработать «грамотную» ценовую стратегию и стратегию поведения на рынке (подходящих моментов для закупки и реализации металлолома).

Таким образом, в ходе анализа реальных данных установлено, что эконометрическая регрессионная модель достаточно точно описывает динамику цен на рынке лома в течение года, поэтому есть основания использовать ее для долгосрочного прогнозирования. С другой стороны, можно отметить существенные отклонения прогнозных данных от реальных в отдельные периоды, что связано, в первую очередь, с действием неучтенных в модели факторов. Учесть их влияние и скорректировать результаты, полученные с помощью метода наименьших квадратов, возможно с помощью метода экспертных оценок. Применение статистических и экспертных методов в совокупности для целей прогнозирования цен на лом черных металлов позволяет ломоперерабатывающему предприятию своевременно реагировать на изменения рыночных факторов и строить с учетом этого свою ценовую политику.

5.5. О ВЫБОРЕ ВИДА РЕГРЕССИОННОЙ МОДЕЛИ

Одни и те же данные можно обрабатывать различными способами. Например, как уже отмечалось, функцию спроса можно приближать линейной зависимостью или степенной. Можно пробовать применить квадратичную зависимость или многочлен третьего порядка и т.д. Какую же организационно-экономическую модель использовать? Очевидно, ту, которая лучше других соответствует реальным данным. Но что значит «лучше соответствует»? Как измерять качество модели?

Основной показатель качества регрессионной модели. 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 (т.е. дисперсии погрешностей e_k), скорректированную на число m параметров, оцениваемых по наблюдаемым данным:

$$\hat{\sigma}^2(m) = \frac{SS}{n-m},$$

где, как и раньше, n — объем данных (число векторов, по которым восстанавливается зависимость). В случае задачи восстановления линейной функции одной переменной, рассмотренной в начале главы, оценка остаточной дисперсии имеет вид

$$\partial^2(m) = \frac{SS}{n-2},$$

поскольку число оцениваемых параметров $m = 2$.

Почему эта формула отличается от приведенной в подразделе 5.1? 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; точнее, отношение этих величин стремится к 1.

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 (существование такого многочлена доказывают в курсе высшей алгебры). 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 степени многочлена (полинома) можно использовать, например, первый локальный минимум скорректированной оценки остаточной дисперсии, т.е.

$$m^* = \min\left[m : v(m-1) > v(m),\ v(m) \leq v(m+1)\right].$$

В работе [77] найдено предельное распределение этой оценки степени многочлена.

Теорема 5.2. При справедливости некоторых внутриматематических условий регулярности

$$\lim_{n\to\infty} P(m^* < m_0) = 0,\ \lim_{n\to\infty} P(m^* = m_0 + u) = \lambda(1-\lambda)^u,\ u = 1, 2, \ldots,$$

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 (одно из известных в теории вероятностей параметрических семейств распределений). Это означает, в частности, что оценка не является состоятельной. При этом вероятность получить меньшее значение, чем истинное, исчезающе мала. Далее имеем:

$$P(m^* = m_0) \to 0{,}68268, \quad P(m^* = m_0 + 1) \to 0{,}68268(1 - 0{,}68268) = 0{,}21663;$$

$$P(m^* = m_0 + 2) \to 0{,}68268(1 - 0{,}68268)^2 = 0{,}068744;$$

$$P(m^* = m_0 + 3) \to 0{,}68268(1 - 0{,}68268)^3 = 0{,}021814.$$

Разработаны и иные методы оценивания неизвестной степени многочлена, например путем многократного применения процедуры проверки адекватности регрессионной зависимости с помощью статистики Фишера [77]. Предельное поведение оценок такое же, как в приведенной теореме, только значение параметра λ иное. Разработаны также состоятельные оценки степени полинома [77].

Регрессия как условное математическое ожидание. Во всех рассмотренных ранее постановках вид регрессионной зависимости задавался исследователем. Это линейная функция, или многочлен, или элемент иного параметрического семейства. Однако откуда уверенность, что семейство выбрано правильно? А если исследователь ошибся? Тогда к случайным ошибкам добавляется методическая. Например, данные приближаются линейной функцией, а на самом деле зависимость квадратичная. Тогда разность между этими функциями (тоже квадратичная функция) и есть методическая ошибка.

Избежать методической ошибки можно, не задавая априори вид регрессионной зависимости. Продемонстрируем это на примере оценивания условного математического ожидания.

Рассмотрим общее понятие регрессии как условного математического ожидания. Пусть случайный вектор $[x(\omega), y(\omega)]$ (здесь ω — элементарный исход опыта) имеет плотность $p(x, y)$. Как известно из курса теории вероятностей, плотность условного распределения $y(\omega)$ при условии $x(\omega) = x_0$ имеет вид

$$p(y \mid x) = p\left[y \mid x(\omega) = x_0\right] = \frac{p(x, y)}{\int_{-\infty}^{+\infty} p(x, y)dy}.$$

Условное математическое ожидание, т.е. регрессионная зависимость y от x, имеет вид

$$f(x) = \int_{-\infty}^{+\infty} yp(y \mid x)dy = \frac{\int_{-\infty}^{+\infty} yp(x, y)dy}{\int_{-\infty}^{+\infty} p(x, y)dy}.$$

Таким образом, для нахождения оценок регрессионной зависимости достаточно найти оценки совместной плотности распределения вероятности $p_n(x, y)$ такие, чтобы

$$p_n(x, y) \to p(x, y) \text{ при } n \to \infty.$$

Тогда непараметрическая оценка регрессионной зависимости, построенная с помощью замены неизвестной исследователю плотности совместного распределения на непараметрическую состоятельную оценку этой плотности, т.е.

$$f_m(x) = \frac{\int_{-\infty}^{+\infty} yp_n(x, y)dy}{\int_{-\infty}^{+\infty} p_n(x, y)dy}$$

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 7 [89].

Регрессионному анализу (т.е. методам восстановления зависимостей) посвящено огромное количество литературы (по нашей оценке, не менее 100 тыс. книг и статей на различных языках). Он хорошо представлен в программных продуктах по анализу данных, особенно та его часть, которая связана с методом наименьших квадратов.

5.6. НЕПАРАМЕТРИЧЕСКОЕ ОЦЕНИВАНИЕ ТОЧКИ ПЕРЕСЕЧЕНИЯ РЕГРЕССИОННЫХ ПРЯМЫХ

Рассмотрим несколько конкретных практически важных задач регрессионного анализа. Они нужны для решения организационно-экономических проблем прогнозирования на промышленном предприятии [58]. В рамках непараметрической вероятностно-статистической модели (т.е. без предположения о нормальности распределения погрешностей) получим асимптотическое распределение точки пересечения (встречи) двух регрессионных линейных зависимостей. Для этого на основе метода линеаризации выпишем выражения для асимптотической дисперсии точки встречи и границ доверительного интервала для нее [57].

Постановка задачи. Пусть зависимость от времени t некоторого показателя $x_1(t)$ технического уровня или качества продукции предприятия «Альфа» описывается линейной функцией

$$x_1(t) = a_1 t + d_1.$$

Пусть аналогичный показатель у его конкурента (предприятие «Бета») также описывается линейной функцией, но с другими коэффициентами:

$$x_2(t) = a_2 t + d_2.$$

Предположим, что предприятие «Альфа» находится в положении догоняющей стороны. Это значит, что в рассматриваемый момент времени t_0 (например, «сегодня») значение показателя у его продукции ниже: $x_1(t_0) < x_2(t_0)$, но темп роста у предприятия «Альфа» выше, чем у конкурента: $a_1 > a_2$.

Возникает естественный вопрос: когда предприятие «Альфа» догонит конкурента? Другими словами, в какой момент времени будет выполнено равенство $x_1(t) = x_2(t)$? Решая относительно t уравнение

$$a_1 t + d_1 = a_2 t + d_2,$$

получаем, что встреча произойдет в момент

$$t_{\text{в}} = \frac{d_2 - d_1}{a_1 - a_2}.$$

Представляют интерес еще две величины. Во-первых, уровень качества, при котором предприятие «Альфа» сравняется с конкурентом, т.е. общий уровень качества в момент встречи:

$$x = x_1(t_{\text{в}}) = x_2(t_{\text{в}}) = \frac{a_1 d_2 - a_2 d_1}{a_1 - a_2}.$$

Во вторых, временной лаг, т.е. величина отставания предприятия «Альфа» в рассматриваемый момент времени t_0. В какой (более ранний) момент времени t_k конкурент имел тот уровень качества, которого предприятие «Альфа» достигло сейчас? Для ответа на этот вопрос надо решить уравнение $x_2(t) = x_1(t_0)$. Решение описывается формулой

$$t_k = \frac{x_1(t_0) - d_2}{a_2}.$$

Следовательно, предприятие «Альфа» отстает на

$$L = t_0 - t_k = \frac{(a_2 - a_1)t_0 + d_2 - d_1}{a_2} = \frac{x_2(t_0) - x_1(t_0)}{a_2} \text{ единиц времени (лет).}$$

В реальных ситуациях линейные зависимости неизвестны. Однако известны исходные данные: $(t_{i1}; x_{i1})$, $i = 1, 2, ..., n(1)$ — для предприятия «Альфа» и $(t_{j2}; x_{j2})$, $j = 1, 2, ..., n(2)$ — для предприятия «Бета». При этом значения показателя $x_1(t_{i1}) = x_{i1}$ у предприятия «Альфа» в моменты времени t_{i1} представляются в виде

$$x_1(t_{i1}) = x_{i1} = a_1 t_{i1} + d_1 + e_{i1},\ i = 1, 2, ..., n(1),$$

где коэффициенты a_1 и d_1 неизвестны статистику, а e_{i1} — погрешности измерения (невязки).

Будем считать, что e_{i1}, $i = 1, 2, ..., n(1)$ — совокупность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием и дисперсией $D(e_{i1}) = \sigma_1^2$, неизвестной статистику.

Для предприятия-конкурента «Бета» справедливо аналогичное представление:

$$x_2(t_{j2}) = x_{j2} = a_2 t_{j2} + d_2 + e_{j2},\ j = 1, 2, ..., n(2),$$

где коэффициенты a_2 и d_2 неизвестны статистику, а e_{j2} — погрешности измерения (невязки).

Примем, что e_{j2}, $j = 1, 2, ..., n(2)$ — совокупность независимых одинаково распределенных случайных величин с нулевым математическим ожиданием и дисперсией $D(e_{i1}) = \sigma_2^2$, неизвестной статистику.

Примем, что две совокупности случайных величин e_{i1}, $i = 1, 2, ...$ ..., $n(1)$, и e_{j2}, $j = 1, 2, ..., n(2)$, независимы между собой. В каждой совокупности случайные величины одинаково распределены, но функции

распределения, соответствующие разным совокупностям (т.е. предприятию «Альфа» и предприятию-конкуренту «Бета»), могут различаться между собой.

Подчеркнем, что в рассматриваемой вероятностно-статистической модели не предполагается, 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.

Метод решения. Рассмотрим 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 подразделе 5.1, а затем рассчитать оценки момента встречи $t_{\text{в}}^*$, уровня качества в момент встречи $x^* = x_1^*(t_{\text{в}}^*) = x_2^*(t_{\text{в}}^*)$ и временно́го 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 подразделе 5.1, 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

$$t_{\text{ср}}(1) = \frac{t_{11} + t_{21} + \ldots + t_{n(1)1}}{n(1)}, \quad t_{\text{ср}}(2) = \frac{t_{12} + t_{22} + \ldots + t_{n(2)2}}{n(2)}. \qquad (5.9)$$

Таким образом,

$$d_k = b_k - a_k t_{\text{ср}}(k),\ k = 1, 2.$$

Дело в том, что асимптотическое описание совместного распределения коэффициентов проще в случае центрированной зависимости, в частности, оценки коэффициентов $a_k^*, b_k^*, k=1,2$ асимптотически независимы.

Как известно (см. подраздел 5.1), оценки метода наименьших квадратов имеют вид:

$$a_k^* = \frac{\sum_{i=1}^{n(k)} x_{ik}\left[t_{ik} - t_{\text{ср}}(k)\right]}{\sum_{i=1}^{n(k)}\left[t_{ik} - t_{\text{ср}}(k)\right]^2}; \qquad (5.10)$$

$$b_k^* = x_{\text{ср}}(k) = \frac{x_{1k} + x_{2k} + \ldots + x_{n(k)k}}{n(k)},\ k=1,2. \qquad (5.11)$$

Точечные оценки момента встречи $t_{\text{в}}^*$, уровня качества в момент встречи x^* и временнóго лага L^* выражаются через оценки коэффициентов линейных зависимостей:

$$t_{\text{в}}^* = \frac{d_2^* - d_1^*}{a_1^* - a_2^*} = \frac{b_2^* - b_1^* + a_1^* t_{\text{ср}}(1) - a_2^* t_{\text{ср}}(2)}{a_1^* - a_2^*}; \qquad (5.12)$$

$$x^* = \frac{a_1^* d_2^* - a_2^* d_1^*}{a_1^* - a_2^*} = \frac{a_1^* b_2^* - a_2^* b_1^* + a_1^* a_2^*\left[t_{\text{ср}}(1) - t_{\text{ср}}(2)\right]}{a_1^* - a_2^*}; \qquad (5.13)$$

$$L^* = \frac{(a_2^* - a_1^*)t_0 + d_2^* - d_1^*}{a_2^*} = \frac{(a_2^* - a_1^*)t_0 + b_2^* - b_1^* + a_1^* t_{\text{ср}}(1) - a_2^* t_{\text{ср}}(2)}{a_2^*}. \qquad (5.14)$$

Из приведенных формул вытекает, что

$$t_{\text{в}}^* = f_1(a_1^*, a_2^*, b_1^*, b_2^*)\ x^* = f_2(a_1^*, a_2^*, b_1^*, b_2^*)\quad L^* = f_1(a_1^*, a_2^*, b_1^*, b_2^*),$$

где

$$f_1(z_1, z_2, z_3, z_4) = \frac{z_4 - z_3 + z_1 t_{\text{ср}}(1) - z_2 t_{\text{ср}}(2)}{z_1 - z_2};$$

$$f_2(z_1, z_2, z_3, z_4) = \frac{z_1 z_4 - z_2 z_3 + z_1 z_2\left[t_{\text{ср}}(1) - t_{\text{ср}}(2)\right]}{z_1 - z_2};$$

$$f_3(z_1, z_2, z_3, z_4) = \frac{(z_2 - z_1)t_0 + z_4 - z_3 + z_1 t_{\text{ср}}(1) - z_2 t_{\text{ср}}(2)}{z_2}.$$

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.

Рассмотрим асимптотическое распределение вектора оценок, полученных методом наименьших квадратов в рамках описанной ранее непараметрической вероятностно-статистической модели (см. подраздел 5.1). Оценки $a_1^*, a_2^*, b_1^*, b_2^*$ являются несмещенными, их дисперсии таковы:

$$D(a_k^*) = \frac{\sigma_k^2}{\sum_{i=1}^{n(k)}\left[t_{ik} - t_{\text{ср}}(k)\right]^2},\ d(b_k^2) = \frac{\sigma_k^2}{n(k)},\ k=1,2.$$

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 5.1. 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

$$\lim_{n(k)\to\infty} \max_{1\le i\le n(k)} \left|t_{ik} - t_{\text{ср}}(k)\right| / \left\{\sum_{i=1}^{n(k)}\left[t_{ik} - t_{\text{ср}}(k)\right]^2\right\}^{1/2} = 0,\ k=1,2,$$

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 ранее параметрами.

Рассмотрим распределение функции от вектора оценок, полученных методом наименьших квадратов. Если функция $f(z_1, z_2,$

z_3, z_4) достаточно гладкая, то согласно методу линеаризации [77, подраздел 4.4]

$$f(a_1^*, a_2^*, b_1^*, b_2^*) - f(a_1, a_2, b_1, b_2) = \frac{\partial f}{\partial z_1}\left(a_1^* - a_1\right) +$$

$$+\frac{\partial f}{\partial z_2}\left(a_2^* - a_2\right) + \frac{\partial f}{\partial z_3}\left(b_1^* - b_1\right) + \frac{\partial f}{\partial z_4}\left(b_2^* - b_2\right)$$

с точностью до бесконечно малых величин более высокого порядка.

Как показано ранее, 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, полученных методом наименьших квадратов, является асимптотически нормальной случайной величиной с математическим ожиданием — функцией от неизвестных параметров $f(a_1, a_2, b_1, b_2)$, совпадающим с теоретическим значением, и дисперсией

$$D\left[f(a_1^*, a_2^*, b_1^*, b_2^*)\right] = \left(\frac{\partial f}{\partial z_1}\right)^2 D\left(a_1^*\right) + \left(\frac{\partial f}{\partial z_2}\right)^2 D\left(a_2^*\right) +$$

$$+\left(\frac{\partial f}{\partial z_3}\right)^2 D\left(b_1^*\right) + \left(\frac{\partial f}{\partial z_4}\right)^2 D\left(b_2^*\right).$$

Подставив приведенные ранее значения дисперсий, получаем, что

$$D\left[f(a_1^*, a_2^*, b_1^*, b_2^*)\right] = \left(\frac{\partial f}{\partial z_1}\right)^2 \frac{\sigma_1^2}{\sum_{i=1}^{n(1)}\left[t_{i1} - t_{\text{ср}}(1)\right]^2} +$$

$$+\left(\frac{\partial f}{\partial z_2}\right)^2 \frac{\sigma_2^2}{\sum_{i=1}^{n(2)}\left[t_{i2} - t_{\text{ср}}(2)\right]^2} + \left(\frac{\partial f}{\partial z_3}\right)^2 \frac{\sigma_1^2}{n(1)} + \left(\frac{\partial f}{\partial z_4}\right)^2 \frac{\sigma_2^2}{n(2)}.$$

Асимптотическое решение. Рассмотрим асимптотическое распределение момента встречи. Начнем с функции $t_{\text{в}}^* = f_1(a_1^*, a_2^*, b_1^*, b_2^*)$. Имеем:

$$\frac{\partial f_1}{\partial z_1} = \frac{z_3 - z_4 + z_2\left[t_{\text{ср}}(2) - t_{\text{ср}}(1)\right]}{\left(z_1 - z_2\right)^2}, \quad \frac{\partial f_1}{\partial z_2} = \frac{z_4 - z_3 + z_1\left[t_{\text{ср}}(1) - t_{\text{ср}}(2)\right]}{\left(z_1 - z_2\right)^2};$$

$$\frac{\partial f_1}{\partial z_3} = \frac{1}{z_1 - z_2}; \quad \frac{\partial f_1}{\partial z_4} = \frac{1}{z_1 - z_2}.$$

В приведенных формулах частные производные можно брать как в точке (a_1, a_2, b_1, b_2), так и в точке $(a_1^*, a_2^*, b_1^*, b_2^*)$. Различие — бесконечно малые величины более высокого порядка. Поскольку истинные значения коэффициентов линейных зависимостей неизвестны, частные производные будем брать в точке.

Из последних формул с помощью несложных преобразований получаем, что

$$D(t_{\text{в}}^*) = \left\{\frac{b_1^* - b_2^* + a_2^*\left[t_{\text{ср}}(2) - t_{\text{ср}}(1)\right]}{(a_1^* - a_2^*)^2}\right\}^2 \frac{\sigma_1^2}{\sum_{i=1}^{n(1)}\left[t_{i1} - t_{\text{ср}}(1)\right]^2} +$$

$$+\left\{\frac{b_1^* - b_2^* + a_1^*\left[t_{\text{ср}}(2) - t_{\text{ср}}(1)\right]}{(a_1^* - a_2^*)^2}\right\}^2 \frac{\sigma_2^2}{\sum_{i=1}^{n(2)}\left[t_{i2} - t_{\text{ср}}(2)\right]^2} +$$

$$+\frac{1}{(a_1^* - a_2^*)^2}\left[\frac{\sigma_1^2}{n(1)} + \frac{\sigma_2^2}{n(2)}\right]. \tag{5.15}$$

Для практического применения полученных результатов остается заменить неизвестные дисперсии невязок σ_1^2 и σ_2^2 на их состоятельные оценки. 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;

$$SS(k) = \sum_{i=1}^{n(k)}\left[x_{ik} - x_k^*(t_{ik})\right]^2, \; k = 1, 2. \tag{5.16}$$

Иногда рекомендуют применение несмещенных оценок дисперсий невязок:

$$(\sigma_1^2)^{**} = \frac{SS(1)}{n(1) - 2}; \quad (\sigma_2^2)^{**} = \frac{SS(2)}{n(2) - 2}. \tag{5.17}$$

Ясно, что с ростом объемов данных $n(1)$ и $n(2)$ различие между двумя последними формулами исчезает.

На основе полученных результатов легко указать методы доверительного оценивания и проверки гипотез для момента встречи $t_{\text{в}}$. Так, асимптотический доверительный интервал, соответствующий доверительной вероятности p, имеет вид

$$\left[t_в^* - U(p)\left[D^*(t_в^*)\right]^{1/2};\ t_в^* + U(p)\left[D^*(t_в^*)\right]^{1/2}\right], \qquad (5.18)$$

где $D^*(t_в^*)$ — только что описанная оценка дисперсии случайной величины $t_в$ (с использованием той или иной оценки дисперсий невязок); $U(p)$ — квантиль стандартного нормального распределения порядка $(1 + p)/2$, т.е. $\Phi[U(p)] = \frac{1+p}{2}$, где Φ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

Пример 5.4. В качестве примера рассмотрим показатели технического уровня продукции (в условных единицах) двух предприятий — ОАО «Альфа» и «Бета». Приведенные в таблице данные показывают, что в 1999 г. первое предприятие отстает от второго, но постепенно сокращает разрыв, более быстрыми темпами наращивая показатель технического уровня. Когда же оно догонит второе предприятие?

Показатели технического уровня продукции двух предприятий (в условных единицах, на конец года)

Показатель	Годы/условные моменты времени $t_{i1} = t_{j2}$						
	1999/1	2000/2	2001/3	2002/4	2003/5	2004/6	2005/7
ОАО «Альфа»							
x_{i1}	0,1	0,2	0,6	0,5	0,8	0,9	1,3
Восстановленные значения x_{i1}^*	0,0715	0,2572	0,4429	0,6286	0,8143	1,0000	1,1857
ОАО «Бета»							
x_{j2}	0,6	0,95	0,8	1,2	1,1	1,2	1,4
Восстановленные значения x_{j2}^*	0,6928	0,8071	0,9214	1,0357	1,1500	1,2643	1,3786

Для проведения расчетов естественным образом введем условные моменты времени (см. таблицу). Методом наименьших квадратов восстановим линейные независимости. По формуле (5.9) получаем, что $t_{ср}(1) = t_{ср}(2) = 4$. По формулам (5.10) и (5.11) находим оценки коэффициентов линейных зависимостей

$$a_1^* = 0{,}1857,\quad a_2^* = 0{,}1143,\quad b_1^* = 0{,}6286,\quad b_2^* = 1{,}0357.$$

Восстановленные зависимости имеют вид

$$x_1^*(t) = 0{,}1857(t-4) + 0{,}6286,\ \ x_2^*(t) = 0{,}1143(t-4) + 1{,}0357.$$

Восстановленные значения приведены в таблице.

Оценку момента встречи $t_в^*$ определим по формуле (5.12)

$$t_в^* = \frac{1{,}0357 - 0{,}6286 + 0{,}1857\times 4 - 0{,}1143\times 4}{0{,}1857 - 0{,}1143} = 9{,}7017.$$

Другими словами, значения показателей технического уровня предприятий сравняются в начале 2008 г. Это общее значение найдем по формуле (5.13):

$$x^* = \frac{0{,}1857\times 1{,}0357 - 0{,}1143\times 0{,}6286 + 0{,}1857\times 0{,}1143(4-4)}{0{,}1857 - 0{,}1143} = 1{,}6877.$$

Для определения временного лага, т.е. величины, показывающей, на сколько предприятие «Альфа» отстает от предприятия-конкурента «Бета» (для определенности — в 2004 г.), воспользуемся формулой (5.14):

$$L^* = \frac{(0{,}1143 - 0{,}1857)6 + 1{,}0357 - 0{,}6286 + 0{,}1857\times 4 - 0{,}1143\times 4}{0{,}1143} = 2{,}3123.$$

Рассчитаем по формуле (5.18) асимптотический доверительный интервал, соответствующий доверительной вероятности $p = 0{,}95$. Для этого значения несмещенных оценок дисперсий невязок найдем по формуле (5.17), а значения остаточной суммы квадратов $SS(1)$ и $SS(2)$ определим по формуле (5.16). Получаем $\sigma_1^2 = 0{,}0137$, $\sigma_2^2 = 0{,}0156$. Асимптотическую дисперсию момента встречи найдем по формуле (5.15): $D^*(t_в^*) = 7{,}4883$. Поскольку $U(p) = 1{,}96$ при $p = 0{,}95$, то доверительный интервал таков (рис. 5.6):

$$\left[9{,}7017 - 1{,}96\sqrt{7{,}4883}; 9{,}7017 + 1{,}96\sqrt{7{,}4883}\right] = [4{,}3383; 15{,}0652].$$

Таким образом, возможно, что обгон уже состоялся (в 2004 или в 2005 г.), но это не отражено в таблице из-за погрешностей, искажающих зависимости.

О практическом применении статистических оценок точки встречи. Необходимость получения приведенных ранее результатов, касающихся оценок момента встречи $t_в^*$, уровня качества в момент встречи x^* и временного лага L^*, была выявлена в результате решения прикладных проблем, возникших при разработке системы автоматического проектирования (САПР) стандартов на продукцию. В этой системе

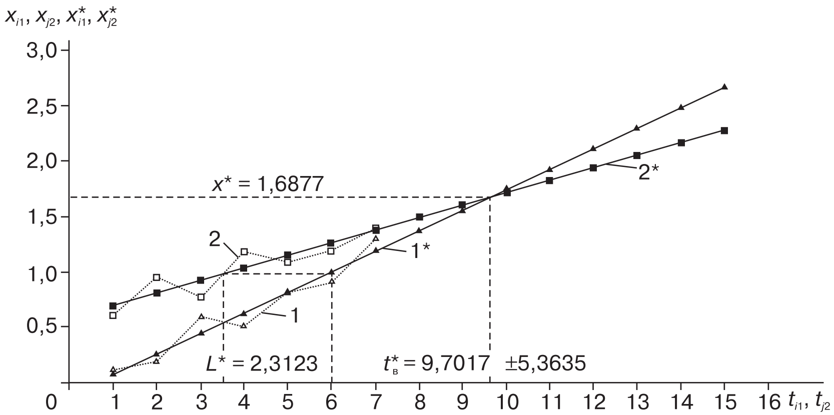


Рис. 5.6. Динамика показателей технического уровня двух предприятий (1, 2), восстановленные зависимости (1*, 2*) и доверительное оценивание момента встречи

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 интерактивного (человеко-машинного) принятия решений по управлению качеством, сертификации и 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 интерактивного решения задач пользователями (инженерами по стандартизации и техническому регулированию), не имеющими специальной подготовки по статистическим методам. Кроме того, возникли специфические задачи, требующие совершенствования статистического аппарата, в частности задача сравнительного анализа тенденций развития отечественной и зарубежной групп продукции, решение которой было приведено ранее. Разработанное математическое и программное обеспечение применялось для анализа данных о характеристиках качества изделий электронной техники.

Разработанные в настоящем подразделе методы могут быть использованы при решении различных практических задач, связанных с интервальной оценкой точки пересечения двух регрессионных прямых. В частности, они использованы для получения асимптотических

дисперсий уровня качества в момент встречи и временнóго лага (величины отставания) и разработки методов доверительного оценивания этих величин, используемых в системах управления промышленными предприятиями [56].

5.7. МОДЕЛЬ С ПЕРИОДИЧЕСКОЙ СОСТАВЛЯЮЩЕЙ

Рассмотрим задачу восстановления зависимости $x = x(t)$ на основе набора n пар чисел (t_k, x_k), $k = 1, 2, ..., n$, где t_k — значения независимой переменной, а x_k — соответствующие им значения зависимой переменной.

При анализе экономических данных возникает необходимость использования моделей временных рядов, включающих три составляющие: трендовую (T), периодическую, или циклическую, (S) и случайную (E). Рассматривают аддитивную модель $T + S + E$ и мультипликативную модель TSE.

Простейшая аддитивная модель имеет вид

$$x_k = a\left(t_k - \bar{t}\right) + d + e_k = a\left(t_k - \bar{t}\right) + d + f(t_k) + E_k,\ k = 1, 2, \ldots, n, \quad (5.19)$$

где трендовая составляющая — линейная функция $a\left(t_k - \bar{t}\right) + d$ (как и в подразделе 5.1, такая запись тренда предпочтительнее для облегчения выкладок); периодическая составляющая $f(t)$ обычно описывает сезонность, т.е период известен (в зависимости от моделируемой органиазционно-экономической ситуации он равен году, неделе, суткам и т.п.); случайная составляющая представлена слагаемыми E_k, которые являются реализациями независимых одинаково распределенных случайных величин с нулевым математическим ожиданием и дисперсией σ^2, неизвестной статистику.

В модели (5.19) имеем:

$$e_k = f(t_k) + E_k, = 1, 2, ..., n.$$

В отличие от модели, изученной в подразделе 5.1, отклонения от линейного тренда e_k в модели, описываемой формулой (5.19), не являются одинаково распределенными. Однако их распределения отличаются лишь сдвигами (на значения детерминированной периодической составляющей).

Соответствующая мультипликативная модель имеет вид

$$y_k = Bt_k^a \times f_1(t_k)\left(1 + \varepsilon_k\right), \quad k = 1, 2, \ldots, n. \quad (5.20)$$

В формуле (5.20) сомножители имеют описанный ранее смысл. При логарифмировании модель, описываемая формулой (5.20), переходит в аналог модели (5.19), следовательно, достаточно рассматривать модель (5.19).

Практическая значимость этой модели очевидна. Однако описанные расчетные методы являются эвристическими. Цель настоящего раздела — ***построить непараметрическую вероятностно-статистическую теорию прогноза временнóго ряда на базе линейного тренда с учетом аддитивной периодической составляющей*** (см. работу [73]).

Следуя эвристическому подходу, изучим асимптотическое поведение оценок, полученных методом наименьших квадратов a^* и d^*, заданных формулами (5.3), установим их асимптотическую нормальность в предположениях модели (5.19), а затем состоятельно оценим периодическую составляющую $f(t)$ и построим интервальный прогноз для $x(t)$. В частности, выявится целесообразность анализа данных за полное число лет (периодов). В отличие от методики, описанной в [77], длину периода оценивать не требуется, поскольку она задана из содержательных соображений (например, для данных подраздела 5.4 — один год).

Асимптотические распределения оценок параметров. Из формул (5.3) следует, что в принятых ранее предположениях и обозначениях настоящего подраздела

$$\begin{aligned} d^* &= \frac{a}{n}\sum_{k=1}^{n}\left(t_k-\bar{t}\right)+d+\frac{1}{n}\sum_{k=1}^{n}e_k=d+\frac{1}{n}\sum_{k=1}^{n}e_k= \\ &= d+\frac{1}{n}\sum_{k=1}^{n}f(t_k)+\frac{1}{n}\sum_{k=1}^{n}E_k. \end{aligned} \quad (5.21)$$

Согласно центральной предельной теореме оценка d^* имеет асимптотически нормальное распределение с математическим ожиданием $d+\frac{1}{n}\sum_{k=1}^{n}f(t_k)$ и дисперсией σ^2/n, оценка которой приводится далее.

Из формул (5.3) и (5.21) вытекает, что

$$x_k-\bar{x}=a\left(t_k-\bar{t}\right)+d+e_k-d-\frac{1}{n}\sum_{k=1}^{n}e_k=a\left(t_k-\bar{t}\right)+e_k-\frac{1}{n}\sum_{k=1}^{n}e_k,$$

$$\left(x_k-\bar{x}\right)\left(t_k-\bar{t}\right)=a\left(t_k-\bar{t}\right)^2+e_k\left(t_k-\bar{t}\right)-\frac{\left(t_k-\bar{t}\right)}{n}\sum_{k=1}^{n}e_k.$$

Последнее слагаемое во втором соотношении при суммировании по k обращается в 0, поэтому

$$a^*=a+\sum_{k=1}^{n}c_ke_k=a+\sum_{k=1}^{n}c_kf(t_k)+\sum_{k=1}^{n}c_kE_k,\; c_k=\frac{\left(t_k-\bar{t}\right)}{\sum_{k=1}^{n}\left(t_k-\bar{t}\right)^2}. \quad (5.22)$$

Формулы (5.22) показывают, что оценка a^* является асимптотически нормальной с математическим ожиданием $a+\sum_{k=1}^{n}c_kf(t_k)$ и дисперсией

$$D(a^*)=\sum_{k=1}^{n}c_k^2D(E_k)=\frac{\sigma^2}{\sum_{k=1}^{n}\left(t_k-\bar{t}\right)^2}.$$

Отметим, что многомерная нормальность имеет быть, когда каждое слагаемое в формулах (5.22) мало сравнимо со всей суммой, т.е.

$$\lim_{n\to\infty}\max\left|t_k-\bar{t}\right|/\left[\sum_{k=1}^{n}\left(t_k-\bar{t}\right)^2\right]^{1/2}=0. \quad (5.23)$$

Условие (5.23) выполнено, если t_k образуют (полную, т.е. без пропусков) арифметическую прогрессию, число членов которой безгранично растет.

Итак, дисперсии оценок, полученных методом наименьших квадратов параметров a^* и b^* линейного тренда, те же, что и при отсутствии сезонных искажений (см. подраздел 5.1). А вот их математические ожидания зависят от периодической составляющей. Однако в случае

$$\sum_{k=1}^{n}f(t_k)=0,\quad \sum_{k=1}^{n}\left(t_k-\bar{t}_{\text{ср}}\right)f(t_k)=0 \quad (5.24)$$

оценки a^* и b^* являются несмещенными.

Условия формулы (5.24) являются принципиально важными. Они являются необходимыми и достаточными для несмещенности и состоятельности оценок, рассмотренных в настоящем разделе.

Справедливости первого из условий (5.21) можно добиться, изменив в случае необходимости свободный член d в модели, описываемой формулой (5.19). Некоторая проблема состоит в том, какую сумму значений периодической составляющей приравнивать 0 — за период (например за год) или за все время наблюдений, как и записано в формуле (5.24). С организационно-экономической точки зрения естественнее первое (т.е. приоритет отдается свойствам за период, интервал наблюдений может меняться, например расширяться со временем). Когда оба варианта дают одно и то же?

Первое из условий формулы (5.24) можно считать выполненным, если t_k образуют (полную, т.е. без пропусков) арифметическую прогрессию, причем целое число шагов составляет один период (например, если измерения проводятся ежемесячно или раз в квартал, а период — год), и, кроме того, данные взяты за целое число периодов. Действительно, тогда естественно принять, что сумма значений периодической составляющей за период равна 0, поскольку в противном случае, как уже отмечалось, можно было бы скорректировать свободный член (т.е. по тем же соображениям, по которым принято условие нулевого математического ожидания случайных составляющих E_i).

Для справедливости второго из условий формулы (5.24) достаточно добавить к сказанному предположения симметричности множества $\{t_k, k = 1, 2, ..., n\}$ относительно $\bar{t}$ (например, начала года) и четности периодической составляющей $f(t)$ относительно той же точки. Последнее выполнено, если, например, график $f(t)$ симметричен относительно середины года.

Несмещенность — в предположениях формулы (5.24), и асимптотическая нормальность оценок метода наименьших квадратов позволяют легко указывать для них асимптотические доверительные границы и проверять статистические гипотезы, например о равенстве определенным значениям, прежде всего 0.

Асимптотическое распределение трендовой составляющей. Из формул (5.21) и (5.22) следует, что при справедливости (5.24)

$$M\left[a^*\left(t-\bar{t}\right)+d^*\right]=M\left(a^*\right)\left(t-\bar{t}\right)+M\left(d^*\right)=a\left(t-\bar{t}\right)+d,$$

т.е. оценка $y^*(t)a^*\left(t-\bar{t}\right)+d^*$ трендовой составляющей $y(t)a^*\left(t-\bar{t}\right)+d$ рассматриваемой зависимости является несмещенной. Поэтому

$$D\left[y^*(t)\right]=D\left(a^*\right)\left(t-\bar{t}\right)^2+2M\left[\left(a^*-a\right)\left(d^*-d\right)\left(t-\bar{t}\right)\right]+D\left(d^*\right).$$

При этом, поскольку погрешности E_k независимы в совокупности и $M(E_k) = 0$,

$$M\left[\left(a^*-a\right)\left(d^*-d\right)\left(t-\bar{t}\right)\right]=\frac{1}{n}\sum_{k=1}^{n}c_k\left(t-\bar{t}\right)M\left(E_k^2\right)=\frac{1}{n}\left(t-\bar{t}\right)\sigma^2\sum_{k=1}^{n}c_k=0.$$

Таким образом,

$$D\left[y^*(t)\right]=\sigma^2\left[\frac{1}{n}+\frac{\left(t-\bar{t}\right)^2}{\sum_{k=1}^{n}\left(t_k-\bar{t}\right)^2}\right]. \qquad (5.25)$$

Итак, оценка $y^*(t)$ является несмещенной и асимптотически нормальной. Для ее практического использования (построения доверительных интервалов, проверки статистических гипотез) необходимо уметь оценивать остаточную дисперсию $\sigma^2 = M(E_k^2)$.

В частности, не представляет труда выписывание нижней и верхней границ для трендовой составляющей прогностической функции:

$$y_{\text{нижн}}(t)=a^*\left(t-\bar{t}\right)+d^*-\delta(t);\ y_{\text{верх}}(t)=$$
$$=a^*\left(t-\bar{t}\right)+d^*-\delta(t),$$

где полуширина доверительного интервала $\delta(t)$ имеет вид:

$$\delta(t)=U(\gamma)\sqrt{D^*\left[y^*(t)\right]}=U(\gamma)\sigma^*\sqrt{\frac{1}{n}+\frac{\left(t-\bar{t}\right)^2}{\sum_{k=1}^{n}\left(t_k-\bar{t}\right)^2}}, \qquad (5.26)$$

где γ — доверительная вероятность, $U(\gamma)$ — квантиль нормального распределения порядка $\frac{1+\gamma}{2}$, 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.

При $\gamma = 0{,}95$ (наиболее применяемое значение) имеем $U(\gamma) = 1{,}96$. В формуле (5.26) $D^*\left[y^*(t)\right]$ — состоятельная оценка дисперсии $y^*(t)$. В соответствии с формулой (5.25) она является произведением состоятельной оценки σ^* среднего квадратичного отклонения σ случайных погрешностей E_k на известную статистику детерминированную функцию от t.

Математическое ожидание остаточной суммы квадратов. В точках t_k, $k = 1, 2, ..., n$, имеются исходные значения зависимой переменной x_k и восстановленные значения $y^*(t_k)$. Рассмотрим остаточную сумму квадратов

$$SS=\sum_{k=1}^{n}\left[y^*(t_k)-x_k\right]^2=$$
$$=\sum_{k=1}^{n}\left[\left(a^*-a\right)\left(t_k-\bar{t}\right)+\left(d^*-d\right)-f(t_k)-E_k\right]^2.$$

Напомним, что при отсутствии периодической составляющей используют (см. подраздел 5.1 и [89, подразделы 5.1 и 5.2]) состоятельные оценки σ^* среднего квадратичного отклонения σ^* случайных погрешностей, построенные на основе остаточной суммы квадратов:

$$\sigma^*=\sqrt{\frac{SS}{n}}\ \text{ или }\ \sigma=\sqrt{\frac{SS}{n-2}}.$$

В соответствии с формулами (5.21) и (5.22) при справедливости условий (5.24)

$$SS = \sum_{k=1}^{n}\left[\left(t_k - \bar{t}\right)\sum_{j=1}^{n} c_j E_j + \frac{1}{n}\sum_{j=1}^{n} E_j - f(t_k) - E_k\right]^2 =$$

$$= \sum_{k=1}^{n}\left\{\sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]E_j - f(t_k) - E_k\right\}^2 = \sum_{k=1}^{n} SS_k.$$

Найдем математическое ожидание каждого из слагаемых:

$$M(SS_k) = M\left\{\sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]E_j - f(t_k) - E_k\right\}^2 =$$

$$= M\left\{\sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]E_j\right\}^2 - 2M\left\{\sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]E_j\right\}\left[f(t_k) + E_k\right] +$$

$$+ M\left[f(t_k) - E_k\right]^2.$$

Поскольку E_k независимы, одинаково распределены и имеют нулевое математическое ожидание, то

$$M\left\{\sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]E_j\right\}^2 = \sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]^2 \sigma^2.$$

Далее,

$$-2M\left\{\sum_{j=1}^{n}\left[c_j\left(t_k - \bar{t}\right) + \frac{1}{n}\right]E_j\right\}\left[f(t_k) + E_k\right] = -2\left[c_k\left(t_k - \bar{t}\right) + \frac{1}{n}\right]\sigma^2.$$

Наконец,

$$M\left[f(t_k) - E_k\right]^2 = f^2(t_k) + \sigma^2.$$

На основе трех последних равенств можно показать, что при выполнении условия асимптотической нормальности (5.23)

$$\lim_{n\to\infty} M(SS_k) = F^2(t_k) + \sigma^2,$$

следовательно,

$$M\left(\frac{SS}{n}\right) = \sigma^2 + \frac{1}{n}\sum_{k=1}^{n} f^2(t_k). \tag{5.27}$$

В правой части формулы (5.27) первое слагаемое соответствует вкладу случайной составляющей, второе — вкладу периодической составляющей.

В некоторых случаях второе слагаемое в правой части формулы (5.27) может быть известно из предыдущего опыта или же оценено экспертами, однако в большинстве ситуаций целесообразно исходить из оценки периодической составляющей.

Оценивание сезонной компоненты. Рассматривают как параметрические, так и непараметрические подходы. Популярный метод исходит из того, что достаточно гладкую функцию можно разложить в ряд Фурье и получить хорошее приближение с помощью небольшого числа гармоник. В простейшем случае — одна гармоника. Так, динамику индекса инфляции можно попытаться изучать с помощью модели

$$x_k = a\left(t_k - \bar{t}\right) + d + f(t_k) + E_k = a\left(t_k - \bar{t}\right) + d +$$
$$+ g\cos(2\pi t_k) + E_k,\ k = 1, 2, \ldots, n.$$

Здесь время t измеряется в годах. Тогда неизвестные параметры a, b, g оцениваются методом наименьших квадратов.

Однако обычно нет оснований предполагать, что периодическая составляющая входит в то или иное параметрическое семейство функций. Приходится строить непараметрические оценки. Опишем одну из возможных постановок.

Пусть в согласии с предположениями формулы (5.24) рассматривается целое число периодов, т.е. $n = mq$, где n — объем наблюдений, m — число периодов, q — число наблюдений в одном периоде. Тогда в соответствии с определением периодической составляющей справедливы равенства

$$f(t_s) = f(t_{q+s}) = f(t_{2q+s}) = \ldots = f(t_{(m-1)q+s}),\ s = 1, 2, \ldots, q. \tag{5.28}$$

Если наблюдения проводятся ежемесячно в течение m лет, s — номер месяца в году, $s = 1, 2, \ldots$, то число наблюдений в одном периоде $q = 12$, общий объем наблюдений $n = 12m$ (далее — 12). Обозначим g_s — общее значение в (5.28). Требуется оценить $g_1, g_2, \ldots, g_q$.

Естественный подход состоит в том, чтобы усреднить m значений $x_k - y^*(t_k)$, соответствующих моментам времени, отстоящим друг от друга на целое число периодов. Другими словами, усреднить «очищенные» от трендовой составляющей исходные данные, соответствующие одноименным месяцам различных лет. Речь идет об оценках

$$g_s^* = \frac{1}{m}\sum_{j=1}^{m}\left[x_{s+(j-1)q} - y^*(t_{s+(j-1)q})\right],\ s = 1, 2, \ldots, q. \tag{5.29}$$

Оценка периодической составляющей распространяется на весь интервал наблюдений очевидным образом:

$$f^*(t_s) = f^*(t_{q+s}) = f^*(t_{2q+s}) = \ldots = f^*(t_{(m-1)q+s}) =$$
$$= g_s^*,\ s = 1, 2, \ldots, q. \tag{5.30}$$

Сложив восстановленные значения трендовой и периодической оставляющей, получим оценку зависимости, «очищенную» от случайной составляющей

$$x^*(t) = y^*(t) + f^*(t) = a^*\left(t - \bar{t}\right) + d^* + f^*(t), \qquad (5.31)$$

где оценки a^* и d^* находят по формулам (5.3), а оценки $f^*(t)$ — по формулам (5.29)—(5.30).

С помощью формулы (5.31) можно строить точечный прогноз, используя ее вне интервала наблюдений. Для этого достаточно распространить сезонную составляющую $f^*(t)$ вплоть до рассматриваемого момента времени по правилу, описанному формулой (5.30), и суммировать ее с прогнозом трендовой составляющей $y^*(t)$. Интерполяция и экстраполяция на моменты времени t, не входящие в исходное множество $\{t_k, k = 1, 2, ..., n\}$ и множества, полученные из него сдвигами на целое число периодов, могут быть осуществлены путем линейной интерполяции ближайших значений или иным методом сглаживания.

Обсудим свойства оценок (5.29)—(5.31).

При безграничном росте объема данных и справедливости условий (5.23) и (5.24) оценки a^* и d^* параметров трендовой составляющей являются состоятельными и несмещенными. Поэтому, как можно показать в рассматриваемых в настоящем подразделе условиях, суммы в соответствии с формулой (5.29) оценивают периодическую составляющую состоятельно (при $m \to \infty$) и несмещенно. Как следствие,

$$\frac{1}{n}\sum_{k=1}^{n}\left[f^*(t_k)\right]^2 - \frac{1}{n}\sum_{k=1}^{n} f^2(t_k) \to 0 \qquad (5.32)$$

по вероятности при $n \to \infty$. В соответствии с формулой (5.27) последнее соотношение дает возможность оценить σ^2, а затем построить интервальный прогноз для трендовой составляющей согласно (5.26).

Отметим, что в рассматриваемой ситуации, как правило, n растет, увеличиваясь на величины, кратные q — числу наблюдений в одном периоде. Как следствие, уменьшаемое в (5.32) — константа, которая не зависит от n. Эти особенности связаны с тем, что выполнение условий формулы (5.24) предполагает рассмотрение целого числа периодов.

Рассмотрим оценки формулы (5.29) подробнее. Как вытекает из (5.19), (5.28) и (5.29),

$$g_s^* = f\left(t_s\right) - \left(a^* - a\right)\frac{1}{m}\sum_{j=1}^{m}\left(t_{s+(j-1)q} - \bar{t}\right) - \left(d^* - d\right) + \frac{1}{m}\sum_{j=1}^{m} E_{s+(j-1)q},$$

$$s = 1, 2, \ldots, q.$$

С учетом (5.21), (5.22) и (5.24) получаем, что

$$g_s^* = f\left(t_s\right) - \left(\sum_{k=1}^{n} c_k E_k\right)\left(\frac{1}{m}\sum_{j=1}^{m}\left(t_{s+(j-1)q} - \bar{t}\right)\right) - \frac{1}{n}\sum_{k=1}^{n} E_k + \frac{1}{m}\sum_{j=1}^{m} E_{s+(j-1)q},$$

$$s = 1, 2, \ldots, q.$$

Таким образом,

$$g_s^* = f\left(t_s\right) + \sum_{k=1}^{n} h_{ks} E_k, \;\; s = 1, 2, \ldots, q, \qquad (5.33)$$

где $h_{ks} = -c_k r_s - \frac{1}{n} + \frac{1}{m}$, если $k \in \{s + (j-1)q, j = 1, 2, ..., m\}$, и $h_{ks} = -c_k r_s - \frac{1}{n}$ при всех остальных значениях индекса суммирования k, и $r_s = \frac{1}{m}\sum_{j=1}^{m}\left(t_{s+(j-1)q} - \bar{t}\right)$.

Соотношение (5.33) означает, что рассматриваемые оценки есть суммы независимых случайных величин, а потому с помощью центральной предельной теоремы можно построить доверительные интервалы для рассматриваемых значений периодической составляющей (в предположении справедливости условий (5.23)).

Интервальный прогноз. Точечный прогноз строят по формуле (5.28) на основе $x^*(t)$ — оценки зависимости, «очищенной» от случайной составляющей, но включающей трендовый и периодический компоненты. Если выполнены условия формулы (5.24), то

$$M\left[x^*(t)\right] = x(t) = a\left(t - \bar{t}\right) + d + f(t),$$

т.е. оценка $x^*(t)$ является несмещенной.

При справедливости условий формулы (5.24) с учетом формул (5.21), (5.22) и (5.33) получаем, что для момента времени t, входящего в исходное множество $\{t_k, k = 1, 2, ..., n\}$ или в множества, полученные из него сдвигами на целое число периодов,

$$x^*(t) - x(t) = \left(t - \bar{t}\right)\sum_{k=1}^{n} c_k E_k + \frac{1}{n}\sum_{k=1}^{n} E_k + \sum_{k=1}^{n} h_{ks} E_k. \qquad (5.34)$$

В формуле (5.34) при определении значений коэффициентов h_{ks} в качестве s следует взять номер наименьшего из исходных моментов времени $\{t_k, k = 1, 2, ..., n\}$, отстоящих от рассматриваемого момента t на целое число периодов. С помощью формулы (5.33) заключаем, что

$$x^*(t) - x(t) = \sum_{k=1}^{n} w_{ks} E_k,$$

где $w_{ks} = c_k(t - \bar{t} - r_s) + \frac{1}{m}$, если $k \in \{s + (j-1)q, j = 1, 2, ..., m\}$, и $w_{ks} = c_k(t - \bar{t} - r_s)$ при всех остальных значениях индекса суммирования k, и r_s — то же, что и в формуле (5.33).

В правой части формулы (5.34) стоит сумма независимых случайных величин, поэтому оценка $x^*(t)$ является асимптотически нормальной при справедливости условий (5.23) с математическим ожиданием $x(t)$ и дисперсией

$$D[x(t)] = \sum_{k=1}^{n} w_{ks}^2 D(E_k) = \sigma^2 \sum_{k=1}^{n} w_{ks}^2. \qquad (5.35)$$

Следовательно, нижняя $x_{\text{нижн}}(t)$ и верхняя $x_{\text{верх}}(t)$ доверительные границы для прогностической функции (с учетом как трендовой, так и периодической составляющих) имеют вид:

$$x_{\text{нижн}}(t) = a^*(t-\bar{t}) + d^* + f^*(t) - \Delta(t), \;\; x_{\text{верх}}(t) = \\ = a^*(t-\bar{t}) + d^* + f^*(t) + \Delta(t),$$

где

$$\Delta(t) = U(\gamma)\sqrt{D^*[x^*(t)]} = U(\gamma)\sigma^* \sqrt{\sum_{k=1}^{n} w_{ks}^2}, \qquad (5.36)$$

где γ — доверительная вероятность; $U(\gamma)$ — квантиль нормального распределения порядка $\frac{1+\gamma}{2}$.

В формуле (5.36) $D^*[x^*(t)]$ — состоятельная оценка дисперсии точечного прогноза $x^*(t)$. В соответствии с формулой (5.35) она является произведением состоятельной оценки σ^* среднего квадратичного отклонения σ случайных погрешностей E_k на известную детерминированную функцию от t. Величину σ^* рассчитывают согласно формулам (5.27) и (5.32).

Подведем итоги. По сравнению с эвристическими алгоритмами разработанная в настоящем подразделе теория позволяет:

1) дать общее обоснование этим алгоритмам в рамках асимптотических методов математической статистики и указать условия их применимости — формула (5.23);
2) выявить принципиально важные условия — формула (5.24), необходимые и достаточные для несмещенности и состоятельности рассматриваемых оценок;
3) построить доверительные интервалы для зависимости (прогностической функции) и ее трендовой составляющей.

В рамках математической статистики удается провести анализ не всех распространенных эвристических алгоритмов. Так, довольно часто рекомендуют вначале провести сглаживание («выравнивание») временнóго ряда, например методом скользящих средних. При этом периодическая (сезонная) составляющая меняется, а погрешности (отклонения от суммы трендовой и периодической составляющих) становятся зависимыми случайными величинами, что делает невозможным применение описанных в настоящем подразделе методов.

Пример 5.5. Рассмотрим применение непараметрического метода наименьших квадратов в модели с периодической составляющей. Обработаем фактические данные ОАО «ММК» о закупочных ценах на лом черных металлов (см. табл. 5.4). Как следует из подраздела 5.4, может быть использована рассмотренная в настоящем подразделе аддитивная модель линейного тренда с периодической составляющей, описанная формулой (5.19). Для облегчения понимания оставим из каждого квартала данные только по одному месяцу. Введем условные моменты времени, а именно будем измерять время в кварталах, начиная с первого квартала 2003 г. Исходные данные для демонстрации примера применения непараметрического метода наименьших квадратов в модели с периодической составляющей — пары чисел (t_k, x_k), $k = 1, 2, ..., 12$ — представлены в табл. 5.8 в столбцах (3) и (4) соответственно.

Таблица 5.8

Построение модели прогнозирования цен на лом марки 3А

№ п/п	Период	Условные моменты времени	Закупочные цены, руб./т	Оценка тренда	Отклонения от оценки тренда	Восстановленные значения	Кажущиеся невязки
k		t_k	x_k	$y^*(t_k)$	$x_k - y^*(t_k)$	x_k^*	$x_k - x_k^*$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			2003 г.				
1	Январь	1	2 750	2 800	–50	2 424	326
2	Апрель	2	3 800	3 012	788	3 545	255
3	Июль	3	2 900	3 224	–324	2 655	245
4	Октябрь	4	3 100	3 437	–337	3 848	–748
			2004 г.				
5	Январь	5	2 761	3 649	–888	3 273	–512
6	Апрель	6	4 602	3 861	741	4 394	208
7	Июль	7	3 540	4 073	–533	3 504	36
8	Октябрь	8	5 268	4 286	982	4 697	571
			2005 г.				
9	Январь	9	4 307	4 498	–191	4 122	185
10	Апрель	10	4 779	4 710	69	5 243	–464
11	Июль	11	4 071	4 922	–851	4 353	–280
12	Октябрь	12	5 723	5 135	588	5 546	177

По формулам (5.3) найдем оценки параметров a^* и d^*, что позволяет построить оценку трендовой составляющей

$$y^*(t) = a\left(t - \bar{t}\right) + d^* = 212{,}26\left(t - 6{,}5\right) +$$
$$+ 3967{,}17 = 212{,}26t + 2587{,}48.$$

Численные значения трендовой составляющей приведены в столбце (5) табл. 5.8.

Рассчитав отклонения исходных значений закупочных цен от оценок трендовой составляющей — столбец (6) табл. 5.8, возведя их в квадрат и сложив, получаем остаточную сумму квадратов SS = 4 539 214 и $SS/n = SS/12$ = 378 267,843.

Сгруппировав отклонения исходных значений закупочных цен от оценок трендовой составляющей по месяцам (табл. 5.9), наглядно убеждаемся в наличии периодической составляющей. Взяв среднее арифметическое отклонение от тренда за конкретный месяц, рассчитываем оценку $f^*(t_s)$ периодической составляющей в соответствии с формулой (5.29). Результаты приведены в табл. 5.9.

Таблица 5.9

Оценивание периодической составляющей

Номер квартала s	Месяц	Отклонения от тренда			Оценка периодической составляющей $g_s^* = f^*(t_s)$
		2003 г.	2004 г.	2005 г.	
1	Январь	–50	–888	–191	–376
2	Апрель	788	741	69	533
3	Июль	–324	–533	–851	–569
4	Октябрь	–337	982	588	411

Рассчитав по формуле (5.30) оценки периодической составляющей на весь интервал времени и сложив их с оценками трендовой составляющей, получаем в соответствии с формулой (5.31) оценку зависимости, «очищенную» от случайной составляющей, т.е. восстановленные значения — столбец (7) табл. 5.8. Кажущиеся невязки, т.е. отклонения исходных значений закупочных цен от восстановленных значений, приведены в столбце (8) табл. 5.8. Сравнивая столбцы (6) и (8), убеждаемся в целесообразности введения в модель периодической составляющей. В 9 случаях из 12 абсолютные величины отклонений уменьшились, в остальных трех хотя и возросли, но лишь до среднего уровня среди остальных.

Возведя в квадрат оценки периодической составляющей (см. табл. 5.9), сложив эти квадраты, умножив на число лет и поделив на n, получаем, что

$$\frac{1}{n}\sum_{k=1}^{n}\left[f^*(t_k)\right]^2 = 229\,535.$$

В соответствии с формулой (5.27) оценкой дисперсии случайной составляющей является

$$\left(\sigma^*\right)^2 = \frac{SS}{n} - \frac{1}{n}\sum_{k=1}^{n}\left[f^*(t_k)\right]^2 = 378\,267{,}83 - 229\,537 = 148\,731,$$

а оценкой среднего квадратичного отклонения

$$\sigma^* = \sqrt{148731} = 385{,}7.$$

В соответствии с формулами (5.21) и (5.22) оценим дисперсии оценок параметров:

$$D^*(a^*) = \sum_{k=1}^{n} c_k^2 D^*(E_k) = \frac{(\sigma^*)^2}{\sum_{k=1}^{n}(t_k - \bar{t})^2} = \frac{143\,731}{143} = 1040;$$

$$D^*(d^*) = \frac{(\sigma^*)^2}{n} = \frac{143\,731}{12} = 12\,394.$$

Средние квадратичные отклонения a^* и d^* оцениваются как 32,25 и 111,33 соответственно, а доверительные интервалы, соответствующие доверительной вероятности 0,95, таковы:

$$[a_{\min};\ a_{\max}] = [149{,}05;\ 275{,}47],\ \ [d_{\min};\ d_{\max}] = [3748{,}96;\ 4185{,}38].$$

Первое из условий (5.24) выполнено в силу построения оценок периодической составляющей по целому числу периодов. Действительно, согласно данным табл. 5.9 сумма оценок периодической составляющей для 12 точек наблюдений равна (–3), незначительное отклонение от 0 вызвано ошибками округления.

В соответствии с формулой (5.22) смещение оценки a^* оценивается как

$$\sum_{k=1}^{n} c_k f^*(t_k) = \frac{\sum_{k=1}^{n}(t_k - \bar{t})f^*(t_k)}{\sum_{k=1}^{n}(t_k - \bar{t})^2} = \frac{5568}{143} = 38{,}94.$$

Таким образом, смещение имеет тот же порядок, что и среднее квадратичное отклонение оценки a^*, и заведомо меньше, чем полуширина доверительного интервала. Дальнейшее сравнение может быть проведено на основе оценки дисперсии смещения — случайной величины

$$Z = \frac{\sum_{k=1}^{n}(t_k - \overline{t})f^*(t_k)}{\sum_{k=1}^{n}(t_k - \overline{t})^2}.$$

Алгоритм вычисления дисперсии Z аналогичен таковым для периодической составляющей и интервального прогноза (см. формулы (5.33) и (5.35) соответственно), но более сложен, поэтому не включен в учебник.

Таким образом, можно считать, что предположения формулы (5.24) модели, описанной формулой (5.19), выполнены для данных табл. 5.8.

Перейдем к оценке дисперсий значений периодической составляющей. Как следует из равенства (5.33),

$$D(g_s^*) = \sigma^2 \sum_{k=1}^{n} h_{ks}^2, \quad s = 1, 2, \ldots, q,$$

где $h_{ks} = -c_k r_s - \frac{1}{n} + \frac{1}{m}$, если $k \in \{s + (j-1)q, j = 1, 2, \ldots, m\}$, и $h_{ks} = -c_k r_s - \frac{1}{n}$ при иных значениях индекса суммирования k, и $r_s = \frac{1}{m}\sum_{j=1}^{m}(t_{s+(j-1)q} - \overline{t})$.

Начнем со значения $s = 1$ (периодическая составляющая для января). Тогда $r_1 = \frac{1}{3}\left[(1-6{,}5)+(5-6{,}5)+(9-6{,}5)\right] = -1{,}5$. Понадобятся значения

$$c_k = \frac{t_k - \overline{t}}{\sum_{k=1}^{n}(t_k - \overline{t})^2} = \frac{t_k - 6{,}5}{143} = \frac{k - 6{,}5}{143}.$$

Расчет удобно проводить с помощью таблицы (табл. 5.10).

В таблице 5.10 столбец (3) получен из столбца (2) умножением на $\frac{r_1}{143} = \frac{-1{,}5}{143} = -0{,}01049$, каждый элемент столбца (6) равен сумме элементов столбцов (3), (4) и (5), стоящих в той же строке, а в столбце (7) стоят квадраты соседних элементов из столбца (6). Цель построения табл. 5.10 — расчет суммы элементов столбца (7). Эта сумма равна 0,28275. Следовательно,

$$\sqrt{D^*(g_1^*)} = \sigma^* \sqrt{\sum_{k=1}^{n} h_{k1}^2} = 385{,}7\sqrt{0{,}28275} = 204{,}8.$$

Таблица 5.10

Расчет дисперсии периодической составляющей

Номер месяца в году	Параметр					
k	$t_k - \overline{t}$	$c_k r_1$	$-1/n$	$+1/m$	h_{k1}	h_{k1}^2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	–5,5	0,0577	–0,0833	0,3333	0,3077	0,09468
2	–4,5	0,0472	–0,0833	—	–0,0361	0,00130
3	–3,5	0,0367	–0,0833	—	–0,0466	0,00217
4	–2,5	0,0262	–0,0833	—	–0,0571	0,00326
5	–1,5	0,0157	–0,0833	0,3333	0,2657	0,07060
6	–0,5	0,0052	–0,0833	—	–0,0781	0,00610
7	0,5	–0,0052	–0,0833	—	–0,0885	0,00783
8	1,5	–0,0157	–0,0833	—	–0,0990	0,00980
9	2,5	–0,0262	–0,0833	0,3333	0,2238	0,05009
10	3,5	–0,0367	–0,0833	—	0,1200	0,01440
11	4,5	–0,0472	–0,0833	—	0,1305	0,01703
12	5,5	–0,0577	–0,0833	—	0,1410	0,01988

Доверительный интервал для значения периодической составляющей в январе (–376 – 1,96 × 204,8; –376 + 1,96 × 204,8) захватывает 0 (при доверительной вероятности 0,95), отличие значения периодической составляющей от 0 незначимо (на уровне значимости 0,05).

Аналогичный случай для значения $s = 2$ (периодическая составляющая для апреля) дает

$$\sum_{k=1}^{n} h_{k2}^2 = 0{,}25524, \quad \sqrt{D^*(g_2^*)} = \sigma^* \sqrt{\sum_{k=1}^{n} h_{k2}^2} = 385{,}7\sqrt{0{,}25524} = 194{,}86.$$

Доверительный интервал для значения периодической составляющей в апреле (533 – 1,96 × 194,86; 533 + 1,96 × 194,86) = = (533 – 381,93; 533 + 381,93) не захватывает 0 (при доверительной вероятности 0,95), отличие значения периодической составляющей от 0 значимо (на уровне значимости 0,05).

Приступим к завершающему этапу анализа данных табл. 5.8 — построению интервального прогноза. Необходимо рассчитать величины $w_{ks} = c_k(t - \overline{t} - r_s) + \frac{1}{m}$, если $k \in \{s + (j-1)q, j = 1, 2, \ldots, m\}$, и $w_{ks} = c_k(t - \overline{t} - r_s)$ при всех остальных значениях индекса суммирования k, где r_s — то же, что и в формуле (5.33), поскольку точечный прогноз $x^*(t)$ является несмещенным, асимптотически нормальным, а его дисперсия оценивается, согласно формуле (5.35), так:

$$D^*\left[x^*(t)\right] = (\sigma^*)^2 \sum_{k=1}^{n} w_{ks}^2.$$

Начнем с прогноза на январь 2006 г. (по данным за 2003—2005 гг.). Тогда $t = 13$, $s = 1$, $r_1 = -1{,}5$, $w_{k1} = 8c_k + \frac{1}{3}$, если $k \in \{1 + 4(j-1), j = 1, 2, 3\}$, и $w_{k1} = 8c_k$ при всех остальных значениях индекса суммирования. При этом

$$8c_k = 8\frac{k-6{,}5}{143} = \frac{8k-52}{143}.$$

Расчет удобно проводить с помощью таблицы (табл. 5.11).

Таблица 5.11

Расчет дисперсии прогностической функции

Номер месяца в году	Параметр			
k	$\frac{8k-52}{143}$	$1/m$	w_{k1}	w_{k1}^2
1	–0,3077	0,3333	0,0256	0,00066
2	–0,2517	—	–0,2517	0,06336
3	–0,1958	—	–0,1958	0,03834
4	–0,1399	—	–0,1399	0,01957
5	–0,0839	0,3333	0,2494	0,06220
6	–0,0280	—	–0,0280	0,00078
7	0,0280	—	0,0280	0,00078
8	0,0839	—	0,0839	0,00700
9	0,1399	0,3333	0,4732	0,22392
10	0,1958	—	0,1958	0,03834
11	0,2517	—	0,2517	0,06336
12	0,3077	—	0,3077	0,09468

Сумма значений, стоящих в последнем столбце табл. 5.11, равна 0,61299. Согласно формуле (5.36)

$$\Delta(13) = U(0{,}95)\sqrt{D^*(x^*(13))} =$$
$$= 1{,}96 \times 385{,}7\sqrt{0{,}61299} = 591{,}88.$$

Согласно формуле (5.31) точечный прогноз таков:

$$x^*(13) = a^*(13 - \bar{t}) + d^* + f^*(13) =$$
$$= 212{,}26 \times 13 + 2587{,}48 + (-376) = 4971.$$

Нижняя и верхняя доверительные границы для прогностической функции (с учетом как трендовой, так и периодической составляющих) имеют вид:

$$x_{\text{нижн}}(13) = 4971 - 592 = 4379, \;\; x_{\text{верх}}(13) = 4971 + 592 = 5563.$$

Реальное значение (см. табл. 5.7) — 4336. Оно практически совпадает с прогнозным значением $x_{\text{нижн}}(13)$. Прогноз оправдался.

Аналогичные расчеты для апреля 2006 г. ($t = 14, s = 2, r_2 = -0{,}5$) дают

$$\Delta(14) = 1{,}96 \times 385{,}7\sqrt{0{,}72480} = 643{,}60.$$

Точечный прогноз равен $x^*(14) = 6092$, а нижняя и верхняя доверительные границы таковы: $x_{\text{нижн}}(14) = 5448$, $x_{\text{верх}}(14) = 6736$. Реальное значение (см. табл. 5.7) — 5430. Оно практически совпадает с прогнозным значением $x_{\text{нижн}}(14)$. Как и в предыдущем случае, прогноз оправдался.

Метод наименьших квадратов — мощный инструмент организационно-экономического моделирования. Этот раздел эконометрики приносит ощутимую пользу экономистам и управленцам (менеджерам).

Контрольные вопросы

1. Как в методе наименьших квадратов используются преобразования переменных?
2. Как связаны коэффициент линейной корреляции Пирсона и непараметрический коэффициент ранговой корреляции Спирмена?
3. Как метод наименьших квадратов используется для прогнозирования цен в отрасли лома черных металлов?
4. Как метод линеаризации позволяет построить доверительный интервал для точки пересечения двух регрессионных прямых?
5. Сравните модели порождения данных при наличии периодической составляющей (подраздел 5.7) и без таковой (подраздел 5.1). Что при расчетах по методу наименьших квадратов является общим и в чем проявляется различие?

Темы докладов и рефератов

1. Практическое использование метода наименьших квадратов.
2. Критерии качества регрессионной модели.

3. Доказательство теоремы о предельном геометрическом распределении первого локального минимума остаточной дисперсии как оценки степени многочлена, описывающего зависимость.
4. Состоятельные оценки степени многочлена, описывающего регрессионную зависимость.
5. Использование непараметрических оценок плотности для восстановления зависимости.
6. Статистические методы прогнозирования и роль в них метода наименьших квадратов (на основе [84]).
7. Методы выявления информативного подмножества признаков в регрессионном анализе (на основе [111]).
8. Варианты метода наименьших квадратов в нелинейных (по параметрам) моделях.
9. Применение матричной алгебры в линейном регрессионном анализе.

ГЛАВА 6
АНАЛИЗ ДИНАМИКИ ЦЕН И ИСПОЛЬЗОВАНИЕ ИНДЕКСОВ ИНФЛЯЦИИ ПРИ ПРИНЯТИИ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ

6.1. ОПРЕДЕЛЕНИЕ И РАСЧЕТ ИНДЕКСА ИНФЛЯЦИИ

Каждый день мы встречаемся с такими экономическими величинами, как цены на товары и услуги. Как правило, они изменяются с течением времени. Вполне естественно подвергнуть динамику цен на товары и услуги анализу с позиций организационно-экономического моделирования и эконометрики.

Под инфляцией в настоящей главе, как и в учебнике [89], понимаем повсеместно наблюдаемый рост цен. Как следствие, покупательная способность денежных единиц (рублей, евро, долларов США и др.) падает. Следовательно, при анализе экономических процессов, протяженных во времени, при сравнении стоимостных характеристик 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. Продемонстрируем свойства индекса инфляции и алгоритмы его расчета на примере минимальной потребительской корзины продовольственных товаров, которая была разработана в Институте высоких статистических технологий и эконометрики на основе физиологических норм потребления. Затем разберем различные применения индексов инфляции в экономических расчетах.

Краткая история инфляции в России на рубеже тысячелетий. Цены на продовольственные товары (хлеб, молоко и т.п.) в СССР определялись государственными органами и не менялись в течение десятилетий — с начала 1960-х гг. Конец этого периода — 2 апреля 1991 г., когда постановлением Правительства СССР цены на основные потребительские товары были подняты в 2—3 раза. По сообщению Государственного комитета по статистике, к концу 1991 г. (к моменту развала СССР) цены выросли в 2,6 раза. Со 2 января 1992 г. началась — уже в Российской Федерации — так называемая «либерализация цен», в ходе которой торговые организации стали самостоятельно устанавливать цены. В результате за год цены выросли в 26,1 раз. С тех пор рост цен не прекращался. К июлю 2007 г. цены выросли примерно в 60 000 раз, а к декабрю 2008 г. — в 100 000 раз. Однако в январе 1998 г. была проведена так называемая «деноминация», во всех записях стоимостных характеристик были отброшены три нуля, т.е. цены (и доходы) формально уменьшились в 1000 раз. Как следствие, цены июля 2007 г. в 60 раз превышают цены марта 1991 г. (и предыдущих лет), а цены декабря 2008 г. — соответственно в 100 раз.

Итак, цены к июлю 2007 г. выросли в 60 раз, т.е. покупательная способность рубля сократилась в 60 раз. Другими словами, 60 руб. июля 2007 г. соответствуют 1 руб. 1990 г. Это простое соотношение позволяет сопоставлять доходы и расходы, разделенные 17 годами.

Замечание. С течением времени любая конкретная дата уходит в прошлое, а вместе с ней и все конкретные численные значения экономических величин. Примем для сравнения цен март 1991 г. и июль 2007 г. Первая из этих дат — конец стабильности цен в СССР, вторая — начало мирового экономического кризиса, прервавшего рост российской экономики в период стабильности (1999—2007). Читатель сможет перейти к интересующему его моменту с помощью эконометрических методов, разобранных в настоящей главе.

Пример приведения к сопоставимым ценам. Рассмотрим часто обсуждаемый экономический показатель — среднюю заработную плату в России. По данным государственных статистических органов, средняя заработная плата в 1990 г. составляла 303 руб., а в апреле 2007 г. — 12 510 руб. (рост в 41,29 раза). Однако цены выросли в 60 раз, поэтому нынешние 12 510 руб. соответствуют 12 510/60 = 208,50 руб. 1990 г., т.е. в настоящее время реальная средняя заработная плата составляет 208,5/303 × 100% = 68,81% от уровня 1990 г., другими словами, сократилась в 1,45 раза. При этом способе расчета мы привели данные 2007 г. к сопоставимым ценам 1990 г.

Можно поступить и другим образом — привести данные 1990 г. к сопоставимым ценам 2007 г. А именно, если проиндексировать заработную плату 1990 г., т.е. умножить ее на индекс инфляции, показывающий рост цен (в рассматриваемом случае — на 60), то получим 303 × 60 = 18 180 руб. — такой должна была бы быть средняя заработная плата в 2007 г., если бы она росла теми же темпами, что и цены. Проиндексированная заработная плата в 18 180/12 510 = 1,45 раза выше реально начисленной, т.е. реальная средняя заработная плата уменьшилась за 17 лет в 1,45 раз, как и было получено при предыдущем расчете.

Рост цен для различных товаров и услуг. Цены на те или иные товары и услуги растут с различной скоростью. Например, в табл. 6.1 приведены данные о ценах на несколько видов товаров и услуг.

Таблица 6.1

Примеры цен в рублях товаров и услуг в 1990 и 2007 гг. (г. Москва)

Товар (услуга)	Цена		Рост цен, раз
	1990 г.	2007 г.	
Одна поездка в метро	0,05	17	340
Батон белого хлеба «Нарезной»	0,13	9,5—11,8	73—91
Газета	0,03	2,0—6,6	67—220
Водка среднего качества (0,5 л)	10,00	80—120	8—12
Электроэнергия (1 кВт·ч)	0,04	2,08	52

Наблюдаем значительное различие в темпах роста цен — в десятки раз. Поэтому для получения сводного показателя необходимо усреднять темпы роста цен для отдельных товаров и услуг. Этот факт подчеркивает и названия рассматриваемых в настоящей главе показателей — индексы инфляции, индексы потребительских цен. Введем используемые в дальнейшем понятия.

Индексы и их применение. Индекс (лат. *index* — показатель, список) — это 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 т.п. Индексы делятся на индивидуальные и сводные. Так, индивидуальный динамический индекс описывает изменение тех или иных явлений во времени, например изменение 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» (обратите внимание на то, что в статистических методах, согласно традиции, термин «индекс» может использоваться во многих разных смыслах). Таким образом, индексы зависят от двух переменных — индексируемой величины x и весов индексов f.

Определение понятия «индекс инфляции». В качестве примера построения и использования индексов рассмотрим индекс потребительских цен, он же — индекс инфляции.

Цены на различные товары меняются по-разному. Как усреднить темпы роста цен? Одна из основных проблем в современной экономике — проблема агрегирования с целью сжатия информации (см., например, монографию [88]). Как свести к одной величине темпы роста цен различных товаров и услуг?

Уровень цен выражается в виде индекса. Он является измерителем соотношения между совокупной ценой определенного набора товаров, называемого «рыночной корзиной» (или «потребительской корзиной»), для данного (текущего) момента времени и совокупной ценой идентичной либо сходной группы товаров в базовый момент времени.

Первое, что приходит в голову, — усреднить индексы для отдельных товаров и услуг. Но какое среднее взять? Среднее арифметическое? Среднее геометрическое? Среднее гармоническое? Среднее квадратичное? В экономике используется много различных видов средних (см., главу 9). Опишем наиболее распространенный подход.

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 число 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:

$$p_i = p_i(t),\ i = 1, 2, ..., n$$

(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. $\{Q_1(t), Q_2(t), ..., Q_n(t)\} \equiv \{Q_1, Q_2, ..., Q_n\}$. Стоимость $S(t)$ потребительской корзины в момент времени t такова:

$$S(t) = \sum_{1 \le i \le n} p_i(t) Q_i.$$

Затем необходимо сравнить стоимости $S(t_1)$ и $S(t_2)$ потребительской корзины $\{Q_1, Q_2, ..., Q_n\}$ в старых $p_i(t_1)$, $i = 1, 2, ..., n$, и новых $p_i(t_2)$, $i = 1, 2, ..., n$, ценах.

Индексом инфляции называется функция

$$I(t_1, t_2) = \frac{S(t_2)}{S(t_1)} = \frac{\sum_{1 \le i \le n} p_i(t_2) Q_i}{\sum_{1 \le i \le n} p_i(t_1) Q_i}.$$

Здесь индексируемая величина — цена, а весами служат объемы потребления, зафиксированные в принятой исследователем потребительской корзине.

С математической точки зрения индекс инфляции — это функция двух переменных, а именно двух моментов времени — начального, или базового, момента t_1 и конечного, или текущего, момента t_2. Когда говорят об инфляции за определенный промежуток времени, то t_1 — начало этого промежутка (года, месяца), а t_2 — его конец. Обычно $t_1 < t_2$, хотя в приведенном определении это не требуется.

Подчеркнем, что каждой конкретной потребительской корзине соответствует свой индекс инфляции. Потребительская корзина — это *инструмент экономиста*, предназначенный для усреднения индивидуальных индексов инфляции

$$I_i(t_1, t_2) \frac{p_i(t_2)}{p_i(t_1)},\ i = 1, 2, \ldots, n,$$

т.е. темпов роста цен отдельных товаров (услуг). Потребительская корзина не имеет отношения к реальному потреблению экономического субъекта. В частности, структура реального потребления в соответствии с законом Энгеля меняется в зависимости от дохода этого субъекта, в то время как потребительская корзина, используемая для расчета индекса инфляции, зафиксирована и никак не связана с доходом субъекта.

Обсудим подробнее различие понятий «реальное потребление» и «фиксированная потребительская корзина, используемая при расчете индекса инфляции». Расходы C на покупки товаров и услуг некоторого экономического субъекта следует сопоставлять с его доходом D. Если $D - C > 0$, то экономическое положение субъекта благоприятно, его доход больше расходов. Если же $D - C < 0$, то его положение неблагоприятно, доход меньше расходов. Это означает, что он расходует ранее накопленные средства, делает долги (в частности, берет кредиты) и т.д.

Из сказанного вытекает, что величина расходов $C(t)$ обязательно регулируется экономическим субъектом в соответствии с его доходом. Изменение (рост) цен $p_i(t)$ с течением времени t делает невозможным сохранение прежней структуры потребления $\{Q_1, Q_2, ..., Q_n\}$, если рост дохода отстает от роста цен. Структура потребления изменяется, сокращается потребление относительно дорогих товаров и услуг, в порядке компенсации увеличивается потребление относительно дешевых. Например, уменьшается потребление мяса и увеличивается — хлеба и картофеля. При быстром росте цен возможен и другой эффект — «бегство от рубля». В связи с обесцениванием сбережений экономические субъекты направляют доход на текущее потребление ценой отказа от накопления средств на приобретение дорогостоящих товаров длительного пользования.

Таким образом, реальное потребление товаров и услуг определяется как действующими ценами, так и величиной доходов экономических субъектов. Чтобы измерять рост цен, нужно избавиться от влияния изменения доходов. Именно для этого зафиксирована потребительская корзина, используемая для измерения инфляции.

Замечание. Приведем выписку из «Экономического словаря» (URL: http://abc.informbureau.com): «Закон Энгеля — зависимость доли расходов на продукты питания в доходах семьи от их уровня, установленная в XIX в. немецким статистиком и экономистом Эрнстом Энгелем (1821—1896)». Согласно этому закону по мере роста доходов семьи падает доля расходов на продовольствие, почти не меняется удельный вес затрат на жилище, отопление, освещение, одежду; зато растет доля расходов на прочие нужды (прежде всего на сбережения). Выведенная Энгелем эмпирическая зависимость подтверждается длительным опытом экономического развития. Статистические ряды показывают, что в структуре расходов американских семей за 1909—1985 гг. устойчиво снижается доля физических и материально-вещественных потребностей (продуктов питания — на 35%, расходов на жилище — на 5,5%, предметов домашнего обихода — на 26%, расходов на одежду и обувь — на 47%). В то же время систематически возрастает удельный вес «гуманитарных» потребностей — в образовании, медицинском обслуживании, организации досуга и отдыха. Об этом же свидетельствует и статистика семейных бюджетов в бывшем СССР (табл. 6.2).

Таблица 6.2

Расходы на питание в СССР (1963 г.)

Совокупный доход в месяц, руб.	Расходы на питание, %
До 75	51,5
75—100	42,7
100—150	35,8
150—200	31,9
Свыше 200	28,4
Все семьи	34,3

Современная западная наука существенно дополнила развитые Энгелем положения, используя для этого новый аналитический аппарат. На основе формулы эластичности спроса по доходам подсчитано, что в конце 1980-х гг. в США с ростом дохода на 1% спрос на продукты питания возрастал на 0,77%, на одежду — на 0,32%, на транспортные средства — на 1,1%, на жилище — на 0,89%, на медицинские услуги — на 1,9%, на предметы роскоши — на 3,6%, на спортивные товары — на 3,7%, на услуги такси — на 2,8%».

Разброс цен в пространстве. В определении индекса инфляции участвуют цены $p_i(t)$. Однако цены меняются при переходе от одной торговой точки к другой. Это отражено и в табл. 6.1. Два полностью идентичных батона хлеба могут продаваться по разной цене даже в соседних магазинах. Обсудим эффект разброса цен в пространстве и его учет при расчете индекса инфляции.

В конкретном акте купли-продажи цена товара или услуги полностью определена. Однако в современных условиях, когда в большинстве случаев продавец, а иногда и покупатель могут влиять на цену товара или услуги, эта цена зачастую меняется от одного акта купли-продажи к другому. Можно выделить несколько вариантов.

1. Конкретный продавец меняет цену в зависимости от поведения конкретного покупателя. Пример: индивидуальный продавец на базаре.

2. В конкретном магазине цена фиксированна, но от магазина к магазину она меняется. Пример: большинство товаров (продаваемых в магазинах и киосках), цены которых указаны для сведения покупателей.

3. Единые цены в регионе, например на электроэнергию, услуги транспорта и почтовой связи.

В первом и втором случаях имеет место разброс цен на однотипный товар. Этот эффект проявляется как и нашей стране, так и за рубежом. Например, в современной Франции цены на определенный товар в фешенебельных центральных магазинах и в окраинных непрестижных супермаркетах могут отличаться в несколько раз. Это одно из проявлений так называемой «ценовой диверсификации».

Какие же цены использовать при расчете индекса инфляции? Возможны два подхода к проведению организационно-экономического исследования — средней цены и фиксированного маршрута.

Подход на основе средней цены предполагает проведение обширного статистического исследования, позволяющего с достаточной степенью точности установить распределение цены определенного товара (рассматриваемой как случайная величина). По распределению рассчитывается средняя цена. Нужные данные получают, например, в ходе проводимого органами Росстата бюджетного обследования нескольких тысяч семей, при котором ежедневно фиксируют все их расходы. Тогда достаточно получить среднее арифметическое всех цен при покупках рассматриваемого товара, осуществленных в определенный день, взвешивая их по объему покупок. В другом варианте планирования исследования на основе анализа расходов семей устанавливают доли покупок (по объему), приходящиеся на торговые организации различных типов (базары, магазины, супермаркеты и т.п.), а затем специально подготовленные наблюдатели снимают цены, действующие в этих торговых организациях.

Подход, основанный на использовании фиксированного маршрута, предполагает постоянное слежение за ценами в торговых точках, расположенных вдоль раз и навсегда выбранного маршрута (в идеальном варианте — в одном и том же магазине, в котором продаются все товары, включенные в потребительскую корзину). С помощью такой методики измерения удается избавиться от разброса цен в пространстве и сосредоточиться на изучении их динамики во времени. Именно так собирали цены сотрудники Института высоких статистических технологий и эконометрики, на основе исследовательского опыта которого подготовлен учебный материал, представленный в настоящей главе и — несколько подробнее — в учебнике [89, глава 7].

6.2. ПРАКТИЧЕСКИ ИСПОЛЬЗУЕМЫЕ ПОТРЕБИТЕЛЬСКИЕ КОРЗИНЫ И СООТВЕТСТВУЮЩИЕ ИНДЕКСЫ ИНФЛЯЦИИ

Каждый человек, семья (домохозяйство), фирма, выбрав подходящую потребительскую корзину, может без больших трудозатрат оценивать влияние роста цен на свое экономическое положение (подробнее — в подразделе 6.4).

Однако обычно индекс инфляции рассматривают для более или менее обширной совокупности экономических субъектов — для жителей региона или страны, предприятий определенной отрасли и т.д. При этом потребительскую корзину (Q_1, Q_2, ..., Q_n) стараются приблизить к суммарным объемам потребления для рассматриваемой совокупности, а в качестве цен рассматриваются средневзвешенные цены в осуществленных актах купли-продажи.

Например, в макроэкономике используют так называемый «дефлятор валового внутреннего продукта (ВВП)» — индекс цен на все произведенные в течение года конечные товары и услуги, составляющие объем ВВП, используемый для учета влияния инфляции на величину

номинального ВВП. Номинальный ВВП исчисляется в текущих рыночных ценах. Чтобы определить реальную величину ВВП, необходимо выразить его в сопоставимых ценах базисного года. Для этого применяется индекс инфляции, который отражает изменение среднего уровня цен максимально широкой группы товаров и услуг (охватывающей все составляющие ВВП) за рассматриваемый период.

Согласно принятому определению индекс инфляции определяется номенклатурой (т.е. набором, перечнем) товаров и услуг, для которых он вычисляется, объемами потребления и ценами этих товаров и услуг на начальный и на текущий моменты времени. Отсюда вытекает, в частности, что индекс инфляции для продовольственных товаров отличается, вообще говоря, от такового для промышленных товаров, для услуг и от индекса инфляции оптовых цен; индекс инфляции для москвича отличается от такового для жителя Краснодара (хотя бы потому, что климат различается, а потому и объем потребляемой энергии); индекс инфляции для машиностроительной продукции меняется в зависимости от индекса инфляции в строительстве; этот индекс меняется в зависимости от индивидуальных структур потребностей в семьях и т.д. Для адекватного рассматриваемому экономическому субъекту определения индекса инфляции необходимо знать типовые объемы купли-продажи и цены в соответствующих актах купли-продажи, иначе можно говорить только о той или иной оценке этого индекса.

Конкретизация задачи вычисления индекса инфляции. Прежде всего необходимо сформулировать цель экономического анализа роста цен. Будем ориентироваться на положение основной массы населения. Это означает, в частности, что персональные компьютеры (в 2001 г. их имели 6% российских семей) и автомашины иностранных марок (в 2001 г. их имели 0,5% российских семей) в потребительскую корзину, предназначенную для использования на рубеже тысячелетий, включать нецелесообразно.

Как показывают бюджетные исследования, число видов товаров и услуг, потребляемых физическими лицами, измеряется тысячами (а в классификаторах промышленной продукции указаны миллионы марок различных товаров). Поэтому первый шаг — ограничение номенклатуры товаров и услуг, используемых для вычисления индекса инфляции.

На рубеже тысячелетий существенная часть доходов населения России (зачастую не менее половины) идет на покупку продовольственных товаров (что, по классическому закону Энгеля, — см. также учебник нобелевского лауреата по экономике П. Самуэльсона [110] — свидетельствует о сравнительно низком жизненном уровне). Поэтому представляется естественным рассчитать индекс инфляции для продовольственных товаров.

Потребительская корзина Центра экономической конъюнктуры. В первой половине 1990-х гг. Центр экономической конъюнктуры (ЦЭК) при Правительстве Российской Федерации и Государственный комитет Российской Федерации по статистике следили за движением цен по фиксированному набору товаров (табл. 6.3), которые относительно постоянно бывают в магазинах (по различным причинам время от времени этот набор меняется).

Таблица 6.3

Объемы годового потребления продовольственных товаров (потребительская корзина Центра экономической конъюнктуры в сравнении с компонентами потребительской корзины Института высоких статистических технологий и эконометрики — ИВСТЭ)

№ п/п	Продукты питания, кг	Потребительские корзины	
		ЦЭК	ИВСТЭ
1	Хлеб ржаной	92,0	65,3
2	Хлеб пшеничный	86,7	59,8
3	Пшено	18,1	4,9
4	Вермишель	7,3	4,9
5	Сахар	24,8	19,0
6	Масло растительное, л	10,0	3,8
7	Масло животное	3,6	2,5
8	Говядина	42,0	4,4
9	Колбаса вареная	2,2	0,7
10	Колбаса полукопченая	1,1	0,7
11	Молоко, л	184,3	110,0
12	Сметана	4,2	1,6
13	Сыр твердый	2,0	2,3
14	Яйца, шт.	183,0	152,0
15	Картофель	146,0	124,2
16	Капуста свежая	29,8	30,4
17	Лук репчатый	10,2	27,9
18	Яблоки	11,0	15,1
19	Сигареты, пачка	96,0	—

Потребительская корзина Института высоких статистических технологий и эконометрики (ИВСТЭ) на основе данных Института питания Российской академии медицинских наук (РАМН). Приведенная в табл. 6.3 потребительская корзина ЦЭК не полностью соответствует перечню продуктов питания, рекомендованному медиками. И дело не только в сигаретах.

Рассмотрим минимальную потребительскую корзину физиологически необходимых продовольственных товаров, разработанную в 1993 г. ИВСТЭ на основе исходных данных Института питания РАМН. Эти данные использовались также Министерством труда Российской Федерации. Рассматриваемую минимальную потребительскую корзину обозначим сокращенно «корзина ИВСТЭ». В отличие от приведенной ранее корзины ЦЭК в ней содержание белков, жиров и углеводов соответствует (минимальным) медицинским нормам. В корзине ИВСТЭ продукты питания разделены на 11 групп:

1) хлеб и хлебопродукты;
2) картофель;
3) овощи;
4) фрукты и ягоды;
5) сахар;
6) мясопродукты;
7) рыба и рыбопродукты;
8) молоко и молочные продукты;
9) яйца;
10) масло растительное и маргарин;
11) прочие.

Общая стоимость «прочих» видов продуктов составляет до 6% стоимости первых 10 групп продуктов данной потребительской корзины.

На основе физиологических норм потребления Института питания РАМН в ИВСТЭ составлена минимальная потребительская корзина, т.е. указан годовой объем потребления по основным продовольственным товарам, необходимый для поддержания нормальной жизнедеятельности человеческого организма (табл. 6.4). При разработке корзины специалисты Института питания исходили из трех принципов:

1) суммарное содержание белков, жиров, углеводов и калорий должно быть не менее нормативов, определяющих, согласно науке о питании (как части медицины), возможность продолжения существования человеческого организма без физиологического вырождения;
2) на основе включенных в корзину продуктов может быть разработано меню ежедневного трехразового питания на год;
3) стоимость корзины должна быть минимальна.

Первый и третий принципы позволяют сформулировать задачу оптимизации (линейного программирования). Ее решение таково (в расчете на день): 812 г черного хлеба, 705 г картофеля, 180 г молока и 10 г сыра. Хотя этот набор продуктов обеспечивает необходимое количество белков, жиров, углеводов и калорий, ежедневно питаться таким образом невозможно. Второй принцип обеспечивает человека полноценным трехразовым питанием. Но стоимость корзины возрастает примерно на четверть.

Потребительская корзина, представленная в табл. 6.4, не описывает реальное потребление большинства граждан. Например, типовой москвич покупает значительно больше колбасы, сала, копченостей, чем включено в корзину, и в несколько раз меньше муки. Корзина (см. табл. 6.4) предназначена прежде всего для измерения инфляции. Однако еще одно ее использование — оценка (снизу) минимально допустимых расходов на продовольственные товары, обеспечивающих нормальную жизнедеятельность человеческого организма. Таковы расходы в некоторых закрытых учреждениях — больницах, тюрьмах, приютах, домах престарелых.

Таблица 6.4

Номенклатура, годовые нормы потребления и цены для потребительской корзины ИВСТЭ

Продукт питания	Годовая норма, кг	Цена на 14.03.1991 г., руб.	Цена на 14.03.1994 г., руб.	Рост цен, раз
1. Хлеб и хлебопродукты				
1.1. Мука пшеничная	18,5	0,46	646,00	1 404
1.2. Рис	3,5	0,88	620,00	705
1.3. Другие крупы («Геркулес»)	4,9	0,62	750,00	1 210
1.4. Хлеб пшеничный	59,8	0,50	720,00	1 440
1.5. Хлеб ржаной	65,3	0,20	390,00	1 950
1.6. Макаронные изделия	4,9	0,70	1 200,00	1 714
2. Картофель	124,2	0,10	490,00	4 900
3. Овощи				
3.1. Капуста	30,4	0,20	500,00	2 500
3.2. Огурцы и помидоры	2,8	0,85	2 500,00	2 941
3.3. Столовые корнеплоды	40,6	0,20	450,00	2 250
3.4. Прочие (лук и др.)	27,9	0,50	900,00	1 800
4. Фрукты и ягоды				
4.1. Яблоки свежие	15,1	1,50	960,00	640
4.2. Яблоки сушеные	1,0	3,00	1 900,00	633
5. Сахар и кондитерские изделия				
5.1. Сахар	19,0	0,90	650,00	722
5.2. Конфеты	0,8	4,50	3 500,00	778
5.3. Печенье и торты	1,2	1,40	14 700,00	3 357
6. Мясо и мясопродукты				
6.1. Говядина	4,4	2,00	2 700,00	1 350
6.2. Баранина	0,8	1,80	1 940,00	1 078
6.3. Свинина	1,4	2,00	2 300,00	1 150
6.4. Субпродукты (печень)	0,5	1,40	3 500,00	2 500
6.5. Птица	16,1	2,40	2 600,00	1 083
6.6. Сало	0,7	2,40	3 300,00	1 375
6.7. Копчености	0,7	3,70	15 000,00	4 054
7. Рыба и рыбопродукты				
7.1. Свежая (минтай)	10,9	0,37	2 200,00	5 946
7.2. Сельди	0,8	1,40	2 500,00	1 786
8. Молоко и молочные продукты				
8.1. Молоко, кефир, л	110,0	0,32	520,00	1 625
8.2. Сметана, сливки	1,6	1,70	2 500,00	1 471
8.3. Масло животное	2,5	3,60	4 000,00	1 111
8.4. Творог	9,8	1,00	2 000,00	2 000
8.5. Сыр и брынза	2,3	3,60	6 000,00	1 667
9. Яйца, шт.	152,0	0,09	100,00	1 111
10. Масло растительное, маргарин				
10.1. Масло растительное, л	3,8	1,80	2 000,00	1 111
10.2. Маргарин	6,3	1,20	2 000,00	1 667
11. Прочие (6% от стоимости товаров групп 1—10)	—	—	—	—

Расчет стоимости минимальной потребительской корзины продовольственных товаров. Чтобы получить индекс инфляции, рассчитаем стоимость минимальной потребительской корзины продовольственных товаров, исходя из объемов потребления, заданных в разработках Института питания, и цен по состоянию на март 1991 г. (т.е. до первого значительного повышения цен в апреле 1991 г. и их «либерализации» в январе 1992 г.) и — в качестве примера — на март 1994 г. (очевидно, расчеты могут быть проведены и на любой иной момент времени) с целью установить динамику цен за полные три года.

Из таблицы 6.4 следует, что темпы роста цен на различные продукты питания существенно отличаются друг от друга. Минимальный рост цен — в 633 раза (яблоки сушеные), максимальный — в 5946 раз (минтай).

Для нахождения расходов на определенные продукты питания (в расчете на год) достаточно умножить цены на объемы потребления, как это сделано в табл. 6.5. Там же приведены годовые расходы для каждой из 11 товарных групп.

Таблица 6.5

Годовые расходы на покупку продуктов, в рублях

Наименование продуктов питания и групп продуктов	Годовые расходы по ценам на 14.03.1991 г.	Годовые расходы по ценам на 14.03.1994 г.
1. Хлеб и хлебопродукты, всего	61,02	92 199
1.1. Мука пшеничная	8,51	11 951
1.2. Рис	3,08	2 170
1.3. Прочие крупы («Геркулес»)	3,04	3 675
1.4. Хлеб пшеничный	29,90	43 056
1.5. Хлеб ржаной	13,06	25 467
1.6. Макаронные изделия	3,43	5 880
2. Картофель	12,42	60 858
3. Овощи, всего	30,53	65 580
3.1. Капуста	6,08	15 200
3.2. Огурцы и помидоры	2,38	7 000
3.3. Столовые корнеплоды	8,12	18 270
3.4. Прочие (лук и др.)	13,95	25 110
4. Фрукты и ягоды, всего	25,65	16 396
4.1. Яблоки свежие	22,65	14 496
4.2. Яблоки сушеные	3,00	1 900
5. Сахар и кондитерские изделия, всего	22,38	20 790
5.1. Сахар	17,10	12 350
5.2. Конфеты	3,60	2 800
5.3. Печенье и торты	1,68	5 640
6. Мясо и мясопродукты, всего	56,65	73 072
6.1. Говядина	8,80	11 880
6.2. Баранина	1,44	1 552
6.3. Свинина	2,80	3 220

Окончание

Наименование продуктов питания и групп продуктов	Годовые расходы по ценам на 14.03.1991 г.	Годовые расходы по ценам на 14.03.1994 г.
6.4. Субпродукты (печень)	0,70	1 750
6.5. Птица	38,64	41 860
6.6. Сало	1,68	2 310
6.7. Копчености	2,59	10 500
7. Рыба и рыбопродукты, всего	5,15	25 980
7.1. Свежая (минтай)	4,03	23 980
7.2. Сельди	1,12	2 000
8. Молоко и молочные продукты, всего	65,00	104 600
8.1. Молоко, кефир	35,20	57 200
8.2. Сметана, сливки	2,72	4 000
8.3. Масло животное	9,00	10 000
8.4. Творог	9,80	19 600
8.5. Сыр и брынза	8,28	13 800
9. Яйца	13,68	15 200
10. Масло растительное, маргарин	14,40	20 200
10.1. Масло растительное	6,84	7 600
10.2. Маргарин	7,56	12 600
Всего по 10 группам:	306,89	490 675
11. Прочие (6% от стоимости товаров групп 1—10)	18,41	29 441
Суммарная стоимость минимальной потребительской корзины продуктов питания в расчете:		
на год	325,30	520 116
на месяц	27,11	43 499

Как следует из табл. 6.5, индекс инфляции (роста цен) по продуктам питания, исходя из минимальной потребительской корзины ИВСТЭ, составленной по физиологическим нормам потребления продуктов питания для г. Москвы (согласно разработкам Института питания РАМН и Министерства труда Российской Федерации), за 3 года (14.03.1991—14.03.1994) составил (520 116) / (325,3) = 1598,88 или 159 788%.

Таблицы, подобные приведенным ранее табл. 6.4 и 6.5, могут быть составлены любым исследователем (студентом, менеджером или иным заинтересованным гражданином, сотрудником той или иной фирмы, органа власти, профсоюзной организации) в целях изучения динамики реального экономического положения. В таблице 6.6 приведены рассчитанные сотрудниками ИВСТЭ по независимо собранной информации значения стоимости потребительской корзины и индекса инфляции за 1991—2007 гг.

Таблица 6.6

Стоимость потребительской корзины ИВСТЭ и индекс инфляции

№ п/п	Дата снятия цен t	Стоимость потребительской корзины $S(t)$, руб.	Индекс инфляции $I(31.03.91;\ t)$
1	31.03.1991	26,60	1,00
2	14.08.1993	17 691,00	665,08
3	15.11.1993	28 050,00	1 054,51
4	14.03.1994	40 883,00	1 536,95
5	14.04.1994	44 441,00	1 670,71
6	28.04.1994	47 778,00	1 796,17
7	26.05.1994	52 600,00	1 977,44
8	08.09.1994	58 614,00	2 203,53
9	06.10.1994	55 358,00	2 081,13
10	10.11.1994	72 867,00	2 739,36
11	01.12.1994	78 955,00	2 968,23
12	29.12.1994	97 897,00	3 680,34
13	02.02.1995	129 165,00	4 855,83
14	02.03.1995	151 375,00	5 690,79
15	30.03.1995	160 817,00	6 045,75
16	27.04.1995	159 780,00	6 006,77
17	01.06.1995	167 590,00	6 300,38
18	29.06.1995	170 721,00	6 418,08
19	27.07.1995	175 499,00	6 597,71
20	31.08.1995	173 676,00	6 529,17
21	28.09.1995	217 542,00	8 178,27
22	26.10.1995	243 479,00	9 153,35
23	30.11.1995	222 417,00	8 361,54
24	28.12.1995	265 716,00	9 989,32
25	01.02.1996	287 472,55	10 807,24
26	05.03.1996	297 958,00	11 201,43
27	05.04.1996	304 033,44	11 429,83
28	08.05.1996	305 809,55	11 496,60
29	05.06.1996	302 381,69	11 367,73
30	03.07.1996	306 065,21	11 506,21
31	03.08.1996	308 963,42	11 615,17
32	07.09.1996	288 835,07	10 858,46
33	01.10.1996	278 235,35	10 459,98
34	05.11.1996	287 094,77	10 793,04
35	04.12.1996	298 024,76	11 203,94
36	03.01.1997	314 287,16	11 815,31
37	04.02.1997	334 738,24	12 584,14
38	04.01.1998	345,72	12,997
39	03.01.1999	630,07	20,395
40	05.01.2000	737,80	27,737
41	03.01.2001	886,84	33,340
42	03.01.2002	1 051,79	39,541
43	03.01.2003	1 210,62	45,512
44	03.01.2004	1 355,91	50,974
45	14.05.2004	1 369,10	51,470
46	11.01.2005	1 463,98	55,037
47	10.01.2006	1 525,62	57,354
48	26.11.2006	1 571,26	59,070
49	10.01.2007	1 580,89	59,432
50	02.07.2007	1 644,38	61,819
51	30.12.2007	2 286,54	85,960

Окончание

Примечание. Учитывается проведенная 01.01.1998 г. деноминация рубля. Стоимость потребительской корзины приводится без включения группы «прочие».

6.3. СВОЙСТВА ИНДЕКСОВ ИНФЛЯЦИИ

Перед тем как переходить к рассмотрению примеров использования индексов инфляции в экономических расчетах, целесообразно рассмотреть некоторые их свойства.

Соотношение индексов инфляции для трех моментов времени. Рассмотрим три произвольных момента времени t_1, t_2, t_3 и соответствующие индексы инфляции $I(t_1,\ t_2)$, $I(t_2,\ t_3)$ и $I(t_1,\ t_3)$. Из определения индекса инфляции как отношения стоимостей потребительской корзины в соответствующие моменты времени вытекает следующее утверждение.

***Теорема 6.1 (теорема умножения)*.** Для любых трех моментов времени t_1, t_2, t_3 справедливо равенство

$$I(t_1, t_3) = I(t_1, t_2)I(t_2, t_3).$$

Доказательство. По определению индекса инфляции,

$$I(t_1\ t_2) = \frac{S(t_2)}{S(t_1)}; \quad I(t_2\ t_3) = \frac{S(t_3)}{S(t_2)},$$

где $S(t)$ — стоимость потребительской корзины в момент времени t. Следовательно,

$$I(t_1\ t_2)I(t_2\ t_3) = \frac{S(t_2)}{S(t_1)}\frac{S(t_3)}{S(t_2)}.$$

В числителе и знаменателе стоит одно и то же выражение $S(t_2)$. Сократив его, получим

$$I(t_1\ t_2)I(t_2\ t_3) = \frac{S(t_3)}{S(t_1)}.$$

Выражение справа — это, по определению, индекс инфляции $I(t_1, t_3)$. Теорема 6.1 доказана.

Пусть, например, t_1 — это 31 декабря 2004 г., t_2 — это 31 декабря 2005 г., t_3 — это 31 декабря 2006 г. Тогда $I(t_1, t_2)$ — это индекс инфляции за 2005 г., равный 10,9% (официальные данные Росстата). А $I(t_2, t_3)$ — это индекс инфляции за 2006 г., согласно тому же источнику он равен 9%. Теорема умножения дает возможность рассчитать по этим данным индекс инфляции за два года (2005—2006), т.е. с 31 декабря 2004 г. по 31 декабря 2006 г.

Согласно приведенному определению индекс инфляции — действительное число. Если цены постоянны, то индекс инфляции равен 1. Если цены растут, то индекс инфляции больше 1. Однако часто приводят индекс инфляции в процентах. При этом в процентах выражают отклонение от ситуации постоянных цен, т.е. отклонение от 1. Обозначим через $a = a(t_1, t_2)$, или $a\% = a(t_1, t_2)\%$, индекс инфляции в процентах за интервал времени (t_1, t_2). Тогда

$$a(t_1, t_2)\% = [I(t_1, t_2) - 1]100\%; \quad I(t_1, t_2) = 1 + \frac{a(t_1, t_2)}{100}.$$

Или в сокращенной форме:

$$a\% = 100(I - 1)\%; \quad I(t_1, t_2) = 1 + \frac{a\%}{100}.$$

Чтобы перейти к выражению индекса инфляции в процентах, надо значение «в разах» уменьшить на 1 и результат умножить на 100. Наоборот, чтобы от процентов перейти к «разам», надо значение «в процентах» поделить на 100 и результат увеличить на 1.

Таким образом, 1,25 и 25% — это две записи одного и того же значения индекса инфляции. Инфляция 9% за 2006 г. означает, что цены выросли в среднем в 1,09 раза. Рост цен в 1992 г. в 26,1 раз означает, что индекс инфляции за этот год составил (26,1 – 1)100% = 2510%.

Итак, используют два основных способа записи индекса инфляции — в «разах» и в «процентах». В «разах» — это именно тот способ, который дан в определении индекса инфляции как отношения стоимостей потребительской корзины в два момента времени. Однако в средствах массовой информации предпочитают приводить инфляцию в «процентах».

В теореме умножения индексы инфляции выражены «в разах». Следовательно, для расчета индекса инфляции за 2 года надо от «процентов» перейти к «разам». Индекс инфляции за 2005 г. составляет

$$I(t_1\ t_2) = 1 + \frac{10{,}9}{100} = 1{,}109,$$

а индекс инфляции за 2006 г. равен

$$I(t_2\ t_3) = 1 + \frac{9}{100} = 1{,}09.$$

По теореме умножения, индекс инфляции за 2 года таков:

$$I(t_1, t_3) = I(t_1, t_2)I(t_2, t_3) = 1{,}109 \times 1{,}09 = 1{,}20881.$$

Переведем в проценты:

$$I(t_1, t_3) = 100(1{,}20881 - 1) = 20{,}881\%.$$

С достаточной для практики точностью можно округлить: $I(t_1, t_3) = 20{,}9\%$.

Обратите внимание, что при сложении индексов инфляции, выраженных «в процентах», получим 10,9% + 9% = 19,9%, что меньше правильного результата 20,881% почти на 1%. К сожалению, неверная рекомендация о сложении «процентов» (вместо умножения «разов») иногда встречается в литературных источниках.

Теорема умножения позволяет переходить от индексов инфляции за отдельные недели к индексам инфляции за месяц (четыре недели), от помесячных индексов инфляции — к квартальным и годовым, от годовых — к индексам инфляции за несколько лет. Например, индекс инфляции за второй квартал — с 1 апреля 1994 г. по 1 июля 1994 г. — т.е. I(01.04.94, 01.07.94), выражается через индексы инфляции за апрель I(01.04.94, 01.05.94), май I(01.05.94, 01.06.94) и июнь I(01.06.94, 01.07.94) соответственно как произведение этих индексов, т.е. находится по формуле

$$I(01.04.94,\ 01.07.94) =$$
$$= I(01.04.94,\ 01.05.94)\ I(01.05.94,\ 01.06.94)\ I(01.06.94,\ 01.07.94).$$

Аналогично индекс инфляции за год равен произведению двенадцати индексов инфляции за каждый месяц года.

Насколько велика ошибка от сложения индексов инфляции в «процентах»? Найдем ее в общем виде. Поскольку для любых чисел x и y справедливо тождество

$$(1 + x)(1 + y) = 1 + x + y + xy,$$

то, как легко проверить, для индексов инфляции «в процентах»

$$i(t_1, t_2) = 100[I(t_1, t_2) - 1]$$

справедливо тождество

$$i(t_1, t_3) = i(t_1, t_2) + i(t_2, t_3) + \frac{i(t_1, t_2)i(t_2, t_3)}{100}.$$

Если индексы инфляции «в процентах» $i(t_1, t_2)$ и $i(t_2, t_3)$ малы, т.е. индексы инфляции «в разах» $I(t_1, t_2)$ и $I(t_2, t_3)$ мало отличаются от единицы, то справедлива приближенная формула

$$i(t_1, t_3) = i(t_1, t_2) + i(t_2, t_3).$$

Погрешность этой формулы, измеряемая в процентах,

$$\Delta = \frac{i(t_1, t_2)i(t_2, t_3)}{100}.$$

Эта величина становится заметной, если сомножители — порядка десятков и сотен процентов. Если формула применяется несколько раз, то погрешность накапливается. Противоположный случай — при малых индексах инфляции погрешность приближенной формулы мала (является бесконечно малой более высокого порядка).

Период удвоения цен. Рассмотрим пример. В известном учебнике экономической теории [47] рассмотрена связь между ежегодным увеличением цен и числом лет, необходимых для увеличения цен вдвое. Приведено правило, которое вначале выглядит совершенно непонятным:

Приблизительное количество лет, необходимое
для удвоение цен = 70 / (*Темп ежегодного увеличения*
уровня цен в процентах).

Действительно, пусть n — число лет, необходимое для удвоения цен, а x — темп ежегодного увеличения уровня цен (в процентах — $100x\%$). При «подходе профана» рост за n лет составит $100nx$, а потому срок удвоения цен должен находиться из условия

$$100nx = 100, \quad n = 100 / (100x),$$

т.е. в числителе дроби должно стоять число 100, а не 70. В чем дело?

А дело в том, что рост описывается не линейной функцией, а экспоненциальной, надо не складывать, а возводить в степень. За n лет рост цен составит $(1 + x)^n$. Период удвоения находится из уравнения

$$(1 + x)^n\ 2,$$

тогда

$$n = \frac{\ln 2}{\ln(1 + x)}.$$

Воспользуемся приближенной формулой математического анализа

$$\ln(1 + x) = x,$$

тогда с точностью до бесконечно малых более высокого порядка

$$n = 100 \ln 2 / 100\, x.$$

Остается заметить, что

$$100 \ln 2 = 100 \times 0{,}69314718,$$

т.е. с достаточной для подобных расчетов точностью 100 ln2 = 70. «Таинственное» правило полностью обосновано.

Следствия теоремы умножения. Эта теорема позволяет без труда сменить начало отсчета. Например, в табл. 6.6 начальный момент времени — 31 марта 1991 г. (можно заменить на 1990 г., поскольку цены были постоянными). Поскольку по теореме умножения

$$I(t_2, t) = \frac{I(t_1, t)}{I(t_1, t_2)},$$

то без труда можно перейти к отсчету инфляции от первого дня третьего тысячелетия. Действительно, в соответствии со строкой 41 табл. 6.6 имеем:

$$I(01.01.01, t) = \frac{I(31.03.91, t)}{I(31.03.91, 01.01.01)} = \frac{I(31.03.91, t)}{33{,}34}.$$

Например, получаем, что инфляция за 6 лет (2001—2006) составляет

$$I(01.01.01, 10.01.07) = \frac{59{,}432}{33{,}340} = 1{,}7826,$$

т.е. 78,26%.

Обсудим соотношение инфляции по месяцам и за год, а также понятие среднего темпа роста цен и среднемесячной инфляции. Пусть I_1 — индекс инфляции за январь, I_2 — за февраль, I_3 — за март, ...

..., I_{12} — за декабрь, а $I_{\text{год}}$ — за год в целом. Тогда согласно теореме умножения

$$I_{\text{год}} = I_1 I_2 I_3 \dots I_{12}.$$

Как ввести понятие среднего индекса инфляции? Естественно исходить из требования, чтобы при подстановке среднего индекса инфляции вместо всех усредняемых величин итог не изменялся. Пусть $I_1, I_2, I_3, \dots, I_k$ — индексы инфляции за k последовательных интервалов времени, а $I_{\text{ср}}$ — средний индекс инфляции для этой совокупности. Тогда исходное требование — это требование выполнения равенства

$$I_1 I_2 I_2 \dots I_k = I_{\text{ср}}^k.$$

Таким образом,

$$I_{\text{ср}} = \sqrt[k]{I_1 I_2 I_2 \dots I_k},$$

т.е. средний индекс инфляции рассчитывается как среднее геометрическое. Например, средний индекс инфляции за 2005—2006 гг. равен (по официальным данным)

$$I_{\text{ср}} = \sqrt{1{,}109 \times 1{,}09} = 1{,}09946.$$

Отметим, что всегда среднее геометрическое меньше среднего арифметического:

$$\sqrt[k]{I_1 I_2 I_3 \times \ldots \times I_k} < \frac{I_1 + I_2 + I_3 + \ldots + I_k}{k},$$

за исключением единственного случая, когда все усредняемые величины равны между собой (равны и их среднему арифметическому, и их среднему геометрическому). Среднее арифметическое индексов инфляции за 2005 г. и 2006 г. равно (1,109 + 1,09)/2 = 2,199/2 = 1,0995, что больше среднего геометрического 1,09946, хотя превышение и невелико.

Среднемесячная инфляция, как и средний темп роста для любого временного ряда, рассчитывается в предположении, что ежемесячный рост цен не меняется от месяца к месяцу. Для данных табл. 6.5 она равна

$$\sqrt[36]{159{,}88} = 1{,}2274, \text{или } 22{,}74\%.$$

Другими словами, с 14 марта 1991 г. по 14 марта 1994 г. цены каждый месяц увеличивались в среднем на 22,74%.

Примеры ошибок при расчетах с индексами инфляции. Информация об индексах инфляции и рассуждения, связанные с ними, постоянно появляются на страницах печати и обсуждаются в иных средствах массовой информации. К сожалению, достаточно широко распространены ошибки.

Так, в одной из экономических (!) газет была помещена публикация, в которой основной исходный материал для обсуждения — индексы инфляции по месяцам (табл. 6.7).

Таблица 6.7

Индексы инфляции по месяцам

Месяц	Индекс инфляции	Месяц	Индекс инфляции	Месяц	Индекс инфляции
Январь	1,00	Май	1,29	Сентябрь	1,22
Февраль	1,23	Июнь	1,30	Октябрь	1,19
Март	1,19	Июль	1,23	Ноябрь	1,23
Апрель	1,25	Август	1,22	Декабрь	1,25

Автору публикации были нужны индексы инфляции за несколько месяцев. Рассчитывая их, он без каких-либо сомнений пользовался ранее рассмотренной приближенной формулой (сложение индексов «в процентах») вместо точной (перемножение индексов инфляции, выраженных «в разах»). В результате он получил для периода январь—декабрь (т.е. за год) значение индекса инфляции 3,6 (поскольку 0% + + 23% + 19% + 25% + 29% + 30% + 23% + 22% + 22% + 19% + 23% + + 25% = 260%), в то время как на самом деле индекс инфляции, рассчитанный в результате перемножения индексов по месяцам, равен 10,23. Допущенная ошибка в 10,23/3,6 = 2,84 раза существенно исказила дальнейшие расчеты (фонда оплаты труда, средней зарплаты и других экономических характеристик) в рассматриваемой публикации, названной в специализированной экономической газете не как-нибудь, а «консультацией»!

В еженедельнике «Аргументы и факты» в апреле 1994 г. в рубрике «Прогноз» помещена беседа журналистки Татьяны Коростиковой с первым заместителем министра экономики России Яковом Уринсоном [33], в которой Я. Уринсон прогнозирует: *«...Мы предполагаем рост цен за 1994 г. в 5 раз... В месяц — 7—8%...»*.

Сказанное противоречиво. Если индекс инфляции за год равен 5, то за месяц, очевидно, рост цен равен в среднем

$$\sqrt[12]{5} = 1{,}1435,$$

т.е. 14,35% в месяц, а не 7—8%. Если же рост цен составляет 7—8% в месяц, то индекс инфляции за год лежит между

$$(1{,}07)^{12} = 2{,}25 \text{ и } (1{,}08)^{12} = 2{,}58,$$

т.е., по крайней мере, в 2 раза меньше, чем названный в беседе достаточно реальный прогноз — рост в 5 раз. Остается неясным, кто дезориентировал читателя многотиражного издания — чиновник или журналист. Наш запрос об этом в редакцию «Аргументов и фактов» остался без ответа.

Покажем на примере этих данных накопление погрешностей при использовании приближенной формулы, основанной на суммировании индексов инфляции «в процентах». Если в месяц имеется рост на 14,35%, то за 1 год по приближенной формуле — на $14{,}35 \times 12 = 172{,}2\%$ (вместо 400% — рост в 5 раз). Если в 1 месяц — на 7%, то за 1 год — на $7 \times 12 = 84\%$ (вместо 125%). Если в 1 месяц — на 8%, то за 1 год — на $8 \times 12 = 96\%$ (вместо 152%).

Приведенных примеров достаточно для констатации того, что к сообщениям в средствах массовой информации, посвященным росту цен, следует относиться с известной осторожностью.

Теорема сложения. Целью введения индекса инфляции была выдвинута необходимость усреднения индивидуальных темпов роста цен (индивидуальных индексов инфляции)

$$I_i(t_1, t_2) = \frac{p_i(t_2)}{p_i(t_1)},\ i = 1, 2, \ldots, n.$$

Однако индекс инфляции был определен не как среднее таких величин, а как отношение стоимостей потребительских корзин. Тем не менее индекс инфляции действительно является средним взвешенным арифметическим индивидуальных индексов инфляции, как показывает следующая теорема.

Теорема 6.2 (теорема сложения). Существуют положительные весовые коэффициенты β_i, $i = 1, 2, ..., n$, такие, что

$$I(t_1, t_2) = \sum_{1 \le i \le n} \beta_i I_i(t_1, t_2),$$

причем $\beta_1 + \beta_2 + ... + \beta_n = 1$. При этом β_i — это доля стоимости потребительской корзины, приходящаяся на соответствующий (i-й) товар (услугу) в начальный (базовый) момент времени,

$$\beta_i = \frac{p_i(t_i)Q_i}{S(t_i)} - \frac{p_i(t_i)Q_i}{\sum_{1 \le k \le n} p_k(t_i)Q_k},\ i = 1, 2, \ldots, n.$$

Доказательство дается следующей последовательностью преобразований:

$$\sum_{1 \le k \le n} \beta_i I_i(t_1, t_2) = \sum_{1 \le k \le n} \frac{p_i(t_1)Q_i}{S(t_1)} \times \frac{p_i(t_2)}{p_i(t_1)} =$$

$$= \sum_{1 \le k \le n} \frac{p_i(t_2)Q_i}{S(t_1)} = \frac{1}{S(t_1)} \sum_{1 \le k \le n} p_i(t_2)Q_i = \frac{S(t_2)}{S(t_1)} = I(t_1, t_2),$$

(сокращается выражение $p_i(t_1)$, оказывающееся как в числителе, так и в знаменателе i-го слагаемого).

Теорема сложения справедлива и в случае, когда вместо индивидуальных коэффициентов инфляции стоят групповые. Например, при расчете индекса инфляции по потребительской корзине ИВСТЭ (см. табл. 6.4) можно сначала рассчитать индексы инфляции по 10 группам, выделенным в этой корзине (хлеб и хлебопродукты, овощи, сахар и кондитерские изделия и др.), а затем объединить их в единый индекс с помощью весовых коэффициентов согласно теореме сложения. Аналогично можно, получив индексы инфляции отдельно по продовольственной корзине, отдельно по товарам повседневного спроса, длительного спроса, отдельно по различным видам услуг (жилищно-коммунальных, транспортных, информационных и др.), получить итоговый индекс инфляции по объединенной корзине, например предусмотренной в Федеральном законе «О прожиточном минимуме в Российской Федерации» (в редакции федеральных законов РФ от 27.05.2000 № 75-ФЗ, от 22.08.2004 № 122-ФЗ). Большое значение имеет теорема сложения при расчета дефлятора ВВП (в целях приведения его к сопоставимым ценам), поскольку потребительская корзина при этом должна охватывать весь спектр конечных товаров и услуг, производимых на территории страны за год.

6.4. ИСПОЛЬЗОВАНИЕ ИНДЕКСОВ ИНФЛЯЦИИ В ЭКОНОМИЧЕСКИХ РАСЧЕТАХ

Хорошо известно, что в любой стране стоимость ее денежных единиц со временем меняется. Например, на 1 дол. США полвека назад можно было купить примерно в 8 раз больше материальных ценностей (например, продовольствия), чем сейчас (см. таблицу пересчета в учебнике [47]), а если сравнивать со временами Тома Сойера — 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 российского предприятия пока что, как правило, игнорируют сам факт наличия инфляции. Отличие финансиста от бухгалтера проявляется, в частности, в том, что бухгалтер имеет дело с величинами, выраженными в номинальных денежных единицах (поскольку в документах первичного бухгалтерского учета используются именно они), а финансист не может игнорировать изменение покупательной способности денежных единиц во времени.

Обсудим некоторые возможности использования индекса инфляции в экономических расчетах в процессе подготовки и принятия решений. Чтобы избежать непродуктивных эмоций при обсуждении современного экономического положения, отнесем большинство рассматриваемых примеров к ушедшей в историю середине 1990-х гг.

Будем пользоваться как данными ИВСТЭ (см. табл. 6.6), так и официальными (табл. 6.8). Сравнение столбцов (4) и (5) табл. 6.8 показывает, что официальная статистика занижала реальную инфляцию в 1,5—2,0 раза. Именно это было причиной того, что ИВСТЭ по заказу Минобороны России в 1990-е гг. проводил сбор и анализ данных о динамике цен. Заказчика интересовали размеры финансирования НИР в реальных (сопоставимых) ценах. Часть полученных результатов была опубликована [29; 89]. Мониторинг цен продолжается [75; 92].

Необходимо отметить, что в начале XXI в. темпы роста цен, фиксируемые независимыми исследователями (в частности, ИВСТЭ) и официальной статистикой, достаточно близки. Прежние расхождения были порождены реалиями 1990-х гг. и остались в прошлом тысячелетии. Однако, начиная с 2007 г., проявились расхождения нового типа. Специалисты отмечают нерешенные проблемы в области измерения инфляции, имеющие расхождения в подходах, отсутствие прозрачности в деятельности официальных статистических органов.

Таблица 6.8

Индексы инфляции в 1992—2007 гг.
(по данным официальных статистических органов)

Год	Индекс инфляции		Накопленная инфляция с января 1992 г.	Накопленная инфляция с марта 1991 г.	Данные ИВСТЭ к столбцу (5)
	«в разах»	«в процентах»			
(1)	(2)	(3)	(4)	(5)	(6)
1991	—	—	—	2,6	—
1992	26,1	25 100	26,1	67,86	—
1993	9,4	840	245,34	637,88	1 235,42
1994	3,15	215	772,82	2 009,33	3 680,34
1995	2,31	131	1 785,22	4 641,56	9 989,32
1996	1,218	21,8	2 174,39	5 653,42	11 815,31
1997	1,11	11,0	2 413,58	6 274,30	12 997
1998	1,844	84,4	4,451	11,572	20,395
1999	1,365	36,5	6,075	15,795	27,737
2000	1,202	20,2	7,303	18,986	33,340
2001	1,186	18,6	8,661	22,517	39,541
2002	1,151	15,1	9,968	25,917	45,512
2003	1,12	12,0	11,164	29,028	50,974
2004	1,117	11,7	12,471	32,424	55,037
2005	1,109	10,9	13,830	35,958	57,354
2006	1,09	9,0	15,075	39,194	59,342
2007	1,119	11,9	17,830	46,358	85,960

Примечание. Накопленные индексы инфляции с 1998 г. даются с учетом деноминации.

Переход к сопоставимым ценам. Индекс инфляции даст возможность перехода к сопоставимым ценам, расходам, доходам и другим экономическим величинам. Например, по данным ИВСТЭ (ср. табл. 6.6, в которой приведены значения индекса инфляции на 02.03.1995 и 30.03.1995), индекс инфляции за 4 года — с 14 марта 1991 г. по 16 марта 1995 г. — составил 5936. Это означает, что покупательной способности 1 руб. марта 1991 г. соответствует примерно 6000 руб. (а точнее, 5936 руб.) марта 1995 г.

Рассмотрим приведение доходов к неизменным ценам. Пусть Иван Иванович Иванов получал в 1990 г. 300 руб. в месяц, а в начале мая 1995 г. ему выдали 1 млн руб. за апрель (т.е за предыдущий месяц). Увеличились его доходы или уменьшились?

Номинальная заработная плата выросла в 1 000 000/300 = 3333 раза. Однако индекс инфляции на 18 мая 1995 г. составлял 6180 (дата взята исходя из того, что средства, полученные в начале мая, Иван Иванович Иванов тратит в течение этого месяца). Это значит, что 1 руб. 1990 г. соответствовал по покупательной способности 6180 руб. в ценах на 18 мая 1995 г. Следовательно, в ценах 1990 г. доход И.И. Иванова составлял 1 000 000 / 6180 = 161 руб. 81 коп., т.е. 53,9% от дохода в 1990 г.

Можно поступить наоборот — привести доход 1990 г. к ценам на 18 мая 1995 г. Для этого достаточно умножить его на индекс инфляции: доход 1990 г. соответствует 300 × 6180 = 1854 тыс. руб. в ценах мая 1995 г., что в 1 854 000 / 100 000 = 1,854 раза больше, чем месячный доход в 1990 г. Следовательно, доход мая 1995 г. составляет 100 (1/1,854)% = = 53,9% от дохода 1990 г. Нетрудно показать, что оба способа расчетов приводят к одному и тому же результату.

Средняя зарплата. Сопоставим инфляцию со средней заработной платой. В марте 1991 г. она равнялась приблизительно 300 руб. в месяц, т.е. минимальная продуктовая корзина ИВСТЭ составляла около 8,9% от средней заработной платы. В марте 1994 г. среднемесячная зарплата в Москве составила 206 076 руб. (данные Московского городского статистического комитета), следовательно, стоимость корзины ИВСТЭ составляла 4 349 900 / 206 076%=21,11% от нее. Если судить по ценам на продукты, за три года уровень жизни основной массы населения понизился в 21,1/11,4 = 1,85 раз.

В октябре 1995 г. в Москве средняя заработная плата — 520 тыс. руб., а стоимость потребительской корзины ИВСТЭ — 196,6 тыс. руб., т.е. 37,8% от средней зарплаты, падение уровня жизни — в 4,2 раза.

По данным Госкомстата России, средняя заработная плата составляла в 1990 г. 303 руб., в октябре 1993 г. — 93 тыс. руб., в январе 1995 г. — 303 тыс. руб. Поскольку зарплата тратится в основном в следующем месяце после получки, то рассмотрим индексы инфляции на 15 ноября 1993 г. и 02 февраля 1995 г., равные 1054 и 4856 соответственно (см. табл. 6.6). В ценах 1990 г. средняя зарплата составила 88 руб. 24 коп. и 62 руб. 40 коп. соответственно, т.е. 26,48% и 20,59% от зарплаты 1990 г.

Среднемесячная заработная плата (номинальная и в процентах от уровня 1990 г.) представлена в табл. 6.9. Она составлена по данным Пенсионного фонда Российской Федерации, использующего сведения о средней заработной плате при расчете величины пенсий. Обратим внимание, что средняя заработная плата в декабре 1994 г. (354 тыс. руб.) больше, чем в январе 1995 г. (303 тыс. руб.). Проявляется эффект конца года — дополнительные выплаты по итогам года в сочетании с некоторым затишьем производственной деятельности в начале следующего года.

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 1991—1995 гг. 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 (медиана распределения), т.е. не более 212 100 руб. по состоянию на январь 1995 г., а наиболее массовой является оплата в 50% от средней (мода распределения), т.е. около 150 тыс. руб. в месяц.

Таблица 6.9

Среднемесячная заработная плата

№ п/п	Дата	Среднемесячная заработная плата в России (по данным Пенсионного фонда Российской Федерации, ноябрь 2004 г.), руб.	Индекс инфляции $I(31.3.91; t)$	Среднемесячная заработная плата в РФ, в процентах от уровня 1990 г.
1	1990 г.	303 (за 1990 г.)	1,00	100
2	Август 1993 г.	65 400	665,08	32,45
3	Декабрь 1994 г.	354 200	3 680,34	32,08
4	Декабрь 1995 г.	735 500	9 989,32	24,30
5	Декабрь 1996 г.	1 017 000	11 815,31	28,32
6	Декабрь 1997 г.	760 000	12 997,00	19,30
7	Декабрь 1998 г.	760	23,395	10,72
8	Декабрь 1999 г.	1 086	32,004	10,97
9	Декабрь 2000 г.	1 584	35,684	14,80
10	Декабрь 2001 г.	1 671	43,321	12,73

Известно, что типичное распределение доходов таково, что мода величин доходов весьма меньше медианы, которая в свою очередь существенно меньше среднего арифметического (центральная часть распределения доходов — за исключением больших доходов — хорошо приближается логарифмически нормальным распределением). Дифференциация доходов в России быстро нарастала вплоть до второй половины 1990-х гг. и сильно превзошла уровень всех промышленно развитых стран. Правда, уровень Бразилии и Кении пока не достигнут, но и климат в этих странах существенно иной, так что минимальное жизнеобеспечение требует существенно меньших затрат.

Доходы отдельных слоев трудящихся снизились еще существеннее. Зарплата профессора Московского государственного института электроники и математики (технического университета) составляла в марте 1994 г. 42 руб. 92 коп. (в ценах 1990 г.), в июле 1995 г. — 43 руб. 01 коп., т.е. с 1990 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 1990 г.

Кроме того, необходимо учесть, что Госкомстат России (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 1994 г. (22 500 руб.) и в мае 1995 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 1995 г. составляла около 26,34% от стоимости корзины, т.е. реально уменьшилась в 1,44 раза по сравнению с сентябрем 1994 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 1995 г., снижение в 1,5 раза покупательной способности минимальной зарплаты, необходимо заключить, что в сентябре 1995 г. доход бюджетников был в 2 раза меньше, чем 1 год назад.

Оценим прожиточный минимум. Бюджетные обследования в 1990 г. показали, что для лиц с низкими доходами расходы на продовольствие составляют около 50% всех расходов, т.е. на промышленные товары и услуги идет около 50% доходов. Это соотношение подтвердило и проведенное ИВСТЭ бюджетное обследование конца 1995 г. Исходя из него, среднедушевой прожиточный минимум можно оценить, умножая на 2,0 стоимость минимальной продовольственной корзины ИВСТЭ (этот метод расчета прожиточного минимума разработан американской исследовательницей бедности польского происхождения М. Оршански [29]). Например, на 28 декабря 1995 г. — 531 432 руб. (см. табл. 6.6), т.е. прожиточный уровень для семьи из трех человек — муж, жена и ребенок — должен был на 28 декабря 1995 г. составлять 1,594 млн руб. (в месяц). Например, муж должен получать 900 тыс. руб., жена — 700 тыс. руб. в месяц (чистыми, т.е. после вычета подоходного налога). В декабре 1995 г. средняя заработная плата составляла 735 500 руб. (табл. 6.9). Сопоставление приведенных численных значений показывает, что среднеоплачиваемые работники не могли обеспечить прожиточный минимум для своей семьи (муж и жена суммарно могли заработать лишь 1,471 млн руб. «грязными», что заметно меньше прожиточного минимума).

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 27 июля 1995 г. — 6598, а для Иванова на 1 августа 1995 г. — 7542. Поскольку потребительская корзина на 14 марта 1991 г. в областном центре г. Иваново была на 95 коп. дешевле, то и на 1 августа 1995 г. она несколько

дешевле, чем была бы при том же индексе инфляции в Москве, и равна 193 452 руб., а прожиточный минимум равен 386 905 руб. Этот и другие подобные расчеты показывают, что приведенные численные значения для Москвы в качестве первого приближения можно использовать для различных регионов России.

Индексы инфляции с помощью описанной 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 мая 1994 г. по 18 мая 1995 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. Другими словами, реальный процент платы за депозит оказался отрицательным, равным –16% в годовом исчислении, при том, что в номинальных рублях договор с банком обеспечивает 214% годовых.

Пусть фирма получила кредит под 200% годовых. Значит, вместо 1 руб., полученного в настоящий момент в кредит, через год ей надо отдать 3 руб. Пусть она взяла кредит 19 мая 1994 г., а отдает 18 мая 1995 г. Тогда в ценах на момент взятия кредита она отдает 3 / 3,73 = 0,8 руб. за 1 руб. кредита (в сопоставимых ценах на момент выдачи 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 1992—1994 гг. Но бесплатных подарков в бизнесе не бывает — за них надо платить по другим каналам, как правило, криминальным.

Сколько стоит доллар? На 14 августа 1993 г. курс доллара США составлял в Российской Федерации 1984,5 руб., а индекс инфляции — 665,08. Следовательно, в сопоставимых ценах 1990 г. реальный курс доллара США равнялся 1984,5 / 665,08 = 2 руб. 98 коп.

В июле 1995 г. индекс инфляции составлял около 6500 (см. табл. 6.6), а курс доллара США — около 4500 руб. за 1 доллар. Следовательно, 1 дол. США стоит 4500 / 6500 = 0,69 руб. в ценах 1990 г., т.е. примерно соответствует официальному обменному курсу в 1980-х гг. Сопоставим с тем, что в сентябре 1994 г. курс доллара был около 2000, а индекс инфляции — около 2200, т.е. 1 дол. стоил около 0,91 руб. в ценах 1990 г. Реальная покупательная способность доллара упала за 10 месяцев в 1,32 раза.

Пик реального курса доллара приходился на время после дефолта 1998 г. На начало января 1999 г. курс был 20 руб. 65 коп. при индексе инфляции 20,395, т.е. в реальном исчислении он составлял 20,65 / 20,395 = 1 руб. 01 коп. За год до этого, в начале 1998 г., индекс инфляции равнялся 12,997, курс доллара — 5 руб. 96 коп., в реальном исчислении 5,96 / 12,997 = 0,46, т.е 46 коп. — в 2 с лишним раза меньше, чем в начале 1999 г.

В середине 2003 г. курс доллара был 30 руб. 38 коп., индекс инфляции составлял 48,56, следовательно, 1 дол. США по своей покупательной способности в России на июль 2003 г. соответствовал 63 коп. начала 1991 г. В начале 2004 г. курс доллара составлял около 28 руб. 50 коп. при индексе инфляции 50,974, следовательно, 1 дол. США по своей покупательной способности в России на начало 2004 г. соответствовал 55 коп. начала 1991 г. Всего за полгода доллар подешевел на 15%.

В конце июля 2007 г. курс доллара США был равен 25 руб. 41 коп. при индексе инфляции 61,819, следовательно, реальный его курс — 41 коп. (в сопоставимых ценах 1990 г.).

В конце декабря 2008 г. курс доллара США — 29 руб. 00 коп. при индексе инфляции около 100, следовательно, его реальный курс 29 коп. (в сопоставимых ценах 1990 г.).

Реальный курс доллара США на любой момент времени можно получить, располагая двумя широко доступными временны́ми рядами — ежедневными данными о номинальном курсе доллара США на Московской межбанковской валютной бирже (ММВБ) и рядом значений индексов инфляции (см. табл. 6.6 и 6.8). В 1990-е г. был распространен миф о том, что можно избавиться от влияния инфляции, ведя расчеты в долларах США. Этот миф опровергается приведенными результатами расчетов. Инфляция уменьшает покупательную способность доллара США как в нашей стране, так и в самих США, а также и в других странах.

Как известно, курс доллара США в Российской Федерации определяется в результате торгов на ММВБ. Каковы свойства валютного рынка? Ясно, что это отнюдь не рынок чистой конкуренции [89; 110]. Игроки не являются равноправными. Участвующие в торгах коммерческие банки административно зависят от Центрального банка Российской Федерации. Другой инструмент влияния Банка России — долларовые или рублевые интервенции. Необходимо признать, что реально курс доллара во многом определяется руководством страны, решающим поставленные перед собой задачи, действуя через Банк России. Не будем пытаться обсуждать эти задачи, констатируем только, что официальный курс доллара США в Российской Федерации определяется во многом административным путем. Возникает естественный вопрос: а каков же реальный курс?

Международные сопоставления на основе паритета покупательной способности. Процитируем типичную публикацию в средствах массовой информации:

«Российский рубль входит в число самых недооцененных мировых валют по паритету покупательной способности» — говорится в очередном исследовании журнала «The Economist». Исходя из результатов исследования, справедливая цена доллара составляет 15,2 руб. Однако участники рынка считают, что подешеветь до 15 руб. доллар сможет лишь через 15—20 лет.

Индекс «Биг Мака» основывается на паритете покупательской способности (ППС), рассчитанной с помощью одинакового во всех странах мира продукта (в данном случае «Биг Мака»). Согласно методике расчета индекса «Биг Мака», курсы валют должны быть такими, чтобы стоимость этого продукта была во всех странах одинаковой.

Индекс «Биг Мака», рассчитанный журналом «The Economist», показал, что рубль недооценен на 41%. При этом китайский юань, например, недооценен на 58% (ежедневная деловая газета «РБК Daily», 09.07.2007).

Обсудим три вопроса. По каким причинам реальное соотношение валют может отличаться от официального? Как установить реальный курс? Какие последствия влечет пересчет с официальных курсов на реальные?

Если официальный курс доллара США по отношению к рублю выше реального, то это означает, что государство защищает отечественных товаропроизводителей, поскольку зарубежные товары (из долларовой зоны) продаются внутри страны дороже, чем было бы при соответствии официального курса реальному. Государство поддерживает также работу отечественных предприятий на экспорт, искусственно занижая издержки производства. Одновременно завышение официального курса ставит препятствия на пути закупки новейших зарубежных технологий, делает невыгодным получение кредитов. Ясно, что возвращать долги в долларах легче при низком курсе доллара, чем при высоком. Остановившись на сказанном, констатируем, что в те или иные периоды своего развития государство, ведущее активную экономическую политику, имеет основания устанавливать официальный курс обмена валюты, отличный от реального, соответствующего свободному рынку.

Как же установить реальный курс? Принцип ППС предлагает исходить из того, что одна и та же потребительская корзина должна стоить одинаково в разных странах. Если потребительская корзина ИВСТЭ стоит в начале июля 2007 г. в Москве 1644 руб., а в Нью-Йорке, к примеру, 100 дол., то приравниваем: 1644 руб. = 100 дол. США, т.е. курс доллара по ППС — 16 руб. 44 коп. В процитированной публикации из СМИ приравнивалась стоимость «Биг Мака» — продукта, который в распространенной по многим странам мира сети ресторанов быстрого питания «Макдоналдс» всюду изготавливается по одной и той же рецептуре. Другими словами, в качестве используемой для сравнения потребительской корзины берется набор товаров и услуг, необходимых для изготовления «Биг Мака» (включая продукты питания, электроэнергию, оплату труда, амортизацию оборудования и т.п.). Надо отметить, что в зависимости от конкретной методики международного сопоставления, выбранной тем или иным экономистом (в частности, конкретной потребительской корзины и способа измерения ее стоимости), оценки реального курса валют по ППС могут заметно различаться, например курс доллара США — от 8 до 15 руб. (по состоянию на 2007 г.). Несмотря на разногласия, общий вывод одинаков — курс доллара США в России завышен в несколько раз.

Международные сопоставления на основе ППС приводят к принципиально иным результатам, чем на основе официальных курсов обмена валют. В качестве примера приведем табл. 6.10. Продемонстрируем это различие на примере валового национального дохода (ВНД) десяти ведущих стран мира. Таблица составлена на основе «Доклада о мировом развитии», представленного Всемирным банком в 2004 г. Валовой национальный доход страны — одна из основных ее макроэкономических характеристик; он меньше валового национального продукта (ВНП) на величину амортизационных отчислений. Другими словами, ВНД аккумулирует добавленную стоимость, произведенную живым трудом в течение года, в то время как в ВНП входят и перенесенные на вновь созданные товары и услуги результаты прошлого труда. Все показатели табл. 6.10 рассчитаны специалистами Всемирного банка по принятым в этой организации методикам.

Таблица 6.10

Валовой национальный доход в 2002 г., млрд дол. США

№	Страна	Объем ВНД	Место в мире по ППС	Объем ВНД по ППС
1	США	10 110	1	10 100
2	Китай	1 210	6	5 625
3	Япония	4 266	2	3 315
4	Индия	502	11	2 691
5	Германия	1 870	3	2 163
6	Франция	1 343	5	1 556
7	Великобритания	1 486	4	1 523
8	Италия	1 098	7	1 467
9	Бразилия	497	12	1 266
10	Россия	308	16	1 127

В таблице 6.10 страны упорядочены в соответствии с убыванием объема ВНД, рассчитанного по ППС. Видно, что упорядочение по этому показателю значительно отличается от упорядочения по ВНД, соответствующего средним обменным курсам 2002 г. Китай с 6-го места поднялся на 2-е, далеко опередив Японию, Германию, Великобританию, Францию, которые предшествовали ему по «официальному» ВНД. Индия поднялась с 11-го места на 4-е, Бразилия — с 12-го на 9-е, Россия — с 16-го на 10-е. Соответственно сдвинулись вниз ведущие европейские страны.

Особенно интересно обсудить положение экономики Китая в мире. По некоторым оценкам, уже сейчас Китай обладает самой мощной экономикой в мире, его ВВП превышает ВВП США. По «экономической силе» на 1-м месте — Китай, на 2-м — Европейский союз, и только на 3-м — США. Это упорядочение необходимо учитывать российским организациям при стратегическом планировании.

Проблема учета инфляции при экономическом анализе финансово-хозяйственной деятельности предприятия. Индексы инфляции используют 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. Однако «в первом приближении» можно использовать потребительскую корзину ИВСТЭ или применять индексы инфляции Росстата, особенно для тех организаций, для которых в структуре себестоимости выпускаемых товаров и услуг большое место занимает оплата труда.

Рассмотрим условное предприятие. В таблице 6.11 представлена информация о прибыли предприятия по годам. Эти значения взяты из ежегодных отчетов, сданных в налоговые органы, и выражены в номинальных денежных единицах. Видим, что прибыль год от года растет, за 6 лет увеличилась на 80%. Напрашивается вывод, что предприятие процветает, его руководители заслуживают похвал и наград.

Таблица 6.11

Динамика прибыли предприятия, млн руб.

№ п/п	Год	Прибыль, млн руб.	Индекс инфляции, раз	Накопленная инфляция, раз	Прибыль в сопоставимых ценах
(1)	(2)	(3)	(4)	(5)	(6)
1	2000	1,0	—	—	1,0
2	2001	1,1	1,186	1,186	1,1 / 1,186 = 0,927
3	2002	1,3	1,151	1,365	1,3 / 1,365 = 0,952
4	2003	1,4	1,12	1,529	1,4 / 1,529 = 0,912
5	2004	1,5	1,117	1,708	1,5 / 1,708 = 0,878
6	2005	1,7	1,109	1,894	1,7 / 1,894 = 0,896
7	2006	1,8	1,09	2,064	1,8 / 2,064 = 0,872

Однако не будем торопиться. Приведем прибыль к сопоставимым ценам. В качестве точки отсчета естественно взять начало тысячелетия, т.е. конец 2000 — начало 2001 г. Другими словами, приведем интересующую нас характеристику работы предприятия к сопоставимым ценам на 1 января 2001 г. Именно в этих ценах выражена прибыль 2000 г. Будем использовать официальные данные Росстата (см. табл. 6.8). Индексы инфляции приведены в столбце (4) табл. 6.11.

Для приведения прибыли 2001 г. к сопоставимым ценам на начало года достаточно разделить ее на годовой индекс инфляции: 1,1 / 1,186 = = 0,927. Вот уже первая неожиданность: реальная прибыль не выросла в 2001 г. на 10% по сравнению с 2000 г., как номинальная, а наоборот, упала на 7,3%.

Чтобы привести к сопоставимым ценам прибыль 2002 г. (и следующих лет), надо сначала найти накопленную инфляцию за прошедшие годы. За 2001—2002 гг. индекс инфляции находится путем перемножения индексов за отдельные годы: 1,186 × 1,151 = 1,365. Аналогично индекс инфляции за 3 года (2001—2003) равен 1,186 × 1,151 × 1,12 = = 1,529, за 4 года 1,529 × 1,117=1,708, за 5 лет — 1,708 × × 1,109 = 1,894, наконец, за 6 лет (2001—2006) — 1,894 × 1,09 = 2,064.

Рассчитанные значения прибыли в сопоставимых ценах приведены в столбце (6) табл. 6.11. Наблюдаем совсем другую картину по сравнению с номинальной прибылью. Реальная прибыль отнюдь не растет, наоборот, имеет устойчивую тенденцию к снижению. К 2006 г. она снизилась на 12,6%, в то время как номинальная прибыль выросла на 80%.

И выводы получаются совсем другие. Нельзя сказать, что предприятие прогрессивно развивается. Констатируем тенденцию к застою и деградации. Руководители вряд ли заслуживают похвал и наград, наоборот, им следует тщательно проанализировать ситуацию и разработать меры улучшения работы предприятия. Хотя надо отметить, что говорить о катастрофе преждевременно: прибыль остается положительной, ее снижение не слишком большое.

Как известно, разработана и широко применяется развернутая система коэффициентов, используемых при экономическом анализе финансово-хозяйственной деятельности предприятия. 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, согласно правилам бухгалтерского учета, инфляцию полностью игнорируют (за исключением периодических корректировок стоимости основных фондов в соответствии с принимаемыми руководством страны нормативными документами). Это приводит к искажению оценки реального положения предприятия, если пользоваться только данными текущего бухгалтерского учета. 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, например для закупки необходимого сырья.

Ясно, что учитывать инфляцию надо. Вопрос в другом: как именно? Потребительская корзина должна, видимо, состоять из тех товаров и услуг, которые предприятие закупает. Стоимость основных фондов может не убывать в соответствии с амортизацией, а возрастать, согласно отраслевому темпу инфляции (уменьшенному на амортизационные коэффициенты), и т.д. Обсуждение конкретных методик расчетов выходит за пределы настоящего учебника.

Сколько стоит предприятие? Специалисты по оценке бизнеса используют три подхода — затратный, доходный и сравнительный [120]. Согласно первому из них, важно оценить основные фонды. Для этого нужно взять их стоимость в определенный момент времени, например в 1990 г., и умножить на индекс инфляции (и учесть амортизационные отчисления). Вспомним здесь, что официальный индекс инфляции Госкомстата—Росстата в 1,5 раза меньше, чем наш, если отсчитывать от 1990 г. (см. табл. 6.8). Есть много способов исказить экономические показатели, и «специалисты» ими умело пользуются. Занижение индекса инфляции выгодно тем, кто хочет обзавестись собственностью по заниженной цене.

В мировой практике известны различные варианты учета инфляции в бухгалтерской деятельности и в работе финансовых аналитиков.

6.5. ДИНАМИКА ЦЕН НА ПРОДОВОЛЬСТВЕННЫЕ ТОВАРЫ

Проведем сравнительный анализ результатов расчетов на основе различных потребительских корзин для оценки точности определения темпов роста цен.

Сравнение индексов инфляции Госкомстата России и ИВСТЭ. Результаты расчетов по различным потребительским корзинам дают, естественно, различные значения индексов инфляции, хотя эти различия, как представляется, не слишком значительны. Так, в таблице 6.3 приведена потребительская корзина Центра экономической конъюнктуры (ЦЭК) при Правительстве Российской Федерации и Государственного комитета Российской Федерации по статистике — корзина Госкомстата России. А в таблице 6.4 дана потребительская корзина ИВСТЭ. Было проведено сравнение соответствующих индексов инфляции. Полученные результаты (табл. 6.12) показали, что эти индексы достаточно близки.

Таблица 6.12

Сравнение результатов подсчета стоимостей потребительских корзин и индексов инфляции Госкомстата России и ИВСТЭ

Промежуток времени	Стоимость корзины и индекс инфляции	
	Госкомстат России	ИВСТЭ
С 14.03.1991 г. по 14.03.1994 г.	30,82 / 48 990,33; 1 589,8	26,85 / 40 889,1; 1 598,88
С 15.11.1993 г. по 14.03.1994 г.	31 255 / 48 990,33; 1,57	28 050 / 40 889,1; 1,46
С 19.05.1994 г. по 26.05.1994 г.	56 670,2 / 57 667,75; 1,02	55 615 / 56 332; 1,01

Примечание. Верхние числа — стоимости потребительских корзин (в руб.) соответственно на первую указанную дату и через дробь — на вторую, нижнее число — индекс инфляции за данный период.

Близость различных индексов инфляции за большой промежуток времени объясняется тем, что цены растут в целом достаточно согласованно, «аномалии» выправляются: если темп роста цены определенного продукта отстает от среднего роста цен, то имеются основания полагать, что его цена в ближайшее время сильно возрастет. Однако на малых и средних промежутках времени проявляется различие роста цен на отдельные товары.

Тем более интересно, что официально публикуемые индексы инфляции Госкомстата России при отсчете с 1990 г. (или, что то же самое, с 14.03.1991 г.) давали в середине 1990-х гг., по крайней мере, вдвое меньшие значения, чем расчеты ИВСТЭ [49].

Изучение динамики цен в условиях реформ. С начала 1990-х гг. в России осуществляется так называемая «радикальная экономическая реформа». Одним из сопутствующих ей эффектов является изменение сложившейся к 1991 г. системы цен на все товары, услуги, труд (рабочую силу). Эти изменения цен приобрели ярко выраженный инфляционный характер. За время «радикальной экономической реформы» произошли изменения не только абсолютных величин цен, но и их пропорций.

Масштабы инфляции были определены не только дисбалансом между скопившейся к 1992 г. у населения значительной массой наличных денег и наличием товаров, но и массовым преобразованием безналичных средств предприятий в наличные деньги в период расцвета совместных предприятий и кооперативов в 1989—1991 гг. (а также отменой монополии внешней торговли, в результате чего, например, около $^1/_3$ произведенных в СССР в 1990 г. товаров массового потребления было вывезено за границу). В дальнейшем в результате применения жестких мер (например, невыплаты заработной платы), ограничивающих поступление наличных денег на рынок товаров и услуг, а также ограничивающих количество покупателей среднего класса и тем самым обеспечивающих искусственное снижение спроса, темпы инфляции заметно снизились, но инфляция не прекратилась. Болезнь не исчезла. Удалось сбить температуру больного, т.е. отключить сигнальную систему, но не вылечить болезнь. Стоит только начать платить людям наемного труда заработанные ими деньги при условии корректной оценки труда как рыночного товара, как инфляционная болезнь возобновится (как это и произошло в 2007 г. — см. далее). В августе 1998 г. инфляция была подстегнута руководством страны путем искусственного подъема курса доллара.

Институт высоких статистических технологий и эконометрики изучает динамику экономического положения граждан России на основе независимо собранной информации. Приведение к сопоставимым ценам (с помощью индексов инфляции и дефляторов) — составная часть любого экономического расчета, связанного более чем с одним моментом времени. Как показали наши наблюдения над ценами, использование публикуемых Госкомстатом России значений индексов инфляции приводит к систематическим ошибкам. Так, например, по нашим данным, цены за шесть с небольшим лет (с марта 1991 г. по май 1997 г.) выросли в среднем примерно в 12 500 раз, а по данным Госкомстата России — примерно в 6000 раз. Сказанное определяет актуальность использования независимой информации о ценах и индексах инфляции при анализе экономического положения России, а также при разработке прикладных моделей и методов управления в современных условиях.

Предметом описанного здесь исследования ИВСТЭ [87; 89] является оценка изменения в ходе реформ фактического среднего и минимального физиологически необходимого уровней жизни граждан Российской Федерации через сравнение индексов инфляции, вычисленных на основании потребительских корзин, и индекса изменения величины средней заработной платы.

Организация сбора и анализа данных. В 1994—1997 гг. еженедельно собирались данные о ценах 35 продуктов в 12 точках Москвы, Подмосковья и Крыма, а именно: в девяти точках Москвы; в двух точках Московской области (г. Раменское и г. Ногинск) и в Крыму (г. Симферополь).

Расчеты по собранным ценам продовольственных товаров проводились для следующих пяти потребительских корзин:

ИВСТЭ — продовольственная потребительская (продуктовая) корзина Института высоких статистических технологий и эконометрики (см. табл. 6.4). Составлена с учетом разработок Института питания РАМН. Является сбалансированной по белкам, жирам и углеводам. Обеспечивает минимальные физиологически необходимые потребности человека;

ГКС-1 — продуктовая корзина из 19 продуктов питания (включая сигареты) Государственного комитета Российской Федерации по статистике, применявшаяся в 1993—1996 гг. (см. табл. 6.3);

ГКС-2 — продуктовая корзина Государственного комитета Российской Федерации по статистике, используемая с 1 января 1997 г. Нормы потребления предложены Министерством труда;

Бюдж-1 — продуктовая корзина, разработанная на основе бюджетного обследования «бедных семей» студентов Московского государственного института электроники и математики (технического университета), среднедушевое потребление в которых не превосходит 90% от медианы обследованной совокупности семей. Эта корзина определяет усредненные фактические объемы потребления (кг/год/человек и, соответственно, кг/мес/человек) 35 продуктовых товаров, по которым производились измерения цен при использовании корзины ИВСТЭ. Общий объем затрат «бедных семей» на продуктовые и иные товары предлагается находить умножением стоимости корзины Бюдж-1 на соответствующие коэффициенты по методу Оршански;

Бюдж-2 — продовольственная потребительская корзина, разработанная на основе бюджетного обследования семей студентов Московского государственного института электроники и математики (технического университета) за октябрь—ноябрь 1995 г. Совокупность обследованных семей в целом характеризуется средним уровнем потребления. Эта корзина определяет усредненные фактические объемы потребления (кг/год/человек и, соответственно, кг/мес/человек) 35 продуктовых товаров, по которым производились измерения цен в корзине ИВСТЭ. Общий объем затрат «семей со средним достатком» на продуктовые и иные товары предлагается находить умножением стоимости корзины Бюдж-2 на соответствующие коэффициенты.

Количество элементарных измерений (значений собранных цен продовольственных товаров) приблизительно равно 30 000. Собранные данные о ценах обрабатывались на компьютерах как известными, так и оригинальными методами. Точность вычислений равна обычной компьютерной точности при работе с числами с плавающей точкой. В дальнейшем все средние величины цен приведены с точностью до 1 руб., а величины процентов — с точностью до 0,1%.

При проведении мониторинга за ценами на продукты питания исследователи обычно создают компьютерную базу данных, в которую заносятся следующие сведения:

1) название продукта питания;
2) объем потребления продукта;
3) цена продукта;
4) дата снятия цены;
5) название торговой точки.

Кроме того, используется вполне понятная последовательность действий (алгоритм) по вычислению индекса инфляции с момента времени t_1 по момент времени t_2:

1) вычислить сумму (по всем составляющим потребительской корзины) произведений цен на объем потребления для момента времени t_1;

2) вычислить сумму произведений цен на объем потребления для момента времени t_2;
3) найти их отношение.

Для нахождения индексов инфляции по товарным группам эти действия выполняются для продуктов искомой группы.

Для вычисления индекса инфляции по продуктам питания разработаны различные программные средства.

Результаты анализа динамики цен. Приведем некоторые результаты анализа данных о ценах. Начнем с временных рядов стоимостей потребительских корзин в Москве. Оказалось, что стоимость потребительской корзины ГКС-2 примерно в 1,5 раза меньше стоимости потребительской корзины ГКС-1. Потребительская корзина ИВСТЭ располагалась по стоимости примерно посередине между ГКС-1 и ГКС-2. Несмотря на различие стоимостей, индексы инфляции для всех трех корзин — ГКС-1, ГКС-2, ИВСТЭ — близки и составляют 8233—8896 на конец декабря 1995 г. и 10 396—10 890 — на конец февраля 1997 г. Любопытно отметить, что ГКС-1 имеет наименьшие значения индекса из трех корзин, а ГКС-2 — наибольшие, если сравнивать с мартом 1991 г. (Госкомстат России такие сравнения не проводит), в то время как рост цен за исследуемый промежуток времени (с конца декабря 1995 г. по конец февраля 1997 г.) наибольший рост цен дает корзина ИВСТЭ (28,05%), а наименьший — ГКС-2 (22,42%).

Совсем иная картина со стоимостями потребительских корзин Бюдж-1 и Бюдж-2. Они относятся к реальному потреблению сравнительно обеспеченных москвичей, включают в себя стоимости не только продуктов, но и других товаров и услуг, в то время как корзины ИВСТЭ, ГКС-1 и ГКС-2 дают представление о стоимости минимального набора товаров и услуг, обеспечивающего физиологические потребности человека. В конце декабря 1995 г. стоимость корзины Бюдж-1 (для «бедных») составляла 659 852 руб., а корзины Бюдж-2 (для «средних» семей) — 726 364 руб., а к февралю 1997 г. они «подросли» до 832 498 руб. (на 26,16%) и 950 989 руб. (на 30,92%) соответственно. Эти величины больше прожиточного минимума, согласно данным Московской федерации профсоюзов (750 тыс. руб. в мае 1997 г.), хотя разницу нельзя назвать заметной.

Интереснее другое — общий рост цен (на февраль 1997 г.) составил 8060—8446, т.е. примерно на 20% меньше, чем рост стоимостей корзин ИВСТЭ, ГКС-1, ГКС-2. Значит, «реформы» сильнее всего ударили по наиболее дешевым товарам, предназначенным для наиболее бедной части населения. Это связано, видимо, с сокращением и прекращением дотаций для таких товаров. Правда, затем темпы роста выровнялись — при сравнении февраля 1997 г. с декабрем 1995 г. они составляют 28,05% для корзины ИВСТЭ, 26,27% — для ГКС-1, 26,16% — для Бюдж-1 и 30,92% — для Бюдж-2. Особняком стоит ГКС-2 — 22,42%, что заметно меньше, чем для других корзин. В то же время наибольший рост для корзины Бюдж-2 может указывать на тенденцию более быстрого роста цен на товары, предназначенные для более состоятельных людей.

Анализ временных рядов стоимостей потребительских корзин и индексов инфляции по Подмосковью в целом подтверждает приведенные выводы, сделанные по московским данным. Снова наблюдаем близость роста цен с 1991 г. для корзин ИВСТЭ, ГКС-1, ГКС-2, снова ГКС-2 в полтора раза дороже ГКС-1, снова темп роста с декабря 1995 г. меньше всего из этих трех корзин у ГКС-2. Снова индексы с 1991 г. для корзин Бюдж-1 и Бюдж-2 на 20—25% меньше, чем для первых трех корзин. Однако с декабря 1995 г. наибольший рост стоимости корзины не у двух последних, а у ГКС-1 (на 2-м месте — корзина ИВСТЭ). Возможно, это отражает меньшую долю состоятельных людей в Подмосковье и соответственно меньшую ориентацию торговцев на их «покупательные» возможности.

Обращает на себя внимание меньшая величина индексов с марта 1991 г. в Подмосковье по сравнению с Москвой. Возможно, дело в том, что стоимости потребительских корзин по состоянию на 31 марта 1991 г. брались те же, что и в Москве, поскольку сведения о ценах на тот момент в Московской области у нас отсутствуют. Это приводит к занижению истинных значений индексов инфляции, поскольку и до 31 марта 1991 г. цены в Подмосковье были несколько ниже, чем в Москве. Это относится, в частности, к ценам на овощи и фрукты, молочные продукты и др.

Вполне естественно, что с марта 1991 г. цены на различные товары выросли по-разному. Так, цены на рыбу (треска, минтай) выросли примерно в 25 000 раз, а цена на сахар — менее чем в 4000 раз. Цены на творог выросли в 2,5 раза больше, чем на сыр, и т.д. В Москве и Московской области рост цен достаточно хорошо согласован. Можно было бы предположить, что в рыночных условиях были исправлены диспропорции прежней дотационной плановой системы. Тогда рост цен после декабря 1995 г. должен был бы быть примерно равномерным, отражающим динамику общеэкономических процессов. Однако конкретные эмпирические данные о динамике цен отвергают это предположение.

В Москве при общем среднем росте цен на 20—30% больше всего выросли цены на огурцы (74,8%), баранину (75,9%), птицу (74,5%), упали цены на капусту (–4,6%), сахар (–5,5%). В Московской области при таком же среднем росте цен больше всего выросли цены на мясо — говядину (82,6%), свинину (88,6%), баранину (107,6%), при этом упали цены на картофель (–10%), капусту (–10%), сахар (–13,1%), конфеты (–21,1%), минтай (–6,5%), растительное масло (–20,6%) и маргарин (–13 %).

Приходится констатировать, что цены растут непропорционально, стабилизация цен не наступила, более того, динамика цен на отдельные товары не только не согласована, но и отнюдь не близка. ***Нет никаких признаков приближения к равновесным ценам***, чего можно было бы ожидать после пяти лет «либерализации» в соответствии с данными учебников по экономической теории. В качестве дополнительного следствия из сказанного вытекает, что, подбирая нужным образом номенклатуру товаров для потребительской корзины, можно получить индекс инфляции желательной величины — от значительного роста (+80%) до падения цен (–20%).

Временные ряды наименьшей, средней и наибольшей из зарегистрированных по Москве цен 35 продовольственных товаров показывают, что такое понятие, как «***цена товара***», строго говоря, ***некорректно***. Оно применимо к единственному акту купли-продажи определенного товара в фиксированном месте, в крайнем случае — к актам купли-продажи в определенном магазине, но не к огромному городу в целом. Действительно, зафиксированные нашими сотрудниками цены на один и тот же товар в один и тот же день могут различаться в несколько раз. Так, 26 июня 1996 г. максимальная зафиксированная цена на рис превышала минимальную в 3,04 раза, а на картофель — в 3,13 раза. Аналогичное превышение для баранины 27 декабря 1996 г. равно 2,79. Типовое же превышение максимальной цены над минимальной — в 1,5 раза. Ничего странного в сказанном нет — всем московским потребителям известно, что наибольшие цены — в центральных престижных магазинах, средние — в рядовых магазинах, наименьшие — на оптовых рынках.

В целях обеспечения сопоставимости данных сотрудники ИВСТЭ собирали сведения в одних и тех же местах (магазинах, киосках, на рынках). Это позволяло отслеживать рост цен и получать корректные значения индексов инфляции. Однако это делало несколько условной стоимость потребительской корзины — потребитель, потратив время и обойдя достаточное число мест продажи, мог обеспечить себя теми же продуктами по менее высоким ценам. Дополнительную сложность вносит большая номенклатура видов одного и того же товара. На какой тип батона белого хлеба ориентироваться? Что понимать под говядиной — отечественную или импортную, вырезку или кости для супа? Объективно существующая свобода при решении организаторами исследования жизненного уровня подобных вопросов дает возможность

для сдвига результатов в заранее заданном направлении. Объективно цены не являются стабильными в пространстве и во времени.

На практике указанные сложности в основном преодолимы. Оказалось, в частности, что стоимость потребительских корзин в различных районах Москвы хотя и отличается, но не более чем на 5—10%. Отклонения в стоимости отдельных продуктов частично компенсируют друг друга.

Нами изучены вклады отдельных продовольственных товаров в стоимости потребительских корзин. Обращает на себя внимание различие между нормативными (т.е. заданными априори) корзинами ИВСТЭ, ГКС-1, ГКС-2 и полученными в результате анализа реального потребления корзинами Бюдж-1 и Бюдж-2. В реальном потреблении гораздо меньше муки, пшена, геркулеса, ржаного хлеба, картофеля, трески, минтая, молока, маргарина, но гораздо больше лука, яблок, конфет, колбасы, сельди, сливочного масла, сыра. Объяснение достаточно очевидное: корзины ИВСТЭ, ГКС-1, ГКС-2 — это «корзины выживания», действительно минимальные по стоимости корзины, в то время как корзины Бюдж-1 и Бюдж-2 — это корзины реального потребления в семьях студентов Московского государственного института электроники и математики (технического университета) различного достатка.

Продовольственные товары, на наш взгляд, можно разделить на две группы. Цены на товары первой группы растут монотонно, без всякой связи со временем года, т.е. ведут себя примерно так же, как промышленные товары. Можно предположить, что индексы инфляции, построенные по подмножеству таких товаров, представляют собой общие индексы, «очищенные от сезонности», а потому лучше описывающие реальное состояние экономики, чем исходные индексы. Однако при их применении теряется связь со стоимостью корзины выживания, обеспечивающей существование без физиологического вырождения.

Вторая группа — это товары с ярко выраженной сезонностью, прежде всего овощи, цены на которые падают во второй половине лета и осенью, а затем начинают возрастать. Наличие этой составляющей приводит к тому, что рост стоимостей корзин практически останавливается летом, а наиболее быстрым является зимой.

Можно ли управлять процессом роста цен? Мы наблюдали результаты явно административного воздействия: в ноябре 1995 г., перед выборами в Государственную Думу, цены в Москве внезапно упали на 9%, хотя в ноябре цены обычно растут быстрее, чем в иное время года. Тем не менее необходимо констатировать, что обычно изменение цен происходит на микроэкономическом уровне, хотя и провоцируется макроэкономическими процессами, в частности монопольными изменениями цен на энергоносители.

Ложная, на наш взгляд, идея монетаристов состоит в том, что они считают необходимым бороться с инфляцией, сокращая денежную массу в стране, например не выплачивая вовремя зарплату и пенсии. Однако, как пишет академик-секретарь Отделения экономики РАН Д.С. Львов, «макроэкономические расчеты показывают, что за каждый процент сокращения инфляции приходится расплачиваться тремя-пятью процентами спада производства» [44, с. 11]. Основной удар монетаристской политики приходится не по инфляции, а по производству.

Процесс инфляции частично управляем административными методами. Осенью 1996 г. спрогнозированного ИВСТЭ роста цен не произошло, что объясняется изменением условий — правительство перешло к борьбе с инфляцией путем гигантского роста задолженностей по зарплате, пенсиям и другим платежам (например, детским пособиям, стипендиям студентов).

Если у населения нет денег, торговцы не поднимают цены. Так, в Москве за 2 года — с лета 1995 г. по лето 1997 г. — цены выросли примерно на 50%, в то время как в Иванове — лишь на 15%, а импортные товары на ивановских рынках стоят на $^1/_3$ дешевле, чем на московских (хотя эти импортные товары закупаются в Москве). Объяснить это можно тем, что экономическое положение в Иванове гораздо хуже, чем в Москве, ниже уровень доходов, больше безработных, что вынуждены учитывать торговцы.

Расчет индекса инфляции — вспомогательная задача, решение которой необходимо для приведения экономических характеристик к сопоставимому виду. Важнейшей задачей является расчет реальной заработной платы, равной частному от деления номинальной заработной платы на индекс инфляции. Известно, что цены на промышленные товары и на услуги, как правило, растут быстрее, чем на продовольствие. Поэтому рассчитываемые по продовольственным потребительским корзинам значения индексов инфляции дают оценку снизу роста потребительских цен и стоимости жизни в целом.

Минимальный прожиточный минимум оцениваем по методу М. Оршански. Коэффициент Энгеля принимаем C = 0,5. Этот метод основан на расчете стоимости минимальной продовольственной корзины и учете стоимостей остальных минимально необходимых затрат с помощью коэффициентов. Так, для «бедных семей» студентов Московского государственного института электроники и математики (технического университета) во время пробного бюджетного обследования в октябре—ноябре 1995 г. затраты на продовольствие составили 52% от всех расходов. Поэтому стоимость прожиточного минимума для них получим, приняв за 52% стоимость минимальной продовольственной корзины ИВСТЭ, т.е. умножив ее стоимость на 1/0,52 = 1,92.

Метод М. Оршански предполагает, что структура затрат практически не меняется. Однако, как уже отмечалось, цены на промышленные товары и на услуги растут быстрее, чем на продовольствие. Поэтому замена 1,92 на 2,00 представляется обоснованной. Полученные значения (на май 1997 г. — 700 тыс. руб. в месяц на человека) хорошо согласуются с уже цитированными данными Московской федерации профсоюзов (750 тыс. руб.). Отметим, что для всей совокупности семей, чьи бюджеты были обследованы в 1996 г., затраты на продовольствие составили 42%, т.е. для них коэффициент Оршански равен 1/0,42 = 2,38.

Средняя (начисленная) заработная плата в Москве составляла в декабре 1996 г. 1,12 млн руб. (в России — 0,84 млн руб.). В сопоставлении со сказанным ранее (с учетом логнормального характера функции распределения доходов и наличия детей) это означает, что даже в Москве, по крайней мере, половина семей живет ниже прожиточного уровня. В 1990 г. средняя зарплата превышала прожиточный минимум в 5,5 раз, а в 1997 г. — лишь в 1,2 раза (по России), т.е. уровень жизни упал в среднем в 4,6 раза и весной 1998 г. соответствовал концу 1950 — началу 1960-х гг. За август—сентябрь 1998 г. корзина ИВСТЭ подорожала в 1,5 раза (а средняя зарплата практически не изменилась), следовательно, уровень жизни упал уже в 7 раз и по покупательной способности зарплаты рядовые граждане «приблизились» к возможностям начала 1950-х гг.

Переход к сопоставимым ценам необходимо использовать также при расчете таких макроэкономических характеристик, как ВВП, объем бюджетных ассигнований и т.п. С учетом сказанного ранее можно утверждать, что экономика России с 1990 по 1998 гг. была «сокращена» в 4—6 раз, что соответствовало сдвигу назад по времени на 35—45 лет.

Материалы описанного исследования ИВСТЭ были опубликованы в 1998—1999 гг. в работе [89].

Инфляция в XXI в. Использование одной и той же потребительской корзины обеспечивает возможность сопоставления результатов расчетов за различные временны́е периоды. Этим работы ИВСТЭ выгодно отличается от подхода официальной статистики. Как известно, Госкомстат России (ныне — Росстат) в 1993—2008 гг. из конъюнктурных соображений неоднократно менял состав потребительской корзины и объемы потребления входящих в нее товаров. Однако в начале XXI в.

потребительская корзина официальной статистики мало отличалась от нашей. Здравый смысл восторжествовал — статистическое ведомство решило исходить из тех же разработок специалистов-диетологов РАМН, на которые мы опирались еще в 1993 г. На основе наших исследований инфляции была составлена глава 7 учебника [89] и соответствующий раздел учебного курса. Следующий шаг был сделан ИВСТЭ весной 2004 г. [92].

Студенты факультета «Инженерный бизнес и менеджмент» МГТУ им. Н.Э. Баумана в течение четырех месяцев один раз в месяц собирали данные о ценах и рассчитывали индекс инфляции. В целях сопоставимости результатов студенты все четыре раза собирали данные на продукты одних и тех же конкретных наименований (марок, сортов) и в одних и тех же торговых организациях. По данным за февраль, март и апрель 2004 г. методом наименьших квадратов строился точечный и непараметрический интервальный прогноз на май, который затем сопоставлялся с реальностью.

Собранные данные позволили изучить разброс значений индекса инфляции в зависимости от конкретных мест сбора данных. Рассмотрим значения индекса инфляции $I(t_1, t_2)$ на текущий момент t_2 = 14 мая 2004 г., соответствующие базовому моменту t_1 = 14 марта 1991 г.

Индексы инфляции в Москве в мае 2004 г.: 39,10; 40,50; 40,56; 40,70; 41,56; 41,73; 44,03; 47,18; 47,18; 47,30; 48,40; 49,27; 51,45; 52,67; 52,70; 53,04; 54,60; 55,00; 55,01; 55,33; 55,62; 56,40; 57,15; 57,29; 57,65; 57,72; 57,80; 58,26; 58,40; 59,59; 62,43.

Индексы инфляции в Московской области в мае 2004 г.: 44,02; 48,11; 50,40; 51,02; 51,08; 54,12; 54,12; 55,65; 57,32.

Статистические характеристики для двух приведенных выборок индексов инфляции содержатся в табл. 6.13. Они показывают, что индекс инфляции — это не число, а типичная нечисловая экономическая величина [78]. Индекс инфляции в Москве можно описать интервалом (39,1; 62,43), а в Московской области — интервалом (44,02; 57,32).

Таблица 6.13

Результаты статистической обработки данных об инфляции (май 2004 г.)

Статистические характеристики	Москва	Подмосковье
Минимум	39,1	44,02
Максимум	62,43	57,32
Объем выборки	31	9
Выборочное среднее арифметическое	51,47	51,76
Среднее квадратичное отклонение	6,80	4,08

Нечисловой характер индекса инфляции позволяет тем не менее сделать полезные для практического применения выводы:

1) в мае 2004 г. индекс инфляции был равен приблизительно 50, т.е. 50 руб. мая 2004 г. по своей покупательной способности соответствовали 1 руб. марта 1991 г.
2) индексы инфляции в Москве и Московской области практически совпадают.

В мае 2004 г. в Москве стоимость минимальной продовольственной корзины оценивается как (27 руб. 11 коп.) × 51,47 = 1400 руб., а прожиточный минимум — как 2800 руб. в месяц.

Индекс инфляции — это эконометрический инструмент, позволяющий доказательно обсуждать и решать те или иные экономические проблемы, например проблему соотношения зарплаты и прожиточного минимума [89]. Средства массовой информации часто рассматривают эту тематику. К сожалению, не всегда обсуждение является доказательным, а выводы — обоснованными. Так, в статье [18] утверждается, что мы в конце 2003 г. «живем, как в 1985 году». Это не так.

Сравним уровни жизни в 1985 г. и в 2003 г. Поскольку цены на основные продовольственные товары до марта 1991 г. не росли, можно признать, что индекс инфляции с 1985 г. по конец 2003 г. совпадает с таковым с марта 1991 г. по конец 2003 г., т.е., согласно [89], с достаточной для расчетов точностью равен 50 (см. табл. 6.6, 6.8). В статье [18] приведены значения средней зарплаты по стране — 199 руб. в 1985 г. и 5722 руб. — в конце 2003 г. Номинальная зарплата выросла в 29 раз, а цены — в 50 раз. Значит, реальная зарплата сократилась в 1,7 раза. В 1985 г. средняя зарплата почти в 4 раза превосходила прожиточный минимум, а в 2003 г. — лишь в 2 с небольшим раза.

В 2004 г. среднестатистический гражданин России живет гораздо хуже, чем в 1985 г. Основную причину назвал Президент Российской Федерации В.В. Путин в Послании 2004 г. Федеральному собранию Российской Федерации: ВВП (в сопоставимых ценах) в 2003 г. меньше, чем в 1989 г. (динамика макроэкономических показателей России анализируется в [87, с. 285]). Большое значение имеет резко возросшая дифференциация доходов. Измеряющий ее децильный коэффициент увеличился за эти годы с 3 до 15; в развитых странах его значение составляет около 7.

Крупное исследование было проведено через три с половиной года. Студенты собрали данные о ценах и рассчитали индексы инфляции в Москве и Московской обл. за период с t_1 = 14 марта 1991 г. до t_2 = 26 ноября 2006 г. (табл. 6.14, 6.15). Обработка данных была проведена О.Ю. Проскуриной.

Таблица 6.14

Индексы инфляции в Москве в ноябре 2006 г.

46,14	46,44	48,26	49,21	49,27	49,71
50,29	51,05	51,07	52,52	53,64	53,75
53,83	54,36	54,68	55,07	57,16	57,83
57,83	58,7	59,11	59,12	60,41	60,41
60,53	60,57	63,81	65,9	68,01	72,15
72,15	72,15	72,15	72,23	73,3	83,61

Таблица 6.15

Индексы инфляции в Московской области в ноябре 2006 г.

39,84	49,15	52,58	58,65
63,51	66,09	68,09	69,18

В Москве индексы инфляции были рассчитаны по ценам в таких торговых организациях, как гипермаркеты «Ашан» и «Метро», супермаркеты SPAR и «Перекресток», другие магазины, рынки. Проверка на однородность двух выборок — индексов инфляции в гипермаркете «Ашан» и индексов инфляции в «других магазинах», не входящих в сети, — с помощью критерия Крамера — Уэлча [89] показала, что выборки однородны, а следовательно, их можно объединить в одну.

Слушатели программы «Топ-менеджер» (Мастер делового администрирования / MBA) Академии народного хозяйства при Правительстве Российской Федерации собрали данные о ценах и рассчитали индексы инфляции 2006 г. в ряде регионов Российской Федерации (табл. 6.16).

Таблица 6.16

Индексы инфляции по регионам России

№	Город, регион	Дата	Значение индекса
1	Владимир	22.02.07	44,5
		22.03.07	46,8
2	Иркутск	09.01.07	42,38
		09.02.07	42,97
3	Красноярск (1)	25.11.06	59,50
		30.01.07	61,77
4	Красноярск (2)	25.11.06	64,60
		08.02.07	66,86
5	Калужская область, г. Малоярославец	20.12.06	46,86
		10.02.07	48,51
6	Нижний Новгород	10.11.06	43,16
		21.05.07	47,0
7	Новосибирск	01.03.07	51,79
		01.05.07	53,74
8	Петропавловск-Камчатский	10.11.06	28,94
		25.01.07	32,96
9	Ростов-на-Дону (1)	01.02.05	51,99
		01.02.07	67,67
10	Ростов-на-Дону (2)	01.02.07	39,66
		01.03.07	46,63
11	Ростов-на-Дону (3)	01.02.07	43,83
12	Татарстан, г. Бавлы	10.11.06	33,72
		25.01.07	36,05
13	Томск	25.12.06	49,86
		01.02.07	51,03
14	Тюменская область, п. Боровский	Январь 07	38,37
		Март 07	41,27
15	Череповец	01.11.06	48,1
		20.01.07	54,3

Примечание. Несколько исследований, проведенных в одном городе, указаны под разными порядковыми номерами. Следует иметь в виду, что стоимости потребительской корзины ИВСТЭ в марте 1991 г. для разных регионов различаются, иногда существенно.

Статистические характеристики для выборок индексов инфляции, приведенных в табл. 6.14, 6.15, содержатся в табл. 6.17. Они показывают, что индекс инфляции имеет заметный разброс, это не число, а типичная нечисловая экономическая величина [78]. В соответствии с приведенными данными индекс инфляции в Москве можно описать интервалом (46,14; 83,61), в Московской области — интервалом (39,84; 69,18). Статистическую обработку данных, приведенных в табл. 6.16, проводить было бы необоснованно, поскольку регионы, в которых проводились исследования, не представляют собой представительную (репрезентативную) выборку из генеральной совокупности регионов России. Кроме того, различаются даты снятия информации о ценах. Поэтому для включения в посвященный Российской Федерации столбец отобрана лишь часть данных. Тем не менее табл. 6.16 дает предварительное представление о динамике цен в регионах России. В частности, подтверждается высказанное ранее утверждение о том, что официальные статистические органы систематические занижают индексы инфляции: приведенное в табл. 6.8 значение 39,194 меньше 24 из 29 индексов инфляции, замеренных слушателями Академии народного хозяйства.

Таблица 6.17

Результаты статистической обработки данных об индексах инфляции *I*(1990, 11.2006) в ноябре 2006 г.

Статистические характеристики	Москва	Подмосковье	Россия
Минимум	46,14	39,84	42,38
Максимум	83,61	69,18	64,6
Объем выборки	36	8	6
Выборочное среднее арифметическое	59,07	58,39	53,40
Среднее квадратичное отклонение	9,74	9,04	7,67

Судя по собранным данным, структура стоимости потребительской корзины в среднем по Москве сравнительно мало изменилась с марта 1991 г. по ноябрь 2006 г. (рис. 6.1).

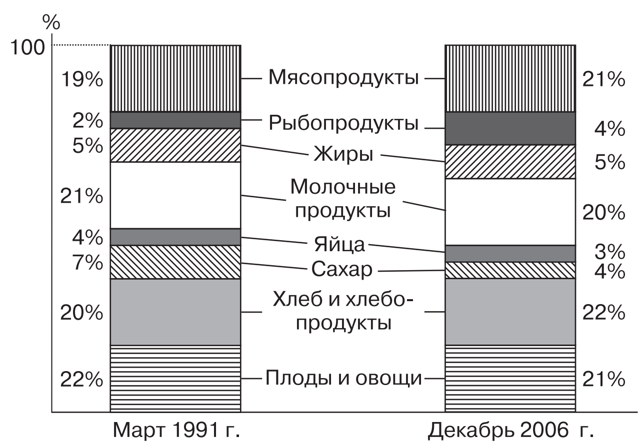


Рис. 6.1. Структура стоимости минимального набора продуктов питания

Нечисловой характер индекса инфляции позволяет тем не менее сделать полезные для практического применения выводы:

1) в ноябре 2006 г. индекс инфляции был равен приблизительно 60, т.е. 60 руб. в ноябре 2006 г. по своей покупательной способности примерно соответствовали 1 руб. марта 1991 г.;
2) индексы инфляции в Москве и Московской области практически совпадают и достаточно близки к индексам инфляции по большинству других регионов России;
3) в ноябре 2006 г. в Москве стоимость минимальной продовольственной корзины оценивается как (27,11 руб.) × 59,07 = = 1601,4 руб., а прожиточный минимум — как 3202,8 руб. в месяц (в соответствии с методом Оршански с коэффициентом Энгеля $C = 2{,}0$).

Отметим для сравнения, что, по методике и данным Росстата, стоимость минимального набора продуктов питания в среднем по России в конце ноября 2006 г. составила 1443,6 руб. в расчете на месяц.

В 2007—2008 гг. наблюдаем всплеск роста цен (табл. 6.17—6.19). Таблица 6.17 (Москва и Подмосковье) рассчитана по данным факультета «Инженерный бизнес и менеджмент» МГТУ им. Н.Э. Баумана, табл. 6.18 — слушателями программы «Топ-менеджер» (МБА) Бизнес-школы Академии народного хозяйства при Правительстве Российской Федерации, табл. 6.19 — слушателями Бизнес-школы МВА МИРБИС, т.е. действующими менеджерами высшего звена организаций и предприятий различных регионов России и Москвы.

Анализ приведенных в табл. 6.17—6.19 результатов измерений роста цен приводит к ряду интересных и практически полезных выводов [75]. В частности, в Екатеринбурге цены за полтора месяца выросли на 67%, в Красноярске за два с половиной месяца — на 32%, в Малоярославце за последний год измерений — на 53%. Данные Росстата — 11,9% за 2007 г. Средний результат (среднее арифметическое по соответствующему столбцу) по табл. 6.18 и 6.19 — 13,8% за последние месяц-два. Средний рост цен с 1990 г. — в 81,69 раза, т.е. на 1 руб. можно было купить в 1990 г. столько же, сколько на 81 руб. 69 коп. в январе 2008 г. А годом ранее индекс инфляции был заметно меньше — 59,07 (см. табл. 6.17). Рост на 38,3%. В три с лишним раза больше, чем по данным Росстата.

Таблица 6.18

Индексы инфляции в России на конец 2007 — начало 2008 гг.

№	Регион	Дата t_1	$I(90, t_1)$	Дата t_2	$S(t_2)$, мес.	$I(90, t_2)$	$I(t_1, t_2)$, %
1	Якутск	01.11.07	76,067	01.01.08	2 074,67	84,519	11,1
2	Хабаровск	30.11.07	66,87	30.12.07	1 759,21	68,79	2,9
3	Петропавловск-Камчатский	09.12.07	93,70	10.01.08	2 576,69	102,31	9
4	Малоярославец	12.06.07	55,09	10.01.08	2 153,49	84,21	53
5	Красноярск	08.12.07	82,81	12.01.08	2 486,29	97,58	18
6	Тюмень	05.12.07	75,36	10.01.08	2 297,28	82,94	10
7	Красноярск	26.10.07	47,77	10.01.08	1 594,42	63,17	32,2
8	Нижневартовск	01.11.07	54,45	01.01.08	1 949,92	62,513	14,8
9	Екатеринбург	01.12.07	50,28	10.01.08	2 142,82	84,10	67
10	Рязань	01.11.07	63,28	01.01.08	1 785,92	69,84	10
11	Москва	26.11.07	88,80	26.12.07	2 280,69	88,84	0,05
12	Новосибирск	01.12.07	84,09	13.01.08	2 167,30	84,75	0,78
13	Самара	28.11.07	67,5	08.01.08	1 762,17	68,9	2

Таблица 6.19

Индексы инфляции на конец 2007 — начало 2008 гг., Москва

№	Дата t_1	$I(90, t_1)$	Дата t_2	$S(t_2)$, мес.	$I(90, t_2)$	$I(t_1, t_2)$, %
1	28.12.07	115,08	30.01.08	2 992,42	117,01	1,68
2	28.12.07	73,61	29.01.08	1 944,66	76,04	3,30

Окончание

№	Дата t_1	$I(90, t_1)$	Дата t_2	$S(t_2)$, мес.	$I(90, t_2)$	$I(t_1, t_2)$, %
3	27.12.07	76,60	31.01.08	2 034,95	79,57	3,88
4	27.12.07	65,53	31.01.08	—	68,37	4,33
5	27.12.07	79,25	31.01.08	2 074,60	81,12	2,36
6	28.12.07	115,492	29.01.08	3 109,94	121,01	5,30
7	01.01.08	75,38	01.02.08	2 035,45	79,59	5,58
8	16.01.08	92,73	31.01.08	2 472,84	96,70	4,28
9	01.01.08	98,96	01.02.08	2 695,68	105,41	6,52
10	10.12.07	66,99	10.01.08	1 735,22	67,85	1,29

Отметим, что расхождение результатов расчетов по независимо собранной информации и данных официальной статистики частично объясняется тем, что в последние годы Росстат в очередной раз сменил потребительскую корзину. Это делает еще более неясной связь сообщаемых им численных значений инфляции с динамикой реальных экономических процессов (на эту неясность обращали внимание участники дискуссии, проведенной в ходе журналистского расследования). Как следствие констатируем, что каждое физическое и юридическое лицо может самостоятельно измерять рост цен с помощью методики, подробно изложенной в настоящей главе. Таким путем целесообразно бороться с монополией Росстата на результаты измерений инфляции.

Отметим еще одну особенность последних лет. Коэффициент Энгеля $C = 2{,}0$ при оценке прожиточного минимума был получен на основе бюджетного исследования середины 1990-х гг. Им можно пользоваться лишь при условии постоянства структуры расходов. Однако в последние годы резко растет доля расходов на оплату жилищно-коммунальных услуг. Это означает, что доля расходов на продовольствие у всех семей, и особенно у бедных, заметно снижается. Следовательно, коэффициент Энгеля $C = 2{,}0$ должен быть повышен, по экспертной оценке, до 3,0 (в 2009 г.).

Инфляция за 80 лет. Нет необходимости связывать возможность расчета индекса инфляции с каким-либо определенным интервалом времени и даже с определенным социально-экономическим строем. Можно формально вычислить индексы инфляции и за весьма длительные промежутки времени. Так, например, рост цен на основные продукты питания с 1913 г. по апрель 1994 г. представлен в табл. 6.20.

Таблица 6.20

Цены в 1913 г. и в апреле 1994 г. (руб./кг)

Продукт	Цена в 1913 г.	Цена в апреле 1994 г.
Хлеб пшеничный	0,05	740
Хлеб ржаной	0,03	400
Молоко	0,14	625
Сыр	0,40	6 150
Масло сливочное	0,55	5 100
Масло растительное.	0,13	2 300
Сметана	0,30	2 500
Говядина	0,23	2 760
Свинина	0,20	4 000
Баранина	0,17	2 000

Используя объемы потребления из потребительской корзины ИВСТЭ, получаем, что индекс инфляции за 1913—1994 гг. составил 11 297, или 1 129 600%. Подобные расчеты позволяют оценить реальное значение количественных экономических величин, используемых в публикациях разных лет.

Отметим, что представление о прожиточном минимуме меняется со временем. Еще 20 лет назад выход в Интернет и мобильная телефонная связь были уделом избранных, а сейчас эти услуги пора включать в прожиточный минимум. С другой стороны, во многих городах дрова перестали быть предметом первой необходимости, а потому нет нужды и возможности отслеживать цены на них. Необходимость модернизации потребительских корзин создает дополнительные проблемы по обеспечению сопоставимости результатов расчетов.

Потребительские корзины, включающие в себя промышленные товары и услуги, и соответствующие индексы инфляции. Недостаточно отслеживать только изменение цен на продовольственные товары. Необходимо также фиксировать инфляцию и в сфере коммунальных, транспортных, медицинских, образовательных и других услуг, а также анализировать цены на промышленные товары широкого потребления. Рост цен в этих областях достаточно заметен (если в 1990 г. проезд в метро в Москве обходился в 5 коп., то в ноябре 1995 г. он стоил 1000 руб., в феврале 1999 г., после деноминации в 1000 раз — 4 руб., в 2001 г. разовая поездка обходилась в 5 руб., а в 2009 г. — уже в 22 руб.). Темпы роста цен на те или иные промышленные товары и услуги не всегда совпадают с темпами ростом цен на продовольственные товары. Например, наблюдалось подорожание хлебобулочных товаров примерно в 2000 раз за три года (1991—1994), а цены на компьютерные товары выросли за это время в среднем только в 80 раз.

Заметная часть доходов каждой семьи идет на оплату коммунальных услуг и покрытие расходов на транспорт и связь. Необходимо учитывать расходы на услуги прачечной, парикмахерской, на ремонт обуви и т.д. Увеличиваются расходы на удовлетворение культурных потребностей из-за роста цен на книги, журналы, газеты, билеты в театры и кино, спортивный инвентарь и т.д. С течением времени подобные расходы конкретных физических лиц могут и сокращаться из-за прекращения покупок книг, журналов, газет, прекращения походов в театры и т.д.

При подсчете индекса инфляции по промышленным товарам возникает ряд трудностей. Например, наблюдается разброс цен по торговым точкам или имеет место временное отсутствие в магазинах некоторых товаров. Меняется мода, многие виды одежды выходят из употребления, вместо них появляются новые. То же самое, в связи с развитием техники, происходит и с товарами длительного пользования (когда-то не было телевизоров, холодильников, стиральных машин, железных дорог и самолетов). Пока еще мы можем пользоваться отдельными бесплатными услугами в области медицины и образования, но скоро, очевидно, и это будет платным, по крайней мере частично.

Для того чтобы подсчитать индекс инфляции по достаточно обширной потребительской корзине, включающей не только продовольственные товары, но и одежду, товары длительного пользования, услуги и т.п., необходимо иметь соответствующие нормы потребления. Определить их трудно. (При нормативном подходе к экономическим явлениям — откуда взять нормы? При позитивном — как в нестабильной ситуации замерить потребительские бюджеты?) Поэтому в настоящей главе мы ограничились индексами инфляции, рассчитанными для продовольственной потребительской корзины. Индекс инфляции можно считать не только для Москвы в целом, но и для отдельных ее районов и даже для покупателей отдельных магазинов — достаточно измерить соответствующие цены; не только для населения в целом, но и для отдельных слоев и даже отдельных семей — достаточно задать соответствующие потребительские корзины.

Как уже отмечалось, в 2009 г., в частности в связи с резким ростом стоимости жилищно-коммунальных услуг, коэффициент 2,0 в методе Оршански расчета прожиточного минимума представляется заниженным. Адекватное значение может быть получено в результате анализа результатов бюджетного обследования подобно тому, что было проведено ИВСТЭ в 1995 г. Альтернативный подход состоит в использовании иной потребительской корзины, например предусмотренной в Федеральном законе «О прожиточном минимуме в Российской Федерации» (в редакции федеральных законов РФ от 27.05.2000 № 75-ФЗ, от 22.08.2004 № 122-ФЗ).

Инфляция и валовой внутренний продукт. Характеристики экономического положения страны — ВВП, ВНД и другие — рассчитываются в текущих ценах. Для перехода к неизменным ценам их надо поделить на индекс инфляции (т.е. умножить на дефлятор). В 2 раза занизишь индекс инфляции — в 2 раза завысишь ВНП, ВВП, ВНД и иные макроэкономические характеристики.

По данным Правительства Российской Федерации, к концу 1998 г. ВВП составил 55,7% от уровня 1990 г. (динамика макроэкономических

показателей России анализируется в [76, с. 285]). Падение больше, чем в Германии в результате разгрома фашизма. И это по официальным данным! Используя же коэффициент занижения инфляции со стороны Госкомстата России, равный 2, получаем более реальную цифру — 25% от уровня 1990 г. Падение в 4 раза! Эта оценка близка к выводам ряда специалистов, не зависимых от правительства.

Напомним, что номинальный ВВП исчисляется в текущих рыночных ценах. Чтобы определить реальный ВВП, необходимо выразить его в сопоставимых ценах базисного года. Для этого применяется так называемый *дефлятор ВВП*, т.е. индекс инфляции, который отражает изменение среднего уровня цен самой широкой группы товаров и услуг за определенный период, охватывающий все составляющие ВВП. Расчеты проводят с помощью так называемой системы национальных счетов.

Нет ничего удивительного в том, что дефлятор ВВП отличается от индекса инфляции Росстата. Так, дефлятор ВВП за 2006 г. по отношению к ценам 2005 г. составил 15,4%, в то время как индекс инфляции Росстата за 2006 г. равен 9%. Разные корзины — разные результаты.

Виды инфляции. Эконометрика описывает инфляцию. Причины инфляции — это предмет иных экономических наук. Однако несколько слов сказать об этом полезно.

Всегда говорят об *инфляции спроса*. Это ситуация, когда у населения много денег, которые оно хочет истратить, а товаров мало. Тогда цены растут. Либо непосредственно, либо через механизм «черного рынка».

Другой вид инфляции — *инфляция издержек*. Производитель вынужден повышать цену на свою продукцию, потому что его поставщики повышают цены на собственную продукцию. Этот порочный круг очень трудно разорвать.

Третий вид инфляции — *административная инфляция*. Цены повышает государство, естественно, на то, что оно контролирует. Например, с августа по декабрь 1998 г. курс доллара США был поднят примерно в 4 раза. Последствия были понятными: адекватный подъем цен на импортные товары, потом рост цен на продукцию, для изготовления которой использовались импортные комплектующие, а затем и рост цен на чисто отечественную продукцию. В результате инфляция за год составила более 80%.

Ранее уже приводились примеры административного регулирования цен. Политика государственных органов в области энергетики, транспорта, экспорта и импорта, налогообложения и других сфер государственного регулирования экономики оказывает непосредственное влияние на инфляцию.

Заключительные замечания. Нобелевский лауреат по экономике Василий Васильевич Леонтьев (1905—1999) подсчитал, что лишь 1% ученых-экономистов анализирует вновь собранные данные, 30% используют данные, приведенные в публикациях предшественников, а остальные в своих рассуждениях вообще не обращаются к реальному миру [39]. Настоящая глава составлена на основе работ ИВСТЭ, относящихся к тому 1%, о котором писал В.В. Леонтьев.

Судя по опыту двух последних десятилетий, инфляционные процессы стали постоянной составляющей отечественной экономической жизни, и экономистам, менеджерам, инженерам различных специальностей придется учитывать их свойства в своей работе. В настоящей главе рассмотрены основы эконометрической теории инфляции. Однако не все проблемы раскрыты достаточно подробно. Кратко рассмотрим некоторые из них.

Прогнозирование индекса инфляции осуществляется с помощью методов наименьших квадратов (см. главу 5), экспертных технологий (часть III), в том числе основанных на сценарном подходе, и различных иных процедур, разработанных в организационно-экономическом моделировании. Обратим внимание на периодическую составляющую во временно́м ряду индексов инфляции. Темп роста цен максимален в зимние месяцы (декабрь—январь), затем постепенно уменьшается до минимума в летние месяцы (июль—август), иногда переходя в дефляцию, затем снова растет. Непараметрический метод выделения периодической составляющей временного ряда рассмотрен в [89, подраздел 6.3], [75, подраздел 10.2], а также — иной подход — в главе 5 ранее.

Стоимости потребительской корзины ИВСТЭ на один и тот же момент времени в наших публикациях, как мог заметить внимательный читатель, несколько отличаются. В этом нет ничего странного, так как исходные цены на продукты также отличались. Строго говоря, цены, стоимости потребительских корзин, индексы инфляции и многие другие экономические величины следовало бы считать нечисловыми данными [78; 89, подраздел 1.5], например интервальными или нечеткими. Развитие нечисловой экономики — перспективное направление научных исследований.

Неоднозначность выбора потребительской корзины, приводящая к неоднозначности индекса инфляции, порождает естественный вопрос: можно ли описать рост цен однозначно, т.е. полностью определенной функцией времени? Строгий ответ известен — нет, нельзя.

Еще в 1930-е гг. В.В. Леонтьев показал, что однозначно можно сравнивать только состояния экономик, имеющих одинаковую отраслевую структуру, так что описывающие их вектора (объемов производства по отраслям) отличаются только множителем [39]. Реально таких двух экономик не существует. Каждый год структура экономики меняется. Поэтому, строго говоря, нельзя сравнивать состояния экономик разных стран и даже состояния экономики одной и той же страны в разные годы. Этому же феномену посвящена теорема профессора В.В. Подиновского: любое изменение коэффициентов весомости единичных показателей качества продукции приводит к изменению упорядочения некоторых пар изделий по средневзвешенному показателю (см. главу 13).

Однако реально мы, несмотря на теоретический запрет, сравниваем экономическое положение в разные годы, зная, что это сравнение проводится с некоторой степенью условности, допустимой в рассматриваемых постановках прикладных задач. Точно на тех же основаниях мы должны принимать во внимание рост цен, выражаемый тем или иным индексом инфляции.

Мы почти не затрагивали историю инфляции. Наиболее быстро цены росли в Германии после Первой и Второй мировых войн и в СССР после Гражданской войны. Немецкий писатель Эрих Мария Ремарк в своем романе «Черный обелиск» описывает Германию 1923 г.: «Доллар стал неистовствовать, он подскакивает ежедневно уже не на тысячи и десятки тысяч, а на сотни тысяч марок. Позавчера он стоил миллион двести тысяч, вчера — миллион четыреста. Ожидают, что завтра он дойдет до двух миллионов, а в конце месяца — до десяти. Рабочие получают теперь заработную плату два раза в день — утром и под вечер, и каждый раз им дают получасовой перерыв, чтобы они успели сбегать в магазины и поскорее сделать покупки — ведь если они подождут до вечера, то потеряют столько, что их дети останутся полуголодными» [105, с. 420]. Рассмотрение методов выхода из инфляции находится вне рамок настоящего учебника.

Знание динамики индекса инфляции повышает обоснованность принятия хозяйственных решений. Отслеживание изменения индекса инфляции полезно и одновременно доступно всем юридическим и физическим лицам. Трудоемкость расчета одного значения индекса инфляции не превосходит 4 ч. Проведение такой работы обеспечивает связь обучения экономическим и управленческим дисциплинам с реальной экономической жизнью и может быть рекомендовано на всех уровнях экономических дисциплин — от средней школы до послевузовского образования. Полученные по независимо собранной информации оценки инфляции, рассмотренные в настоящей главе, используются в научных исследованиях и учебном процессе различных образовательных структур, а также в производственной деятельности предприятий и организаций, например на Магнитогорском металлургическом комбинате.

Контрольные вопросы

1. Каков будет индекс инфляции с 14.03.1991 г. по 14.03.2001 г. с учетом следующих потребительской корзины и цен:

Номенклатура, годовые нормы потребления и цены (руб.)

№ п/п	Продукт питания	Годовая норма, кг	Цена	
			на 14.03.1991 г.	на 14.03.2001 г.
1	Хлеб ржаной	65,3	0,2	10
2	Столовые корнеплоды	40,6	0,2	9
3	Колбаса докторская	0,4	2,3	95
4	Молоко, кефир	110,0	0,32	17
5	Сметана, сливки	1,6	1,7	50
6	Маргарин	6,3	1,2	35

2. Если гражданин Иванов в марте 1991 г. получил 200 руб., а в марте 2001 г. — 5000 руб., то во сколько раз изменился его доход? Увеличился он или уменьшился? (Используйте индекс инфляции из задачи 1.)

3. Если за январь индекс инфляции составил 50%, а за февраль — 200%, то чему равен индекс инфляции за два месяца? Каков средний темп (уровень) инфляции?

4. Каков будет текущий курс доллара США в ценах марта 1991 г. (индекс инфляции можно принять равным 100)?

5. Какова будет динамика индекса инфляции в России?

6. Почему для определения индекса инфляции (в процентах) за два года нельзя складывать индексы инфляции за первый год и за второй год, выраженные в процентах?

Темы докладов и рефератов

1. Место индексов инфляции в системе экономических индексов (в сравнении с индексами Ласпейреса, Пааше, И. Фишера).
2. Теоремы умножения (в случае четырех и более моментов времени) и сложения (для групповых индексов инфляции), их доказательства и использование.
3. Прогнозирование индекса инфляции: методы, практическая реализация, использование для принятия управленческих решений.
4. Учет инфляции при проведении анализа финансово-хозяйственной деятельности предприятия.
5. Обеспечение сопоставимости результатов расчетов при модернизации потребительской корзины.
6. Влияние инфляции на хозяйственную жизнь.
7. Методы выхода из инфляции.

Часть III

ЭКСПЕРТНЫЕ ТЕХНОЛОГИИ ПРИНЯТИЯ РЕШЕНИЙ

ГЛАВА 7
ПРОЦЕДУРЫ ЭКСПЕРТНЫХ ОЦЕНОК

7.1. ВИДЫ ЭКСПЕРТНЫХ ОЦЕНОК, ИНДИВИДУАЛЬНЫЕ ЭКСПЕРТНЫЕ ОЦЕНКИ[1]

Экспертные оценки бывают индивидуальные и коллективные. *Индивидуальные оценки* — это оценки одного специалиста. Например, преподаватель единолично ставит на экзамене оценку студенту. Врач ставит диагноз больному и назначает лечение. Инспектор ГИБДД экспертно оценивает соблюдение правил дорожного движения водителем и выписывает штраф за нарушение правил.

Но в сложных случаях заболевания или при угрозе отчисления студента за плохую учебу обращаются к *коллективному* мнению *экспертной комиссии*. Классический пример коллективной экспертной оценки — решение суда присяжных. По простым делам судья принимает решение единолично, при рассмотрении тяжких преступлений законодательством предусмотрена возможность участия в принятии решений комиссии экспертов — присяжных заседателей.

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 главнокомандующий русской армией М.И. Кутузов поставил вопрос: «Ожидать ли неприятеля на позиции, чтобы дать ему сражение или сдать Москву без сражения?»

Работа экспертной комиссии может быть растянута во времени. Например, лечащий врач может отправить пациента на обследование врачам-специалистам, дать распоряжение сдать различные анализы,

[1] Согласно англо-русскому словарю *expert* — это специалист. Однако в русском языке слово «эксперт» приобрело дополнительные нюансы. Под экспертом понимают не просто специалиста (например, выпускника вуза), а только такого, кто обладает высокой квалификацией, и, кроме того, умеющего использовать свою интуицию для решения поставленных перед ним задач, например для диагностики, прогнозирования, выбора варианта технического или управленческого решения.

Ударение в слове «эксперт», как и в словах «маркетинг» и «творог», можно ставить как на первый слог, так и на второй. Оба варианта признаются нормой. Ударение на первый слог соответствует английскому языку, ударение на второй слог больше подходит для русского языка.

сделать флюорографию и т.п. Собрав мнения экспертов (в данном случае врачей-специалистов) и проанализировав объективные данные, лечащий врач формулирует окончательное решение, выражающее мнение всей экспертной комиссии.

Индивидуальная экспертная оценка может потребовать от специалиста выполнения большого объема работы, например подготовка рецензии на рукопись книги или заключения оппонента о диссертации, представленной к защите на соискание ученой степени. Обычно эксперт должен следовать тем или иным правилам, приведенным в нормативной и методической документации по определенному виду экспертной деятельности. Например, при оценке диссертации эксперт должен исходить из нормативных документов Высшей аттестационной комиссии Российской Федерации.

В структуры государственной власти постоянно поступают научно-технические проекты, подготовленные различными организациями и отдельными гражданами. По каждой заявке требуется принять решение о целесообразности осуществления проекта и необходимом для этого содействии со стороны структур государственной власти (финансировании, организационных решениях).

Первый шаг — проект направляется на экспертизу. Эксперт Российского исследовательского научно-консультационного центра экспертизы (РИНКЦЭ) получает следующий документ.

Вопросы, которые должны быть отражены в заключении эксперта

1. Актуальность проекта.
2. Краткая характеристика положения в данной области в стране и за рубежом.
3. Научное значение проекта.
4. Научная новизна предлагаемых решений.
5. Прикладное значение проекта.
6. Новизна предлагаемых технических (технологических) решений.
7. Существующие отечественные и зарубежные аналоги (марка, тип, фирма, страна).
8. В чем заключается преимущество предлагаемых решений по сравнению с существующими в данной области в стране и за рубежом.
9. Сравнительные данные экономических показателей объекта и его аналогов (в сопоставимом виде).

10. Оценка потенциала разработчика:
- наличие научно-технического задела в данной области и в чем он выражается;
- наличие научно-производственной базы.

11. Обоснованность стоимости работ, оценка структуры затрат.

12. Реальность достижения поставленных целей:
- в предлагаемые сроки;
- предлагаемыми способами (методами) и ресурсами.

13. Возможность серийного освоения предлагаемого проекта.

14. Последствия создания и использования проекта:
- научные и научно-технические;
- экологические;
- гуманитарные;
- экономические;
- социальные.

15. Выводы:
- необходимость реализации проекта (полная, частичная);
- целесообразность финансирования (в целом, частично);
- рекомендации эксперта.

Мнение эксперта должно быть выражено в специальном документе — *заключении*. На все 15 приведенных вопросов эксперт должен ответить в своем заключении. Ясно, что этот документ должен быть достаточно объемным, а подготовка его трудоемка.

Когда нужна формализация мнений экспертов? Цели экспертизы могут быть различны. Так, отзыв официального оппонента заканчивается выводом о том, соответствует или нет рассмотренная им диссертация требованиям ВАК Российской Федерации. Рецензент научного журнала делает в конце своего заключения вывод о том, может или нет данная статья быть опубликована в журнале. В этих двух случаях нет необходимости сравнивать между собой различные объекты экспертизы.

Однако часто необходимо проводить такое сравнение. Научно-технические или инвестиционные проекты нельзя рассматривать отдельно друг от друга, поскольку ограничено суммарное финансирование, выделенное на всю совокупность проектов.

Насколько подходят для сравнения объектов экспертизы обширные заключения, подготовленные различными экспертами? С одной стороны, эти заключения содержат результаты высококвалифицированного труда по оценке содержания проектов. С другой — написанные

в свободной манере заключения не всегда позволяют сопоставить между собой отдельные характеристики проектов. Поэтому эксперты РИНКЦЭ заполняют еще один формализованный документ.

Карта оценки объекта экспертизы

Научная значимость.

1. Исключительно высокая.
2. Значительная.
3. Невысокая.
4. Неопределимая (в настоящее время).
5. Отсутствует.

Практическая значимость.

1. Исключительно высокая.
2. Значительная.
3. Невысокая.
4. Неопределимая (в настоящее время).
5. Отсутствует.

Научная новизна, оригинальность.

1. Не имеет аналогов.
2. Нет аналогов в стране, есть за рубежом.
3. Нет аналогов за рубежом, есть в стране.
4. Есть сведения об отдельных отечественных и зарубежных аналогах.
5. Научная новизна отсутствует.

Методы и способы достижения цели.

1. Новые.
2. Современные.
3. Традиционные.
4. Устаревшие.
5. Неадекватные.

Потенциал исполнителей в рассматриваемой области.

1. Достаточный.
2. Недостаточный в части научного задела (опыта работы).
3. Недостаточный в части материально-технической (лабораторно-экспериментальной) базы.
4. Недостаточный в части состава исполнителей.
5. Данных для оценки недостаточно.

Срок работы.

1. Реальный.
2. Завышенный.
3. Заниженный.
4. Данных для оценки недостаточно.

***Стоимость работ (объем финансирования)*.**

1. Приемлемая.
2. Завышена.
3. Занижена.
4. Данных для оценки недостаточно.

***Рекомендуемый приоритет осуществления*.**

1. Работа первостепенной важности.
2. Работа высокой важности.
3. Работа представляет определенный интерес.
4. Работа представляет незначительный интерес, но заслуживает поддержки при наличии достаточных средств.
5. Работа поддержки не заслуживает.

Дата__________ Эксперт_______________ Подпись___________
(Ф.И.О.)

При заполнении «Карты оценки объекта экспертизы» ничего писать не надо, следует лишь обвести номера тех пунктов в каждом из разделов, которые соответствуют мнению экспертов. В разделе «Потенциал исполнителей» могут быть обведены несколько номеров, в остальных разделах — по одному. По «Карте оценки объекта экспертизы» легко сравнивать мнения экспертов между собой, а также сопоставлять различные объекты экспертизы.

Обратим внимание, что в конце «Карты оценки объекта экспертизы» предусмотрена подпись эксперта. Это связано с тем, что эксперт несет ответственность за свое заключение — административную, материальную, гражданско-правовую, уголовную. Экспертные исследования принципиально отличаются от маркетинговых и социологических, в которых подчеркивается анонимность опрашиваемых [89, глава 2].

Различные типы вопросов. В экспертных исследованиях, а также в выборочных маркетинговых и социологических опросах используют три типа вопросов — закрытые, открытые и полузакрытые, они же полуоткрытые. При ответе на закрытые вопросы можно выбирать лишь из заранее сформулированных составителями анкеты вариантов ответа. В качестве ответа на открытый вопрос опрашиваемого просят изложить свое мнение в свободной форме. Полузакрытые, они же полуоткрытые вопросы занимают промежуточное положение — кроме выбора среди перечисленных в анкете вариантов, можно добавить свои соображения. Ясно, что «Вопросы, которые должны быть отражены в заключении эксперта», являются открытыми, а «Карта оценки объекта экспертизы» состоит из закрытых вопросов.

Каждый из этих типов вопросов имеет свои достоинства и недостатки. Преимущество открытых вопросов состоит в том, что эксперт может свободно высказать свое мнение так, как сочтет нужным. Их недостаток — в сложности сопоставления мнений различных экспертов. 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 эксперт. 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 эксперта «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 именно — эксперт или организатор экспертизы — будет шифровать ответы.

Отметим, что на этапе подготовки важного экспертного опроса проводят так называемое «пилотное» исследование — апробацию документов и процедур анализа ответов, которые будут собраны в ходе будущего опроса. В пилотном исследовании участвует небольшое число экспертов, цель работы которых — проверить доступность задач опроса и документации пониманию экспертов, работоспособность расчетных процедур, уточнить формулировки вопросов и способы сбора и анализа экспертных мнений. В частности, в рамках пилотного исследования может быть проведена предварительная экспертиза, специально посвященная отработке перечня и формулировок вопросов.

7.2. ОЦЕНКА И ВЫБОР ВАРИАНТОВ С ПОМОЩЬЮ ЭКСПЕРТОВ, КОЛЛЕКТИВНЫЕ ЭКСПЕРТНЫЕ ОЦЕНКИ

Рассмотрим несколько процедур коллективных экспертных оценок, начиная с простейших, при этом вводя и обсуждая используемые в дальнейшем понятия.

Оценка номеров в КВН. Простейший пример коллективных экспертных оценок — оценка номеров в известной игре КВН (Клуб веселых и находчивых). Экспертной комиссией является жюри. Просмотрев номер, каждый из членов жюри поднимает планшет со своей оценкой. Затем симпатичная девушка (технический работник, не член жюри) вычисляет среднюю арифметическую оценку, которая и объявляется как коллективное мнение жюри (ниже увидим, что такой подход некорректен с точки зрения теории измерений). Обратим внимание на эту девушку (технического работника), которая после обработки экспертных мнений выставляет оценку на стенд, делая результаты экспертизы доступными всем желающим. Она представляет коллектив тех, кто обеспечивает организацию и проведение экспертизы. Этот коллектив называют «рабочей группой» (РГ) [89, гл. 12] или «группой сопровождения» [115].

Таким образом, два основных объекта рассмотрения в настоящем учебнике — это *экспертная комиссия* (ЭК) и *рабочая группа* (РГ).

Фигурное катание. В фигурном катании процедура обработки оценок экспертов 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.

Пусть $X_1, X_2, ..., X_n$ — оценки n экспертов. При проведении КВН в качестве коллективной экспертной оценки используют среднее арифметическое всех n оценок:

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n}.$$

В фигурном катании нужно переставить элементы выборки в порядке возрастания (точнее, неубывания) и получить вариационный ряд $X(1) \le X(2) \le ... \le X(n)$, исключить минимум $X(1)$ и максимум $X(n)$, а затем в качестве коллективной экспертной оценки взять урезанное среднее арифметическое, т.е. среднее арифметическое оставшихся $(n-2)$ членов вариационного ряда:

$$X^* = \frac{X(2) + X(3) + ... + X(n-1)}{n-2}.$$

С точки зрения прикладной математической статистики, X^* — это робастная оценка теоретического среднего, нацеленная на борьбу с аномальными (резко выделяющимися) результатами наблюдений. Поскольку ясно, что аномальные результаты порождены внешними влияниями на судей фигурного катания, искажающими их профессиональные экспертные оценки, то простое изменение правил расчетов итоговой оценки (переход от среднего арифметического к урезанному среднему) позволяет уберечь экспертов от вызванных извне уклонений от решения поставленных перед ними задач.

Итак, правила обработки оценок экспертов существенно влияют на объективность выводов экспертной комиссии.

Экспертный выбор. Экспертные оценки часто используют 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; инвестиционных проектов для реализации — среди представленных и т.д.

Типовая ситуация такова. Заказчик формулирует технические требования к будущему изделию. Объявляется конкурс (тендер), итогом которого должен быть выбор той или иной разработки для серийного выпуска. Допущенные к конкурсу организации к заданному сроку представляют опытные образцы. Как правило, оказывается, что эти образцы несравнимы, каждый из них по каким-то важным показателям качества лучше других, а по другим важным показателям — хуже того или иного из остальных образцов. Например, у одного опытного образца дальность полета больше, у другого — расход топлива на 1000 км меньше, у третьего — потолок полета выше, у четвертого — броня крепче, у пятого — под крыльями можно дополнительно подвесить две ракеты. Какой стратегический бомбардировщик (из разработанных разными конструкторскими бюро и представленных на тендер) выбрать для серийного производства?

Задача экспертной комиссии — выбрать опытный образец для запуска в серийное производство. Есть два принципиально разных подхода к решению этой задачи.

Первый из них основан на ***сравнении образцов***. Например, каждый из экспертов упорядочивает образцы в соответствии со своими предпочтениями. Полученные от экспертов *упорядочения* (*ранжировки*) обрабатываются теми или иными математическими методами с целью расчета итогового мнения комиссии экспертов. В другом варианте организации экспертизы образцы предъявляют эксперту попарно для сравнения, математический анализ результатов *парных сравнений* позволяет найти итоговое мнение. В третьем варианте каждого эксперта просят выбрать три лучших образца и т.д.

Второй подход имеет целью соизмерить сравнительную важность различных показателей качества, построить интегральный показатель качества (рейтинговую оценку), с помощью которого можно упорядочить образцы по качеству (рассчитать *рейтинг* образцов). Пусть, например, выделено (с помощью предварительного экспертного исследования) m показателей качества. Для конкретного объекта экспертизы экспертная комиссия оценивает эти показатели $Y_1, Y_2, ..., Y_m$, затем РГ рассчитывает значение интегрального показателя качества

$$Y = a_1 Y_1 + a_2 Y_2 + ... + a_m Y_m .$$

На основе полученных значений Y можно выбрать наилучший образец, упорядочить образцы по качеству, указав рейтинг образцов, т.е. значения интегрального показателя, соответствующие образцам. Значения коэффициентов a_i (коэффициентов важности, весомости, значимости) обычно определяются с помощью той или иной экспертной процедуры.

Кроме аддитивной формы интегрального показателя, часто используют мультипликативный вариант этого показателя:

$$Z = \prod_{j=1}^{m} Y_j^{b_j} ,$$

в котором показатели степени b_j обычно также определяются экспертным путем.

Вопросы построения рейтингов подробно рассмотрены в главе 13.

Кроме задачи выбора наилучшего (с точки зрения экспертов) образца, описанные методы позволяют решить ряд иных практических задач, в частности задачу распределения финансирования. Пусть имеется ряд объектов экспертизы, нуждающихся в финансировании, например инвестиционных проектов или заявок на выполнение научно-технических проектов (работ). Естественно упорядочить объекты экспертизы по качеству (рентабельности, привлекательности и т.п.), а затем выделять необходимые объемы финансирования, начиная с наилучшего объекта. Тогда начальная часть вариационного ряда показателей качества будет соответствовать профинансированным объектам экспертизы, а заключительная — тем, кому финансирования не досталось.

На границе между этими двумя группами возможны нюансы. Например, объект экспертизы A нельзя профинансировать в необходимом объеме из-за недостатка средств, а вот на финансирование худшего, чем A, объекта экспертизы B средств достаточно. Тогда объект B будет финансироваться, а объект A — нет, вопреки рейтингу.

Военные советы как форма экспертной деятельности. С тех пор как люди научились говорить, проводились совещания специалистов. Поэтому можно сказать, что экспертным оценкам столько же лет, сколько человеческому обществу. Конечно, постепенно технологии экспертного оценивания развивались. Например, появилась идея *независимой экспертизы*, которую можно сопоставить с идеей разделения власти на законодательную, исполнительную и судебную ветви в предположении независимости ветвей власти.

Весьма важен *регламент* проведения заседания комиссии экспертов. Вот как в «Капитанской дочке» (глава X) А.С. Пушкин приводит слова, с которыми генерал, комендант Оренбурга, обратился к членам военного совета:

«Теперь, господа, — продолжал он, — надлежит решить, как нам действовать противу мятежников: *наступательно* или *оборонительно*? Каждый из оных способов имеет свою выгоду и невыгоду. Действие наступательное представляет более надежды на скорейшее истребление неприятеля; действие оборонительное более верно и безопасно.... Итак, начнем собирать голоса по законному порядку, то есть, начиная с младших по чину. Г-н прапорщик! — продолжал он, обращаясь ко мне. — Извольте объяснить нам ваше мнение».

Военный совет в данном случае — это собрание экспертов (военных специалистов). Председатель собрания четко поставил задачу: надо выбрать либо наступление, либо оборону. Обсуждение идет в однозначно заданном порядке — от младших к старшим. Младшие могут спокойно высказывать свои мысли, не боясь, что их предложения будут противоречить мнению старших. Старшие имеют возможность учесть высказанные аргументы и сделать свои выступления более обоснованными.

Важность соблюдения регламента проведения заседания экспертной комиссии становится особенно ясной при сопоставлении с распространенным в XVII в. местничеством. Бояре постоянно спорили, кто из них главнее и, следовательно, кто должен сидеть ближе к царю и говорить раньше и больше других. Заседание постоянно прерывалось схватками, иногда не только словесными, между его участниками. Повышению эффективности заседаний весьма способствовало введение Петром I системы чинов и регламентации служебных взаимоотношений в соответствии с нею. И в настоящее время общепринятой практикой является выбор (или назначение) в начале собрания председателя и секретаря и утверждение регламента.

Наиболее известный в истории России военный совет состоялся 1 сентября 1812 г. в Филях, вскоре после Бородинского сражения. Обсуждался вопрос: «Дать французам сражение под Москвой или оставить Москву без боя?» Решение должен был принять главнокомандующий фельдмаршал Кутузов. Обратим внимание, что военный совет, как и любая комиссия экспертов, — совещательный орган, а окончательные решения принимает тот, кому это поручено. В современной литературе такой человек обозначается как лицо, принимающее решение (ЛПР).

Большинство экспертов, рассказав о состоянии своих войск, высказалось за сражение. Однако, учитывая тяжелые потери русской армии, Кутузов (т.е. ЛПР) принял решение оставить Москву без боя. Аргументировал это решение Кутузов так: «Оставив Москву, мы сохраним армию; потеряв армию, мы потеряем и Москву, и Россию». И 2 сентября 1812 г. русские войска без боя оставили Москву, с ними ушла и половина московского населения (около 100 тыс. человек). Как известно, это решение Кутузова предопределило поражение Наполеона в войне и изгнание захватчиков.

Итак, ЛПР поступило вопреки мнению большинства экспертов. Значит ли это, что работа экспертной комиссии проделана впустую? Отнюдь! Собранная экспертами информация была использована ЛПР. Продемонстрированный генералами русской армии боевой дух, готовность сражаться с врагом также были учтены ЛПР наряду с теми соображениями, которые не могли знать эксперты и которые были приняты во внимание ЛПР.

Обсуждение регламента проведения заседаний и организации экспертного исследования в целом, взаимоотношений ЛПР и ЭК касается всех видов экспертных оценок, отнюдь не только военных советов.

7.3. ЭКСПЕРТНОЕ ПРОГНОЗИРОВАНИЕ

Перейдем к развитию экспертных исследований в XX в.

Кибернетика — основа управления. Большое влияние на развитие исследований в области управления в целом и менеджмента в частности оказало появления в 1948 г. книги американского математика Норберта Винера (1894—1964) «Кибернетика, или управление и связь в животном и машине». Через два года вышла его книга «Кибернетика и общество». Началось мощное научное движение, ключевые слова которого — кибернетика, исследование операций, системный анализ, математическое моделирование, оптимальное управление, экспертные оценки и др. Оно до сих пор определяет лицо современной науки об управлении. В нашей стране огромную роль в развертывании исследований по кибернетике сыграл академик АН СССР адмирал-инженер Аксель Иванович Берг (1893—1979). С 1950-х гг. до последних дней жизни он возглавлял Научный совет АН СССР по комплексной проблеме «Кибернетика».

Один из вождей отечественного кибернетического движения академик РАН Никита Николаевич Моисеев (1917—2001) в своей книге [54] приводит ряд фактов, позволяющих проследить историю кибернетических идей. В частности, он обращает внимание на книгу профессора Бронислава Трентовского «Отношение философии к кибернетике как

искусству управления народами», вышедшую в Познани в 1843 г. (за 105 лет до книги Н. Винера) на польском языке. Для образованных людей XIX в., знакомых с древнегреческим языком, слово «кибернетика» было вполне понятно. Оно означало систему взглядов, знаний, навыков, которой должен был обладать управляющий, для того чтобы эффективно управлять людьми и ресурсами, находящимися в его распоряжении. Большой вклад в кибернетику в целом и в теорию систем в частности внесли отечественные ученые — член Петербургской академии наук Евграф Степанович Федоров (1853—1919) и особенно Александр Александрович Богданов (настоящая фамилия — Малиновский) (1873—1928), деятель российского революционного движения, врач, философ, экономист, с 1926 г. организатор и директор Института переливания крови. Погиб, производя на себе медицинский опыт. Основное сочинение А.А. Богданова — трехтомная «Всеобщая организационная наука (тектология)». Первый том напечатан в 1913 г. Полностью книга выходила в 1925—1929 гг.

Многие идеи кибернетики были известны задолго до Н. Винера (хотя сам он об этом, скорее всего, и не догадывался). Почему же именно книга Н. Винера послужила толчком к развитию работ по теории управления, а не работы Трентовского, Федорова, Богданова? Одно из возможных объяснений в том, что «Кибернетика» Винера появилась вовремя, после Второй мировой войны, когда стали выделять большие ресурсы на развитие науки (это было реакцией правительств на продемонстрированную в Хиросиме и Нагасаки роль науки в практике).

После Второй мировой войны в рамках научного движения, включающего кибернетику, информатику, теорию управления, менеджмент и исследование операций, стала развиваться самостоятельная научно-практическая дисциплина — теория и практика экспертных оценок.

Метод Дельфи. Один из наиболее известных методов экспертных оценок — это метод Дельфи. Название дано по ассоциации с древним обычаем для получения поддержки при принятии решений обращаться в Дельфийский храм. Он был расположен у выхода ядовитых вулканических газов. Жрицы храма (пифии), надышавшись отравы, начинали пророчествовать, произнося непонятные слова. Специальные «переводчики» — жрецы храма — толковали эти слова и отвечали на вопросы пришедших со своими проблемами паломников. Те спрашивали, отправляться ли в морское путешествие, вступать ли в брак, заключать ли договор с тем или иным деловым партнером, начинать ли войну и т.д.

Технология экспертного оценивания состояла в следующем. Получив «заказ на экспертное прогнозирование», жрецы передавали его пифиям, выслушивали пророчества пифий, а затем толковали услышанное заказчику. С течением времени в храме накапливались пожертвования и памятные доски от тех, для кого прогнозы сбылись. Если же прогноз не осуществился, то сообщить об этом зачастую было некому — заказчик лежал на морском дне или был убит в битве, разорен и продан в рабство и т.п.

По традиции говорят, что Дельфийский храм находился в Греции. Но там нет вулканов. Видимо, он был в Италии — у Везувия или Этны, а сами описанные предсказания происходили в XII—XIV вв. Это вытекает из высшего достижения современной исторической науки — новой статистической хронологии.

В США в 1960-х г. 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 экспертов, 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 15 лет она использовалась не менее 40 тыс. раз. Это объяснялось впечатлением от беспрецедентного успеха предсказания даты высадки на Луну. Можно констатировать, что именно этот успех выдвинул методы экспертных оценок на роль самостоятельного научно-практического направления, с которым должны быть знакомы все инженеры и управленцы, а также деятели иных специальностей.

Средняя стоимость экспертного исследования по методу Дельфи — 5 тыс. дол. США, но в ряде случаев приходилось расходовать и более крупные суммы — до 130 тыс. дол.

Метод сценариев. Несколько в стороне от основного русла экспертных оценок лежит метод сценариев, применяемый прежде всего для экспертного прогнозирования.

Рассмотрим основные идеи технологии сценарных экспертных прогнозов.

Социально-экономическое или, скажем, экологическое 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. Вполне очевидно, что после первого тура президентских выборов 1996 г. о дальнейшем развитии событий можно было говорить лишь в терминах сценариев: если во втором туре победит Б.Н. Ельцин, то будет то-то и то-то, если же победит Г.А. Зюганов, то события пойдут так-то и так-то.

Метод сценариев необходим не только в социально-экономической или экологической областях. Например, 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 в целях 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 — прилет инопланетян, падение астероида, массовые эпидемии ранее неизвестных болезней и т.д. 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 ранее 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* (как говорят, при *ситуационном анализе*), в том числе анализа 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.

Мозговой штурм. 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 записываемое на магнитофон или видеокамеру заседание заканчивается и начинается второй этап мозгового штурма — анализ высказанных идей. Обычно за время дискуссии высказывается около 100 идей. Из них примерно 30 заслуживают дальнейшей проработки, 5—6 идей дают возможность сформулировать прикладные проекты, а 2—3 идеи оказываются в итоге приносящими полезный эффект — прибыль, перевод конфликта в сотрудничество, повышение экологической безопасности, оздоровление окружающей природной среды 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, сбивающие торпеду с курса.

7.4. ЭКСПЕРТНЫЕ ОЦЕНКИ НА СОВРЕМЕННОМ ЭТАПЕ

В настоящее время практически все виды трудовой деятельности так или иначе связаны с проведением экспертиз. Врачи и преподаватели, управленцы (менеджеры) и инженеры, юристы и экономисты — все они в той или иной степени эксперты. Классифицировать основные виды экспертной деятельности можно по областям конкретной профессиональной деятельности, а также по тем задачам, которые решают с помощью экспертных исследований.

По областям конкретной профессиональной деятельности выделяют, в частности, следующие виды экспертиз:

- строительную;
- медицинскую;
- судебную;
- экологическую, в том числе объектов недропользования;
- товароведческую;
- качества товаров;
- патентную;
- страховую;
- аудит;
- при оценке имущества, бизнеса, нематериальных активов и т.д. [40].

Экспертная деятельность в конкретных областях обычно регулируется соответствующими нормативными актами и осуществляется в соответствии с теми или иными методическими материалами. В дальнейших главах в качестве примера нормативного регулирования экспертной деятельности будем рассматривать Федеральный закон от 23 ноября 1995 г. № 174-ФЗ «Об экологической экспертизе».

При классификации по решаемым задачам выделяют оценочные и управленческие экспертизы [40].

Результатами ***оценочных экспертиз*** являются:

- численные оценки объектов (значений показателей, параметров, характеристик объектов);
- отнесение объектов экспертизы к тому или иному виду объектов, классу объектов, сорту;
- ранжирование объектов по тому или иному свойству, качеству, показателю, критерию;
- рейтинги, позволяющие определить численные значения, характеризующие сравнительную предпочтительность объектов экспертизы;
- индексы, позволяющие оценить (характеризующие) состояние объектов экспертизы,

- иные объекты числовой или нечисловой природы, используемые для оценивания объектов экспертизы (конкретные виды объектов рассматриваются в следующих главах учебника).

Примерами результатов оценочных экспертиз, в частности, являются:

- результаты определения победителей конкурсов, тендеров, подрядных торгов, иных соревнований;
- рейтинги организаций (промышленных предприятий, вузов, банков, страховых компаний), ценных бумаг, политических деятелей, бизнесменов и спортсменов;
- индексы (Доу — Джонса и др.), характеризующие движение курсов ценных бумаг на биржах.

Результатом ***управленческих экспертиз*** является подготовка рекомендаций и заключений на всех этапах цикла выработки, принятия и реализации управленческих решений. К их числу относятся экспертизы:

- при выработке стратегии и тактики (определении стратегических целей, приоритетов деятельности, планов, организационных структур, разработке бизнес-планов и т.д.);
- подготовке аналитических материалов и проведении ситуационного анализа, включая разработку прогнозов и сценариев;
- генерировании и отборе альтернативных вариантов решений;
- оценке альтернативных вариантов решений и определении наиболее предпочтительного из них;
- контроле хода реализации принятых решений;
- корректировке принятых ранее управленческих решений на основании оценки хода реализации принятых решений.

Конечно, эти перечни не являются исчерпывающими. Они позволяют составить представление о том, насколько разнообразны задачи экспертных оценок и области их практического применения.

Нельзя не согласиться с мнением Б.Г. Литвака, что экспертизы необходимы на всех стадиях управленческого цикла, в какой бы области деятельности ни принималось решение [40]. Без профессиональной экспертизы нет сегодня профессионально принятого решения!

Разработана 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 — далее). Не меньше существует и методов обработки ответов экспертов, в том числе весьма насыщенных математикой и компьютеризированных. В дальнейших главах книги на основе методологии, развитой в [91], будет рассмотрен ряд современных методов экспертных оценок (см. также [114]).

Контрольные вопросы

1. Какие экспертные оценки называют индивидуальными?
2. Почему необходима формализованная карта оценки объекта экспертизы?
3. Какие экспертные оценки называют коллективными?
4. В чем состоят задачи выбора вариантов с помощью экспертов?
5. Почему большое внимание уделяют регламенту проведения экспертных исследований?
6. Что представляет собой метод Дельфи в экспертном прогнозировании?
7. В чем суть метода сценариев?
8. Что называют мозговым штурмом?
9. В каких конкретных областях используют методы экспертных оценок?

Темы докладов и рефератов

1. Индивидуальное экспертное оценивание (на примере работы преподавателя).
2. Варианты коллективного экспертного оценивания в медицине.
3. Робастное оценивание в экспертизе.
4. Экспертные технологии распределения финансирования.
5. Технологии экспертного прогнозирования.
6. Метод сценариев и экспертная оценка рисков в инвестиционном менеджменте.
7. Экспертные технологии в технико-экономическом анализе.
8. Статистика нечисловых данных в оценочных экспертизах.
9. Управленческие экспертизы в контроллинге.

ГЛАВА 8
ОРГАНИЗАЦИЯ РАБОТЫ ЭКСПЕРТНОЙ КОМИССИИ

8.1. ОСНОВНЫЕ СТАДИИ ЭКСПЕРТНОГО ОПРОСА

Более подробно рассмотрим отдельные этапы типового экспертного исследования. Как показывает практический опыт, с точки зрения менеджера — организатора такого исследования — целесообразно выделять следующие стадии проведения экспертного опроса.

1. Принятие решения о необходимости проведения экспертного опроса и формулировка его цели ЛПР. Инициатива должна исходить от руководства, что в дальнейшем обеспечит успешное решение организационных и финансовых проблем. Очевидно, что исходный толчок может быть дан докладной запиской одного из сотрудников или дискуссией на совещании, но реальное начало работы — решение ЛПР. Цель экспертного исследования ЛПР может сформулировать по-разному, и от этой формулировки зависит выбор процедуры экспертизы.

2. Подбор и назначение ЛПР основного состава рабочей группы — РГ (обычно научного руководителя и ответственного секретаря). При этом научный руководитель отвечает за организацию и проведение экспертного исследования в целом, а также за анализ собранных материалов и подготовку заключения экспертной комиссии. Он участвует в формировании коллектива экспертов и выдаче задания каждому эксперту (вместе с ЛПР или его представителем). Научный руководитель — высококвалифицированный эксперт и признаваемый другими экспертами формальный и неформальный руководитель экспертной комиссии. Дело ответственного секретаря — ведение документации экспертного опроса, решение организационных задач. Назначение научного руководителя и ответственного секретаря оформляется распорядительным документом (приказом, постановлением и т.п.). Остальной состав РГ обычно формируется позже, в процессе развертывания исследования, причем по предложениям научного руководителя и ответственного секретаря.

3. Разработка РГ (точнее, ее основным составом, прежде всего научным руководителем и ответственным секретарем) ***и утверждение у ЛПР технического задания (ТЗ) на проведение экспертного опроса.*** На этой стадии решение о проведении экспертного опроса приобретает четкость во времени, финансовом, кадровом, материальном и организационном обеспечении. В частности, формируется костяк РГ со своей внутренней структурой. Обычно в РГ выделяются различные группы специалистов — аналитическая, эконометрическая (специалисты по методам анализа данных), компьютерная, по работе с экспертами (например, интервьюеров), организационная (конечно, возможно совмещение ролей — один и тот же сотрудник может и отвечать за выбор метода анализа экспертных мнений, и сам же проводить этот анализ). Очень важно для успеха, чтобы все перечисленные позиции были включены в ТЗ и утверждены ЛПР.

4. Разработка аналитической группой РГ подробного сценария (т.е. регламента, правил) проведения сбора и анализа экспертных мнений (оценок). Термин «сценарий» имеет примерно тот же смысл, что и в театре, и кинематографе. Сценарий включает в себя, прежде всего, анкеты и опросные листы (планы интервью), определяющие конкретный вид информации, которая будет получена от экспертов (например, слова, условные градации, числа, ранжировки, разбиения или иные виды объектов нечисловой природы). Довольно часто экспертов просят высказаться в свободной форме, ответив при этом на некоторые заранее сформулированные вопросы. Кроме того, их просят заполнить формальную карту, в каждом пункте выбрав одну из нескольких градаций (см. главу 7).

Сценарий должен содержать и конкретные методы анализа собранной информации, например вычисление медианы Кемени, статистический анализ люсианов, а также иные методы статистики объектов нечисловой природы и других разделов прикладной статистики (о некоторых из названных методов речь пойдет далее, см. также [89]). 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 печальный практический опыт, информация используется не более чем на 1—2%. Причины в том, что в большом ворохе беспорядочно собранных фактов, как правило, отсутствует необходимая упорядоченность. А именно, значения отдельных показателей собраны с пропусками, способы измерения изменяются от одного эксперта к другому, от одного объекта экспертизы к другому (как говорят, определения «плывут»), сам перечень показателей не позволяет ответить на интересующие ЛПР вопросы и т.д.

Сценарий утверждается научным руководителем экспертной комиссии (ЭК).

5. Подбор экспертов в соответствии с их компетентностью. На этой стадии РГ составляет список возможных экспертов и оценивает степень их пригодности для планируемого исследования. Итоговый перечень должен включать, по крайней мере, в 1,5 раза больше потенциальных экспертов, чем то их число, которое планируется реально привлечь к работе.

6. Формирование ЭК. На этой стадии РГ проводит переговоры с экспертами, получает их согласие на работу в ЭК. Возможно, часть намеченных РГ (на стадии 5) экспертов не сможет войти в экспертную комиссию (болезнь, отпуск, командировка и т.п.) или откажется по тем или иным причинам (занятость, условия контракта и т.п.). В обязательном порядке ЛПР утверждает состав ЭК, возможно, вычеркнув или добавив часть экспертов к предложениям РГ. Проводится заключение договоров с экспертами об условиях их работы и ее оплаты. На этой же стадии завершается формирование РГ.

7. Проведение сбора экспертной информации в соответствии с разработанным на стадии 4 сценарием. 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 компьютеризация экспертных мнений, т.е. создание и наполнение соответствующих баз данных или электронных таблиц.

9. При применении (согласно сценарию) экспертной процедуры, состоящей 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. Форма заключения ЭК обычно задается в ТЗ. В Федеральном законе «Об экологической экспертизе» требованиям к заключению ЭК посвящена обширная глава 18.

11. Официальное ***окончание деятельности ЭК и РГ***, в том числе ***утверждение ЛПР заключительного документа ЭК***, подготовка и утверждение научного и финансового отчетов РГ о проведении экспертного исследования, оплата труда экспертов и сотрудников РГ, официальное прекращение деятельности (роспуск) ЭК и РГ.

Научный отчет РГ должен позволять восстанавливать все подробности деятельности ЭК на основе документов. В частности, в него должны быть включены все полученные от экспертов материалы и протоколы компьютерной обработки данных. Этот отчет может быть использован в суде и арбитражном суде, в случае если заинтересованные организации и лица сочтут нужным оспорить выводы ЭК в судебном порядке.

8.2. ПОДБОР ЭКСПЕРТОВ

Разберем подробнее отдельные стадии экспертного исследования. Начнем с подбора экспертов: кадры решают все! Каковы эксперты — таково и качество заключения ЭК.

Проблема подбора экспертов является одной из наиболее сложных в теории и практике экспертных исследований. Очевидно, в качестве экспертов необходимо использовать тех людей, чьи суждения помогут принятию адекватного решения. Но как выделить, найти, подобрать таких людей? Надо прямо сказать, что *нет методов подбора экспертов, наверняка обеспечивающих успех экспертизы*. Сейчас не будем обсуждать проблему существования различных «партий» среди экспертов и обратим внимание на иные стороны процедур подбора экспертов.

В проблеме подбора экспертов можно выделить две составляющие — *составление списка возможных экспертов* и *выбор из них ЭК в соответствии с компетентностью кандидатов*.

Составление списка возможных экспертов. Данный процесс 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 (или таблицы) 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.

Метод «снежного кома». Для формирования списка есть полезный метод «снежного кома». Это вспомогательное экспертное исследование. Название ассоциируется с известной всем процедурой, когда небольшой снежок много раз поворачивается по поверхности свежевыпавшего снега. При каждом повороте на снежок налипает новый слой, и в результате получается большой снежный ком. В качестве «снежка» используется подобранная РГ небольшая (3—5 человек) группа потенциальных экспертов. В методе «снежного кома» от каждого специалиста, привлекаемого в качестве эксперта, получают определенное количество (обычно 5—10) 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 или когда список достигает необходимого размера. В результате получается достаточно обширный список возможных экспертов.

Рассмотрим условный пример. Для начала РГ подобрала 5 потенциальных экспертов. Каждый из них назвал 10 новых фамилий. Всего РГ получила 50 фамилий. После исключения повторов и лиц, которые не смогут быть экспертами, в списке осталось 40%, т.е. 20 новых фамилий. На следующем туре РГ получает суммарно 200 фамилий. Пусть из них только 30% тех, которые можно добавить к списку. Это 60 человек. При их опросе получаем 600 фамилий. Если из них только 20% реально добавляется к списку, то итог этого тура — 120 фамилий. Подведем итог. В списке уже 5 + 20 + 60 + 120 = 205 фамилий. Можно остановиться, поскольку на основе этого списка, очевидно, можно сформировать ЭК (типовое число членов ЭК — 10—30).

Метод «снежного кома» имеет и недостатки. Число туров до остановки процесса наращивания кома нельзя заранее предсказать. Нельзя априори надеяться, что в обозримой окрестности имеется достаточное число экспертов. 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, лиц из этого же «клана». Мнения и аргументы других «кланов» будут упущены.

Здесь речь идет о том, что сообщество специалистов реально разбито на группы, названные выше «кланами», и общение идет в основном внутри «кланов». Неформальная структура науки, к которой относятся «кланы», достаточно сложна для изучения. Отметим здесь, что «кланы» обычно образуются на основе крупных формальных центров (вузов, научных институтов), научных школ.

Компетентность экспертов. Вопрос об оценке компетентности экспертов не менее сложен. Ясно, что 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 и т.п.), очевидно, в современных быстро меняющихся 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 отклонения и в другую сторону — излишне критичное отношение к своим возможностям. Нам известен доцент МГУ им. М.В. Ломоносова, написавший добротный университетский учебник по математической статистике, заявивший, что он не является специалистом по математической статистике. Видимо, он признает себя специалистом лишь в той узкой научной области, которой посвящены его последние научные статьи. Подобный гиперкритицизм по отношению к себе представляется непродуктивным. Более естественной выглядит рекомендация Е.С. Вентцель: «Если вы хотите изучить какой-либо предмет, напишите по нему книгу». Действительно, при написании книги необходимо разбираться в рассматриваемом вопросе и к концу составления текста становиться высококвалифицированным специалистом — экспертом.

При использовании *метода взаимооценки*, когда оценку компетентности конкретного эксперта дают другие эксперты (или кандидаты в эксперты), помимо возможности проявления личностных и групповых симпатий и антипатий, играет роль малая осведомленность экспертов о профессиональных возможностях друг друга. В современных условиях достаточно хорошее знакомство с работами и возможностями друг друга может быть лишь у специалистов, много лет (не менее трех) работающих совместно, в одном кабинете, над одной темой. Именно про такие пары можно сказать, что они «вместе пуд соли съели» (по примерному расчету, если каждый рабочий день обедать вместе и солить блюда из одной солонки, пуд соли будет съеден за 3,5 года). Однако привлечение таких пар специалистов в ЭК 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 — одна из основных функций РГ, и никакие методики подбора не снимают с нее ответственности. Другими словами, именно на РГ 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.

Нормативное регулирование состава экспертов. Существует ряд нормативных документов, регулирующих деятельность экспертных комиссий в тех или иных областях. Примером является Закон Российской Федерации от 23 ноября 1995 г. «Об экологической экспертизе», в котором регламентируется процедура экспертизы «намечаемой хозяйственной или иной деятельности» с целью выявления возможного вреда, который может нанести рассматриваемая деятельность окружающей природной среде. В этом законе указаны дополнительные требования к экспертам, призванные обеспечить их независимость от внешних влияний. Так, в ст. 16, ч. 2 данного закона сказано: «Экспертом государственной экологической экспертизы не может быть представитель заказчика документации, подлежащей государственной экологической экспертизе, или разработчика объекта государственной экологической экспертизы, гражданин, состоящий в трудовых или иных договорных отношениях с указанным заказчиком или с разработчиком объекта государственной экологической экспертизы, а также представитель юридического лица, состоящего с указанным заказчиком или с разработчиком объекта государственной экологической экспертизы в таких договорных отношениях».

Используется и принципиально иной подход к подбору экспертов, согласно которому совокупность экспертов состоит из тех, кто сам себя объявил таковыми. Примерами являются разнообразные опросы, приводимые в Интернете и регулярно публикуемые на сайтах РБК — РИА «РосБизнесКонсалтинг» (http://rbc.ru) и «Глас РУНЕТа — служба опросов интернет-аудитории» (http://voxru.net). В отличие от метода самооценки, здесь требуется и волевой импульс от эксперта — решение об участии в опросе. В случаях когда какие-либо материалы предлагаются к обсуждению, от самовыдвинувшихся экспертов получают ответы на открытые вопросы (а не на закрытые, как в случае опросов РБК). Письма и обращения, поступающие самотеком в средства массовой информации и в государственные органы, также можно рассматривать в рамках теории экспертных оценок. Однако надо подчеркнуть, что распределение самовыдвинувшихся экспертов по социально-экономическим группам (например, по полу и возрасту) обычно существенно отличается от того, которое имеется в обществе. Частично от этого смещения можно избавиться с помощью методов стандартизации («ремонта») выборки, разработанных в эконометрике и прикладной статистике [77; 89].

8.3. О ВЫБОРЕ ЦЕЛИ ЭКСПЕРТИЗЫ

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 на классификацию методов экспертных оценок может принести лишь вред.*

Однако для рассказа о многообразии экспертных оценок необходима какая-либо рабочая классификация методов. Приведем одну из таких возможных классификаций, перечисляя основания, по которым делим методы экспертных оценок.

Один из основных вопросов: что именно должна представить ЭК в результате своей работы — информацию для принятия решения ЛПР или проект самого решения? От ответа на этот методологический вопрос зависит организация работы ЭК, и он служит первым основанием для разбиения методов.

Цель — сбор информации для ЛПР. В этом случае РГ 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, а также и к первому, который имеет возможность дополнить свою аргументацию. 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 (среднестатистического), т.е. инакомыслящие (диссиденты). Именно от них следует ожидать наиболее оригинальных аргументов.

Цель — подготовка проекта решения для ЛПР. Основная задача при этом — разработка (формулировка, получение) коллективного мнения ЭК. Математические методы анализа экспертных оценок 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 обоснований 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 (членов Ученого совета НИИ) 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. Поэтому при голосовании с целью выявления лучшей научно-исследовательской работы за год результат зависел не от рассматриваемых работ, а от численности представителей групп «теоретиков» и «практиков», присутствующих на заседании.

Иногда заявляют, что в случае обнаружения двух или нескольких групп экспертов (вместо одной согласованной во мнениях) опрос не достиг цели. Это не так! *Цель достигнута: установлено, что единого*

мнения нет. Это весьма важно. И при принятии решений ЛПР 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 РГ (или даже «подсказанной» ЛПР).

Правильное решение было принято руководством одного НИИ после обнаружения отсутствия единомыслия среди членов Ученого совета: вместо одной премии стали присуждать две — отдельно за теоретические работы и отдельно за прикладные.

Часто не учитывают еще одного чисто математико-статистического обстоятельства. Поскольку число экспертов обычно не превышает 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 укажем на конкретные методы расчетов с помощью коэффициентов конкордации (т.е. в переводе — согласия) на основе коэффициентов ранговой корреляции Кендалла или Спирмена. Необходимо напомнить, что согласно математико-статистической 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 по дурной традиции 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 математико-статистического аппарата для проверки согласованности — непараметрических методов, основанных на так называемых люсианах [77; 89] и входящих в современный раздел эконометрики — *статистику нечисловых данных*.) Невозможность получения обоснованного заключения о согласованности мнений экспертов по ограниченным данным можно сопоставить с невозможностью проверки нормальности теоретического распределения в случае, когда объем выборки менее 50 (это утверждение подробно обосновано в статье [113]).

Отметим, что группы экспертов с близкими мнениями можно выделить методами кластер-анализа [77].

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 (см., например, учебник [77]).

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) оценки и аргументы диссидентов. Другим примером является принятие решений при судействе в фигурном катании, когда с целью повышения устойчивости выводов жюри отбрасываются минимальная и максимальная из оценок судей.

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 ЭК и РГ по проведению экспертного опроса и перекладывает эту ответственность и труд на плечи ЛПР.

Догма одномерности. В устаревшей, а иногда и в современной научно-технической, управленческой и экономической литературе распространен довольно спорный подход так называемой «квалиметрии», 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». Практически все реальные объекты достаточно сложны, а потому сколько-нибудь точно описать их можно лишь с помощью многих и многих чисел, а также математических объектов нечисловой природы. Жизнь, в том числе экономическая, многомерна, а не одномерна!

Вместе с тем нельзя полностью отрицать саму идею поиска обобщенных показателей качества, технического уровня, конкурентоспособности и т.п. Так, каждый объект можно оценивать по многим показателям качества. Например, легковой автомобиль можно оценивать по таким показателям и группам показателей:

- расход бензина на 100 км пути (в среднем);
- надежность (в том числе число отказов и средняя стоимость ремонта за год);
- безопасность эксплуатации;
- экологическая безопасность, оцениваемая по содержанию вредных веществ в отработавших газах;
- легкость в управлении;
- маневренность (в том числе радиус поворота);
- быстрота набора заданной скорости (например, 100 км/ч) после начала движения;
- максимальная достигаемая скорость;
- длительность сохранения в салоне положительной температуры при низкой наружной температуре и выключенном двигателе;
- эстетичность (дизайн, привлекательность и «модность» внешнего вида автомобиля и отделки салона);
- масса и т.д.

Можно ли свести оценки по этим показателям вместе? Ясно, что 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 служб спасения и государственного управления, видимо, надежность важнее, чем для частника, а расход бензина — наоборот. Для районов Крайнего Севера важна теплоизоляция салона, а для центральных районов — нет. И так далее.

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). 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 *качественно* 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.

8.4. ОСНОВАНИЯ ДЛЯ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ МЕТОДОВ

Первому основанию — цели экспертизы — посвящен предыдущий подраздел. Экспертные методы подразделяются на два класса в соответствии с ответом на вопрос: «Что именно должна представить экспертная комиссия в результате своей работы — информацию для принятия решения ЛПР или проект самого решения?»

Рассмотрим еще несколько оснований.

Число туров. Второе основание классификации экспертных процедур — число туров. Экспертизы могут включать один тур, некоторое фиксированное число туров или неопределенное число туров.

Экспертиза в один тур предполагает, что эксперты не обмениваются информацией, поскольку не общаются друг с другом. Технология такой экспертизы напоминает технологии маркетинговых и социологических выборочных обследований. Это наиболее быстрая и дешевая технология, но и в наименьшей степени использующая творческие способности экспертов, а потому дающая наименьшие полезные результаты.

Наличие нескольких туров предполагает, что эксперты получают информацию друг от друга, обрабатывают ее, получают новое знание и в соответствии с ним корректируют свои выводы. 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.

Наибольшие сложности вызывают процедуры с неопределенным заранее 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.

Порядок вовлечения экспертов. 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 эксперт 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.

Итак, экспертные процедуры можно классифицировать на основании того, как эксперты вовлекаются в работу — одновременно или последовательно. Первый вариант более быстрый, но и более затратный (дорогой), второй — дешевле, но более продолжительный.

Организация общения экспертов. Четвертое основание классификации экспертных процедур — способ организации 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. В будущем с распространением телеконференций грань между очным и заочным общением экспертов начнет стираться.

Общение соответствует также многим реальным процедурам принятия управленческих решений. Координация действий организаций и менеджеров с помощью заочного общения без анонимности происходит и при подготовке документов — планов, приказов, предложений, направляемых в другие организации, ответов на распоряжения и запросы властей и др. Управленческие решения обычно оформляются в виде подобных документов.

Обычно один из сотрудников — назовем его Исполнителем — готовит первоначальный вариант документа, который размножают и рассылают на отзыв заинтересованным в нем менеджерам, а иногда и в другие организации. Исполнитель составляет сводку отзывов, с одними из замечаний соглашается, против других высказывает возражения. Затем собирают так называемое «согласительное совещание», на которое приглашают всех тех, с чьим мнением Исполнитель не согласен. В результате дискуссии по ряду позиций достигается компромисс и возражения снимаются. Окончательное решение по проекту документа с учетом оставшихся возражений принимает ЛПР, например генеральный директор или совет директоров, т.е. высшая инстанция в данной организации. Именно такова процедура подготовки законов Российской Федерации, государственных стандартов и иных ответственных документов.

Во многих случаях эта процедура упрощается и отзывы заменяются *визированием,* при котором свое согласие менеджеры выражают, накладывая на документ *визу,* т.е. расписываясь (иногда добавляя несколько слов по затрагиваемой проблеме). Например, подготовленное для отправки в другую организацию письмо или приказ по организации визируют руководители нескольких отделов, и генеральный директор его подписывает от имени фирмы, не вникая в суть (поскольку каждый день он подписывает десятки писем и приказов, то вникать некогда). Адресату уходит письмо, на обратной стороне которого указаны фамилия и телефон Исполнителя (поскольку адресат тоже хорошо знаком с процедурой подготовки документов, он понимает, что по конкретным вопросам надо обращаться к Исполнителю, а не к генеральному директору). В архиве фирмы остается письмо с визами, так что в случае необходимости легко выяснить, кто составил и одобрил документ.

Как ясно из сказанного, заочные экспертизы часто используются совместно с очными.

При очных экспертизах эксперты говорят, а не пишут, как при заочных, и потому успевают за то же потраченное время сообщить существенно больше. ***Очное общение с ограничениями*** весьма распространено. Это собрание, идущее по фиксированному регламенту. Примером является военный совет в императорской русской армии,

когда эксперты (офицеры и генералы) высказывались в фиксированном порядке от младшего (по чину и должности) к старшему. Другой пример — разработка и принятие решений в Государственной думе в соответствии с регламентом, определяющим последовательность и продолжительность выступлений на заседаниях комиссий, комитетов, других структур, на пленарных заседаниях. Вспомним также технологию «мозгового штурма».

Наконец, ***очное общение 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 пять лейтенантов и три генерала. Независимо от того, какая информация имеется у того или иного участника встречи, ход ее предсказать нетрудно: генералы будут беседовать, а лейтенанты — помалкивать. При этом вполне очевидно, что лейтенанты получили образование позже генералов, а потому могут обладать полезной информацией, которой нет у генералов.

Веса экспертов. Следующее основание классификации экспертных процедур — по способам введения весов для мнений экспертов. Простейший способ — все эксперты равноправны, при голосованиях по отдельным положениям разрабатываемого решения имеют по одному голосу.

Часто вводят понятия решающего голоса и совещательного голоса. Например, при защите дипломного проекта члены Государственной аттестационной комиссии (ГАК) имеют решающие голоса, а все остальные участники заседания — совещательные. В Федеральном законе «Об экологической экспертизе» (1995 г.) подробно расписано, представители каких организаций и структур управления могут присутствовать на заседании экспертной комиссии государственной экологической экспертизы с правом совещательного голоса.

В регламент принятия решений иногда включают положение, согласно которому при делении голосов ровно пополам принимается мнение той половины, к которой относится председатель ЭК. Это означает, что вес голоса председателя на бесконечно малую величину больше веса рядового эксперта. Впрочем, иногда председателю дают два голоса.

При голосованиях на собраниях акционеров вес каждого эксперта (участника заседания) определяется числом акций, которыми он распоряжается.

Комбинация различных видов экспертизы. Реальные экспертизы часто представляют собой комбинации различных описанных ранее 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 защиту дипломного проекта. Наконец — очная экспертиза без ограничений (для членов ГАКа). Дипломный проект оценивается коллегиально, по большинству голосов, при этом один из экспертов (научный руководитель) знает работу подробно, а остальные — в основном лишь по докладу. Отметим, что мнения экспертов учитываются с весами, а именно: мнения членов ГАКа — с весом 1, мнения всех остальных — с весом 0 (совещательный голос). Таким образом, имеем сочетание многотуровой и однотуровой, заочной и очной экспертиз. Подобные сочетания характерны для многих реально проводящихся экспертиз.

8.5. ИНТУИЦИЯ ЭКСПЕРТА И КОМПЬЮТЕР

Обсудим две, казалось бы, далекие друг от друга, но на самом деле тесно связанные между собой темы — роль интуиции эксперта в экспертизе и применение вычислительной техники в технологиях экспертных исследований.

Интуиция эксперта. Примером хорошего эксперта служит врач, чьи диагнозы чаще, чем у его коллег, оправдываются при вскрытии. Дело в том, что только вскрытие позволяет патологоанатому дать достоверное заключение о том, чем болел пациент и правильно ли его лечили. Хотя это достоверное заключение уже не может принести пользы пациенту, его можно применить для оценки профессиональных возможностей врача и корректировки лечебных технологий. Причем чем лучше врач, тем дольше придется ждать подтверждения его высокого профессионализма.

В целях создания систем компьютерной диагностики математики пытались выяснить, как работают выдающиеся врачи [13]. Для этого их просили описать используемые ими в лечебной работе методы умозаключений. Практикующие врачи приводили примерно те же формулировки, что и авторы медицинских учебников. И это вполне естественно. Однако при попытках применить сформулированные таким путем правила для диагностики вновь поступающих пациентов качество принимаемых врачебных решений резко ухудшалось — вплоть до уровня рядового выпускника мединститута. Таким образом, оказалось, что выдающиеся врачи не в состоянии описать, как именно они работают. При попытке вербализации процесса диагностики интуиция исчезала, а вместе с ней — и отличие высококвалифицированного эксперта от рядового специалиста.

Важную роль интуиции в работе эксперта трудно, а точнее, практически невозможно промоделировать математически. Как следствие, нельзя и мечтать о замене экспертных оценок компьютерными расчетами. Экспертиза — это творчество.

Роль интуиции весьма велика в различных творческих профессиях. Например, математическое творчество, по свидетельству выдающегося французского математика Ж. Адамара, основано на интуиции [2].

Экспертные оценки и экспертные системы. Хотя названия этих двух научно-практических дисциплин похожи, различие между ними колоссально. *Теория экспертных оценок* — это наука о методах сбора и анализа мнений людей (экспертов), опирающихся на свою интуицию. *Экспертная система* — это программа для компьютера, которая оперирует со знаниями в определенной предметной области с целью выработки практических рекомендаций для решения возникших проблем. Значит, в экспертных системах не участвуют живые люди, есть только ранее полученные знания — результат прошлой деятельности специалистов. При формализации знаний невозможно учесть интуицию экспертов. Однако компьютерной обработке может быть подвергнут огромный объем знаний, что человек сделать не в состоянии.

Сравнительные возможности живых экспертов и экспертных систем видны при сопоставлении шахматистов и шахматных программ. Люди опираются на интуицию, а компьютеры — на расчеты. Результат известен — за пятьдесят лет компьютеры достигли уровня гроссмейстеров.

Однако речь идет об анализе довольно простой игры — шахматные правила изложены на нескольких страницах, и они строго выполняются. Реальные ситуации гораздо сложнее, и самое интересное — правила игры могут меняться.

В настоящее время экспертные системы, как и другие достижения искусственного интеллекта, — помощники человека. Например, на рыболовном судне или в отдаленном поселении целесообразно иметь экспертную систему неотложной медицинской помощи. Она позволит сохранить жизнь пострадавшему, пока не появится врач. Врачу она тоже поможет для различных справок. Но лечить будет именно врач.

В обозримом будущем та или иная рутинная работа будет передаваться компьютерным системам, например составление бухгалтерского баланса. Но за человеком всегда останется целеполагание. Компьютер, в отличие от человека, не может знать, чего он хочет.

Эксперт и компьютер. Обсудим разные варианты взаимодействия живых экспертов и компьютерных систем.

1. Эксперту нужна различная справочная информация, и наиболее быстро он может ее получить с помощью компьютера. Так, всемирная сеть Интернет — хороший помощник эксперта. К сожалению, в сети циркулирует масса ошибочных сведений. Но ведь и информация, полученная из книг или от людей, не всегда достоверна.

2. Быстрая электронная связь с организаторами экспертизы, с другими экспертами, возможность удаленного общения (чаты, телеконференции и другие формы) резко повышают эффективность экспертной работы.

3. Автоматизированное рабочее место (АРМ) эксперта, например АРМ МАТЭК (математика в экспертизе), обеспечивает как сбор экспертной информации, так и ее анализ с помощью разнообразных математических методов.

4. Экспертные процедуры могут многократно использоваться на различных этапах процесса принятия решений, например для оценки значений признаков, описывающих объекты, или для оценки коэффициентов важности (весомости) самих признаков. При этом процесс принятия решений опирается на ту или иную форму компьютерной поддержки.

5. Интегрированные системы принятия решений включают в себя разнообразные базы данных и знаний, АРМ ЛПР, экспертов и сотрудников группы сопровождения, блоки имитационных, экономико-математических и иных компьютерных моделей (в том числе блоки соответствующих экспертных систем). Такие системы действуют в составе аналитических центров крупных организационных структур, например в Администрации Президента Российской Федерации, Центре управления полетами космических аппаратов, в штабах высокого уровня Вооруженных сил России или в руководящих структурах транснациональных корпораций.

В качестве примера рассмотрим подробнее АРМ МАТЭК (математика в экспертизе).

Автоматизированное рабочее место МАТЭК[1]. Разработано и применяется весьма большое число методов (и особенно их разновидностей) организации и проведения экспертных исследований. Для решения конкретной задачи можно использовать, как правило, не

[1] МАТЭК — МАТематические методы в ЭКспертных оценках

один, а много методов, и выбор наиболее подходящего из них лежит на организаторах экспертизы (попытки стандартизовать правила принятия подобных решений в настоящее время рассматриваются как нецелесообразные — таков один из результатов развития стандартизации в нашей стране и в мире в последние десятилетия, начиная с 1970-х гг.). Автоматизированное рабочее место МАТЭК предоставляет организаторам экспертизы большие возможности для выбора тех или иных методов планирования, организации, проведения экспертизы, анализа экспертных оценок, обеспечивает необходимую компьютерную поддержку в проведении экспертного исследования.

Автоматизированное рабочее место МАТЭК предназначено для подготовки и проведения экспертизы по определенной теме. С помощью АРМ МАТЭК можно автоматизировать процесс подбора экспертов, работу комиссии экспертов и анализ экспертных мнений, а также подготовку опросных листов, бланков и всей отчетной документации.

Работа на АРМ в соответствии с методологией работы [91] состоит из двух этапов:

А. Подготовка экспертизы;

В. Проведение экспертизы.

Этап А подготовки экспертизы включает в себя ввод всей информации, необходимой для проведения экспертизы. Итогом этого этапа являются два документа: техническое задание (ТЗ) и сценарий.

Рассмотрим этап *А* подробнее. Сначала ЛПР должно сформулировать цель экспертизы, сформировать руководство РГ.

Далее к работе приступает РГ. Руководитель РГ должен ввести данные для формирования ТЗ. Затем собираются данные для компоновки сценария.

Рабочая группа может включать в себя руководителя, группу обработки, группу связи и интервьюеров.

Данные для ТЗ следующие:

- основание для проведения экспертизы;
- задачи экспертных опросов, сформулированные в соответствии с целью экспертизы;
- требования к ЭК, опросному листу;
- сроки выполнения экспертизы и порядок контроля за ними;
- финансовое обеспечение проекта.

В зависимости от того, введены или нет те или иные данные для ТЗ, они соответственно будут или не будут включены в него. Техническое задание можно просмотреть на экране и распечатать.

Данные для документа сценария следующие:

- вводный текст (в этом тексте должна содержаться собственно последовательность действий при проведении экспертизы);

- календарный план (КП);
- список используемых методов анализа экспертных мнений (ЭМ).

Как и при формировании ТЗ, сценарий может иметь разную структуру, в зависимости от того, какие пункты будут в него включены. Как приложение к сценарию могут быть использованы примеры бланков опросных листов, анкеты «Согласие» (для выявления согласия экспертов участвовать в экспертизе), анкеты «Снежный ком», «Взаимооценка» (если соответствующие этапы включены в КП). Для этих бланков также требуется ввести оповещение (либо выбрать стандартное). Сценарий можно просмотреть на экране и распечатать.

При формировании сценария будет сформирован опросный лист экспертизы. Он состоит из оповещения (стандартного или оригинального — по выбору РГ) и собственно вопросов. Вопросы группируются по задачам из ТЗ. При формулировке вопросов учитывается список методов обработки ответов. Точнее, пользователь, сформулировав вопрос, должен точно знать формат ответа. Для каждого формата ответа в АРМ предусмотрен список методов обработки ответов (краткое описание каждого из них можно будет просмотреть при выборе метода). Если пользователя не устраивает ни один из этих методов, он должен будет переформулировать вопрос (т.е. изменить формат ответа) так, чтобы в списке соответствующих методов оказался подходящий ему. Тем самым при формировании опросного листа будет одновременно сформулирован список используемых методов анализа экспертных мнений (ЭМ).

Этап В проведения экспертизы недоступен до тех пор, пока не будет завершен этап подготовки экспертизы. После того как подготовка создана, можно запустить или открыть проведение экспертизы. Тем самым возможно проведение нескольких экспертиз с одной и той же подготовкой (для каждой экспертизы выделяется собственная, идентифицируемая по названию экспертизы база данных).

На этапе проведения экспертизы формируется ЭК, проводится сбор и анализ ЭМ, формируется отчет и заключение для ЛПР.

Формирование ЭК — многоступенчатый процесс. Сначала член РГ (руководитель) в соответствии с информацией об экспертах из базы данных может отобрать подходящих кандидатов в ЭК. Далее с помощью анкеты «Согласие» из этого списка отбираются согласившиеся быть членами ЭК. Два последних шага могут проводиться или нет, в зависимости от того, включены ли они в КП. Это этапы «Снежный ком» и «Взаимооценка».

После того как сформирована ЭК, можно проводить сбор ЭМ. Это осуществляется с помощью бланка вопросника. Храниться ЭМ будут

так, чтобы доступ к ним был удобен (т.е. по любому эксперту и любому вопросу можно было получить ответ и т.д.). Анализ ЭМ по каждому вопросу проводится методом, выбранным пользователем АРМ (руководителем РГ) на этапе подготовки экспертизы для этого вопроса.

По всем предыдущим этапам формируются отчеты, из которых в результате получается общий отчет о проведении экспертизы. В соответствии с задачами из ТЗ формируется заключение для ЛПР.

В соответствии с КП ведется контроль за сроками проведения экспертизы.

Ведется протокол экспертизы, т.е. при выходе из системы фиксируется текущее состояние этапа проведения экспертизы, и при открытии данной экспертизы происходит возврат именно на тот этап экспертизы, на котором произошел выход из системы (на этапе подготовки экспертизы протокол не ведется). Разграничены права доступа к базе данных экспертов, ЭМ и результатам обработки ЭМ.

На этом заканчивается «гуманитарная» часть обсуждения теории и практики экспертных оценок. Конкретные методы сбора и анализа экспертной информации рассмотрены в дальнейших главах учебника с привлечением современного математического аппарата.

Контрольные вопросы

1. Каковы основные стадии экспертного опроса?
2. Почему сценарий проведения сбора и анализа экспертных мнений необходимо разрабатывать до подбора экспертов?
3. Что называется методом «снежного кома»?
4. Как выбор цели экспертизы влияет на экспертные технологии?
5. Какова роль диссидентов в комиссии экспертов в зависимости о регламента сбора и анализа экспертных мнений?
6. По каким основаниям классифицируют экспертные методы?
7. Чем отличаются экспертные оценки и экспертные системы?
8. Какова роль компьютеров в экспертных технологиях?

Темы докладов и рефератов

1. Роль ЛПР в организации экспертного исследования.
2. Внутренняя структура РГ экспертного исследования.
3. Типовые сценарии проведения сбора и анализа экспертных мнений.
4. Требования к экспертам, зафиксированные в действующем законодательстве.
5. Сравнительный анализ методов самооценки и взаимооценки.
6. Догма согласованности.
7. Догма одномерности.
8. Подходы к выбору способа организации общения экспертов.
9. Роль интуиции в экспертизе.

ГЛАВА 9
ТЕОРИЯ ИЗМЕРЕНИЙ И ЭКСПЕРТНЫЕ ОЦЕНКИ

9.1. ОСНОВНЫЕ ШКАЛЫ ИЗМЕРЕНИЯ

Почему необходима теория измерений? Теория измерений (ТИ) необходима для разработки технологий экспертного оценивания. За последние десятилетия она прошла путь от малоизвестного раздела математической психологии до общенаучной концепции, знакомство с которой признается обязательным для исследователей и студентов самых разных специальностей (в качестве примеров укажем книги [89; 122]). Теория измерений является одной из составных частей наук, посвященных анализу данных статистики и эконометрики. Принято считать, что ТИ входит в состав *статистики объектов нечисловой природы* [1]. Эта теория исходит из того, что арифметические действия с используемыми в практической работе числами не всегда имеют смысл. Действительно, использование чисел в жизни и хозяйственной деятельности людей отнюдь не всегда предполагает, что эти числа можно складывать и умножать, производить иные арифметические действия. Что бы вы сказали о человеке, который занимается сложением или умножением телефонных номеров? Не всегда выполнимы привычные арифметические соотношения. Например, сумма знаний двух двоечников не равна знаниям «хорошиста», т.е. для оценок знаний 2 + 2 не равно 4. Если вы вечером поместите в клетку двух животных, а потом еще двух, то отнюдь не всегда можно утром найти в этой клетке четырех животных. Их может быть и много больше — если вечером вы загнали в клетку овцематок или беременных кошек. Их может быть и меньше — если к двум волкам вы поместили двух ягнят. Итак, отнюдь не всегда 2 + 2 = 4. Числа используются гораздо шире, чем арифметика.

Приведенные примеры показывают, что практика использования чисел для описания результатов наблюдений (измерений, испытаний, анализов, опытов) заслуживает методологического анализа.

При применении тех или иных экспертных технологий прежде всего необходимо разобраться с проблемами измерения различных величин, используемых в процессе сбора и анализа экспертных мнений. Они могут быть измерены в тех или иных количественных или качественных шкалах. Поскольку в выборе конкретной шкалы имеется

некоторый произвол (например, расстояние можно измерять в аршинах, саженях, верстах, метрах или парсеках), то естественно потребовать, чтобы принимаемое решение не зависело от этого произвола (например, от того, в каких единицах измерено расстояние).

Так, например, мнения экспертов часто выражены в *порядковой шкале* (подробнее о шкалах говорится далее), 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* он 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 значений характеристики у различных объектов. Такой ряд в статистике называется *вариационным*. Формально ранги выражаются числами 1, 2, 3, ..., но с этими числами нельзя делать привычные арифметические операции. Например, хотя в арифметике 1 + 2 = 3, но нельзя утверждать, что для объекта, стоящго на 3-м 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 ТИ.

При чтении литературы надо иметь в виду, что в настоящее время термин «теория измерений» применяется для обозначения целого ряда научных дисциплин, а именно классической метрологии (науки об измерениях физических величин), рассматриваемой здесь ТИ, некоторых других направлений, например алгоритмической теории измерений. Обычно из контекста понятно, о какой конкретно теории идет речь.

Краткая история теории измерений. Сначала ТИ развивалась как теория психофизических измерений. В послевоенных публикациях американский психолог С.С. Стивенс основное внимание уделял шкалам измерения. Во второй половине XX в. сфера применения ТИ стремительно расширялась. Посмотрим, как это происходило.

Один из томов выпущенной в США в 1950-х гг. «Энциклопедии психологических наук» назывался «Психологические измерения».

Значит, составители этого тома расширили сферу применения ТИ 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 [119] упор был сделан на «гомоморфизмы эмпирических систем с отношениями в числовые» (в эти математические термины здесь вдаваться нет необходимости), и математическая сложность изложения заметно 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 ТИ. 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.

Итоги этого этапа были подведены в монографии [88]. В качестве двух основных проблем ТИ наряду с *установлением типа шкалы* измерения конкретных данных был выдвинут поиск алгоритмов анализа данных, результат работы которых не меняется при любом допустимом преобразовании шкалы (т.е. является инвариантным относительно этого преобразования).

Метрологи вначале резко возражали против использования термина «измерение» для качественных признаков. Однако постепенно возражения сошли на нет, и к концу XX в. ТИ стала рассматриваться как общенаучная теория.

Необходимость использования ТИ в теории и практике экспертного оценивания рассмотрим, в частности, в связи с агрегированием мнений экспертов, построением обобщенных показателей и рейтингов.

Рассмотрим основные идеи теории измерений.

Шесть основных типов шкал. В соответствии с ТИ при математическом моделировании реального явления или процесса следует, прежде всего установить *типы шкал*, в которых измерены те или иные переменные. Тип шкалы задает *группу допустимых преобразований шкалы*. Допустимые преобразования не меняют объективно существующих соотношений между объектами измерения.

Например, при измерении длины переход от аршина к метру 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. Если первый длиннее второго в 5 раз при измерении в дюймах, то и при измерении в саженях первый будет длиннее второго ровно в 5 раз. Обратите внимание, что при этом численное значение длины в аршинах отличается от численных значений длины в метрах, дюймах и саженях — не меняется лишь результат сравнения длин двух объектов и отношение длин.

Укажем основные виды шкал измерения и соответствующие группы допустимых преобразований.

В ***шкале наименований*** (другое название этой шкалы — ***номинальная***, это термин на основе латыни; иногда называют также классификационной шкалой) допустимыми являются все взаимнооднозначные 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. Номера страховых свидетельств государственного пенсионного страхования, медицинского страхования, штрих-коды товаров, индивидуальные номера налогоплательщиков (ИНН) измерены в шкале наименований. В этой шкале измерены и многие иные величины, с формальной точки зрения выраженные числами. 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 ИНН или номера паспортов — 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.

Итак, наиболее простой способ использования чисел — применение их для различения объектов. Например, телефонные номера нужны для того, чтобы отличать одного абонента от другого. При таком способе измерения используется только одно отношение между числами — равенство (два объекта описываются либо равными числами, либо различными). С прикладной точки зрения шкала измерения — это способ приписывания чисел рассматриваемым объектам, соответствующий имеющимся между объектами отношениям. Шкалы наименований соответствуют эмпирическим системам, в которых есть только одно отношение — равенства (эквивалентности) элементов этих систем.

Отметим, что числа могут быть приписаны объектам разными способами. Переход от одного способа к другому наблюдаем при замене паспортов или телефонных номеров. Каковы свойства допустимых преобразований? Для шкалы наименований естественно потребовать только взаимной однозначности. Другими словами, применив к результатам измерений взаимно-однозначное преобразование, получаем новую шкалу, столь же хорошо описывающую систему исходных объектов, как и прежняя шкала.

Допустимые преобразования проводятся время от времени в реальной жизни, например при замене телефонных номеров или паспортов. При этом каждому прежнему телефонному номеру соответствует один и только один новый. Не допускается, чтобы два прежних номера «слились» в одном новом или чтобы из одного прежнего получилось два новых. Это и означает, что преобразование является взаимнооднозначным.

В ***порядковой шкале*** числа используются не только для различения объектов, но и 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 школе ровно тот же смысл выражается словесно — известными всем терминами: «неудовлетворительно», «удовлетворительно», «хорошо», «отлично». Этим подчеркивается «нечисловой» характер оценок знаний учащихся. Ведь фактически преподаватель относит учащихся к одной из четырех упорядоченных категорий, а обозначения этих категорий используются лишь для удобства управления образовательным процессом.

Порядковые шкалы соответствуют эмпирическим системам, в которых, кроме отношения равенства (эквивалентности) элементов, есть отношение порядка (нестрогого) между элементами этих систем. Известно, что в таком случае элементы эмпирической системы можно разбить на классы эквивалентности, между которыми имеется отношение строгого линейного порядка [88].

В порядковой шкале допустимыми являются все строго возрастающие преобразования. Так, автор настоящего учебника при участии в российско-французском образовательном проекте пользовался наряду с российской и традиционной французской шкалой оценок, в которой знания учащихся оцениваются числами от 1 до 20. Приходилось постоянно осуществлять преобразования, в которых оценка «неудовлетворительно» переходила в 8, «удовлетворительно» — в 12, «хорошо» — в 15, «отлично» — в 18 (французская шкала оценок позволяла дать численное выражение и дополнительным вариантам оценок, например «четыре с плюсом» — это 16, а «три с двумя минусами» — это 10).

Установление типа шкалы, т.е. задания группы допустимых преобразований шкалы измерения, — дело специалистов соответствующей прикладной области. Так, оценки привлекательности профессий мы в монографии [88], 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. Например, можно предложить каждому опрашиваемому ставить оценки одновременно по двум шкалам, а затем изучить соотношения между выставленными оценками и выявить вид реально используемых преобразований. Пока же подобный эксперимент не поставлен, целесообразно принимать порядковую шкалу, так как это гарантирует от возможных ошибок при анализе данных и получении выводов.

Оценки экспертов, как уже отмечалось, часто следует считать измеренными в порядковой шкале. Типичным примером являются задачи ранжирования и классификации промышленных объектов, подлежащих экологическому страхованию.

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 примерную массу в граммах.

В различных областях человеческой деятельности применяется много других видов порядковых шкал. Так, например, в минералогии используется шкала Мооса, по которой 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. Минерал с большим номером является более твердым, чем минерал с меньшим номером, при нажатии царапает его.

Порядковыми шкалами в географии являются бофортова шкала ветров («штиль», «слабый ветер», «умеренный ветер» и т.д.) и шкала силы землетрясений. Очевидно, нельзя утверждать, что землетрясение в 2 балла (лампа качнулась под потолком — такое бывает и в Москве) ровно в 5 раз слабее, чем землетрясение в 10 баллов (полное разрушение всего на поверхности земли).

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) и т.д. Все эти шкалы построены по схеме: заболевание не обнаружено; первая стадия заболевания; вторая стадия; третья стадия. Иногда выделяют стадии 1а, 1б и др. Каждая стадия имеет свойственную только ей медицинскую характеристику. При описании групп инвалидности числа иногда используются в противоположном порядке: самая тяжелая — первая группа инвалидности, затем — вторая, самая легкая — третья.

Номера домов также измерены в порядковой шкале — они показывают, в каком порядке стоят дома вдоль улицы. Номера томов в собрании сочинений писателя или номера дел в архиве предприятия обычно связаны с хронологическим порядком их создания.

При оценке качества продукции и услуг, в так называемой 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. Иногда применяют четыре градации: имеются критические дефекты (делающие невозможным использование контролируемой единицы продукции) — есть значительные дефекты — присутствуют только незначительные дефекты — нет дефектов. Аналогичный смысл имеет сортность продукции — высший сорт, первый сорт, второй сорт, ...

При оценке экологических воздействий первая, наиболее обобщенная оценка — обычно порядковая, например: природная среда стабильна — природная среда угнетена (деградирует). Аналогично в эколого-медицинской шкале: нет выраженного воздействия на здоровье людей — отмечается отрицательное воздействие на здоровье.

Все шкалы измерения делят на две группы — шкалы качественных признаков и шкалы количественных признаков.

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. В них к отношениям равенства и порядка добавляются отношения, связанные с наличием единицы измерения и начала отсчета.

По ***шкале интервалов*** измеряют величину потенциальной энергии, координату точки на прямой (а также координаты точки на плоскости или в пространстве), географическую долготу (отсчитываемую в настоящее время от произвольно выбранного меридиана Гринвичской обсерватории в Великобритании), температуру по Цельсию, Фаренгейту или Реомюру. Во всех этих случаях на шкале нельзя отметить ни естественное начало отсчета, ни естественную единицу измерения. Исследователь должен сам задать точку отсчета и сам выбрать единицу измерения. Часто путем соглашения договариваются о выборе определенной единицы измерения, фиксируют начало отсчета, но произвольность подобного договора очевидна (например, в случае географической долготы).

Допустимыми преобразованиями в шкале интервалов являются линейные возрастающие преобразования, т.е. линейные функции. Температурные шкалы Цельсия и Фаренгейта связаны именно такой зависимостью:

$$^\circ\mathrm{C} = \frac{5}{9}\left(^\circ\mathrm{F} - 32\right), \tag{9.1}$$

где $^\circ$C — температура (в градусах) по шкале Цельсия, а $^\circ$F — температура по шкале Фаренгейта.

При допустимых преобразованиях в шкале интервалов сохраняется отношение длин интервалов:

$$\frac{x_1 - x_2}{x_3 - x_4} = \frac{\varphi(x_1) - \varphi(x_2)}{\varphi(x_3) - \varphi(x_4)} \tag{9.2}$$

для любых чисел x_1, x_2, x_3, x_4 (результатов измерений) и любого допустимого преобразования $\varphi(x) = ax + b$, $a > 0$. Верно и обратное: если равенство (9.2) справедливо для любых чисел x_1, x_2, x_3, x_4, то $\varphi(x) = ax + b$, $a > 0$ при некоторых значениях коэффициентов a и b.

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, работа, мощность, заряд, напряжение, а также цены в экономике. Допустимыми преобразованиями в шкале отношений являются подобные (изменяющие только масштаб), другими словами, линейные возрастающие преобразования без свободного члена. Примером является пересчет цен из одной валюты в другую по фиксированному курсу.

При допустимых преобразованиях в шкале отношений сохраняется отношение измеряемых величин:

$$\frac{x_1}{x_2} = \frac{\varphi(x_1)}{\varphi(x_2)} \tag{9.3}$$

для любых чисел x_1 и x_2 (результатов измерений) и любого допустимого преобразования $\varphi(x) = ax$, $a > 0$. Верно и обратное: если равенство (9.3) справедливо для любых чисел x_1 и x_2, то $\varphi(x) = ax$, $a > 0$ при некотором значении коэффициента a.

Предположим, мы сравниваем экономическую эффективность двух инвестиционных проектов, используя цены в рублях. Пусть первый проект оказался лучше второго. Теперь перейдем на валюту самой экономически мощной державы мира — юани, используя фиксированный курс пересчета (в работе [89] с помощью расчетов на основе паритета покупательной способности установлено, что в настоящее время валовой внутренний продукт Китая больше, чем у какой-либо иной страны, в частности больше, чем у Европейского союза — 2-е место и США — 3-е место). Очевидно, первый проект должен опять оказаться более выгодным, чем второй. Это следует из общих экономических соображений. Однако алгоритмы расчета не обеспечивают автоматически выполнения этого очевидного условия. Надо проверять, что оно выполнено. Результаты подобной проверки для средних величин описаны далее. Оказалось, что нельзя произвольно выбирать вид средних величин, необходимо согласовывать вид средней со шкалами измерения.

В ***шкале разностей*** есть естественная единица измерения, но нет естественного начала отсчета. Допустимыми преобразованиями в шкале разностей являются сдвиги. Время измеряется по шкале разностей, если год (или сутки — от полудня до полудня) принимаем естественной единицей измерения, и по шкале интервалов — в общем случае. На современном уровне знаний естественного начала отсчета времени указать нельзя. Дату сотворения мира различные авторы рассчитывают по-разному, равно как и момент рождества Христова. Так, согласно новой статистической хронологии, разработанной группой известного историка А.Т. Фоменко, Господь Иисус Христос родился примерно в 1054 г. Позже те же исследователи обосновали несколько иную дату — 1152 год н.э. [63].

При допустимых преобразованиях в шкале разностей сохраняется разность измеряемых величин:

$$x_1 - x_2 = \varphi(x_1) - \varphi(x_2) \tag{9.4}$$

для любых чисел x_1 и x_2 (результатов измерений) и любого допустимого преобразования $\varphi(x) = x + b$. Верно и обратное: если равенство (9.4) справедливо для любых чисел x_1 и x_2, то $\varphi(x) = x + b$ при некотором значении коэффициента сдвига b.

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.

Шесть основных типов шкал измерения описаны в табл. 9.1.

Таблица 9.1

Основные шкалы измерения

Тип шкалы	Определение шкалы	Примеры	Группа допустимых преобразований $\Phi = \{\varphi\}$
Шкалы качественных признаков			
Наименований	Числа используют для различения объектов	Номера телефонов, паспортов, ИНН, штрих-коды	Все взаимнооднозначные преобразования
Порядковая	Числа используют для упорядочения объектов	Оценки экспертов, баллы ветров, отметки в школе, полезность, номера домов	Все строго возрастающие преобразования
Шкалы количественных признаков (описываются началом отсчета и единицей измерения)			
Интервалов	Начало отсчета и единица измерения произвольны	Потенциальная энергия, положение точки, температура по шкалам Цельсия и Фаренгейта	Все линейные преобразования $\varphi(x) = ax + b$, a и b произвольны, $a > 0$
Отношений	Начало отсчета задано, единица измерения произвольна	Масса, длина, мощность, напряжение, сопротивление, температура по Кельвину, цены	Все подобные преобразования $\varphi(x) = ax$, a произвольно, $a > 0$
Разностей	Начало отсчета произвольно, единица измерения задана	Время	Все преобразования сдвига $\varphi(x) = x + b$, b произвольно
Абсолютная	Начало отсчета и единица измерения заданы	Число людей в данном помещении	Только тождественное преобразование $\varphi(x) = x$

Кроме перечисленных в табл. 9.1, используют и иные типы шкал [6; 122]. Отметим, что в табл. 9.1 выражение «единица измерения произвольна» означает, что она может быть выбрана по соглашению специалистов, но не вытекает из каких-либо фундаментальных соотношений. При измерении времени естественная единица измерения задается периодами обращения небесных тел. Начало отсчета при измерении длины задается длиной отрезка, у которого начало и конец совпадают, и т.д.

В настоящее время считается необходимым перед применением тех или иных алгоритмов анализа данных установить, в шкалах каких типов измерены рассматриваемые величины. Отметим, что в процессе развития соответствующей области знания тип шкалы измерения конкретной величины может меняться. Так, сначала температура измерялась по порядковой шкале (холоднее—теплее). Затем — по интервальной (использовались шкалы Цельсия, Фаренгейта, Реомюра). Так, температура °C по шкале Цельсия выражается через температуру °F по шкале Фаренгейта с помощью линейного преобразования (9.1). Наконец, после открытия абсолютного нуля температуру можно 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 как необходимый этап и определение типа шкалы (вместе с обоснованием выбора определенного типа шкалы).

9.2. ИНВАРИАНТНЫЕ АЛГОРИТМЫ И СРЕДНИЕ ВЕЛИЧИНЫ

Основное требование к алгоритмам анализа данных формулируется в ТИ 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.

Требование инвариантности (адекватности) выводов. Выяснение типов используемых шкал необходимо для адекватного выбора методов анализа данных. Основополагающим требованием является независимость выводов от того, какой именно шкалой измерения воспользовался исследователь (среди всех шкал, переходящих друг в друга при допустимых преобразованиях). Например, если речь о длинах, то выводы не должны зависеть от того, измерены ли длины в метрах, аршинах, саженях, футах или дюймах.

Таким образом, одна из основных целей теории измерений — борьба с субъективизмом исследователя при приписывании численных значений реальным объектам. Так, расстояние можно измерять в аршинах, метрах, микронах, милях, парсеках и других единицах, массу (вес) — в пудах, килограммах, фунтах и т.д. 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. *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.* Очевидно, что при разработке управленческих решений можно опираться только на инвариантные выводы.

Другими словами, выводы должны быть инвариантны относительно группы допустимых преобразований шкалы измерения. Только тогда их можно назвать адекватными, т.е. избавленными от субъективизма исследователя, выбирающего определенную шкалу из множества шкал заданного типа, связанных допустимыми преобразованиями. В статье [6] требование инвариантности (адекватности) выводов формулируется как условие «содержательности» («состоятельности») высказывания.

Требование инвариантности выводов накладывает ограничения на множество возможных алгоритмов анализа данных. В качестве примера рассмотрим порядковую шкалу. Одни алгоритмы анализа данных позволяют получать адекватные выводы, другие — нет. Например, в задаче проверки однородности двух независимых выборок алгоритмы ранговой статистики (т.е. использующие только ранги результатов измерений) дают адекватные выводы, а статистики Крамера — Уэлча и Стьюдента — нет. Значит, для обработки данных, измеренных в порядковой шкале, критерии Смирнова и Вилкоксона можно использовать, а критерии Крамера — Уэлча и Стьюдента — нет [89].

Оказывается, требование инвариантности является достаточно сильным. Из многих алгоритмов анализа статистических данных ему удовлетворяют лишь некоторые. Покажем это на примере сравнения средних величин.

Средние величины. Среди всех методов анализа данных важное место занимают алгоритмы усреднения. Еще в 1970-х гг. удалось полностью выяснить, какими видами средних можно пользоваться при анализе данных, измеренных в тех или иных шкалах.

Пусть $X_1, X_2, ..., X_n$ — выборка объема n. Часто используют среднее арифметическое

$$X_{\text{ср}} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Использование среднего арифметического настолько привычно, что второе слово в термине часто опускают. И говорят о средней зарплате, среднем доходе и других средних для конкретных экономических данных, подразумевая под «средним» среднее арифметическое. Такая традиция может приводить к ошибочным выводам. Покажем это на примере расчета средней заработной платы (среднего дохода) работников условного предприятия (табл. 9.2).

Таблица 9.2

Численность работников различных категорий, их заработная плата и доходы (в условных единицах)

№ п/п	Категория работников	Число работников	Заработная плата	Суммарные доходы
1	Низкоквалифицированные рабочие	40	100	4 000
2	Высококвалифицированные рабочие	30	200	6 000
3	Инженеры и служащие	25	300	7 500
4	Менеджеры	4	1000	4 000
5	Генеральный директор (владелец)	1	18 500	18 500
6	*Всего:*	100	—	40 000

Первые три строки в табл. 9.2 вряд ли требуют пояснений. Менеджеры — это директора по направлениям, а именно: по производству (главный инженер), по финансам, по маркетингу и сбыту, по персоналу (по кадрам). Владелец сам руководит предприятием в качестве генерального директора. В столбце «заработная плата» указаны доходы одного работника соответствующей категории, а в столбце «суммарные доходы» — доходы всех работников соответствующей категории.

Фонд оплаты труда составляет 40 000 единиц, работников всего 100, следовательно, средняя заработная плата составляет 40 000/100 = 400 единиц. Однако эта средняя арифметическая величина явно не соответствует интуитивному представлению о «средней зарплате». Из 100 работников лишь 5 имеют заработную плату, ее превышающую, а зарплата остальных 95 существенно меньше средней арифметической. Причина очевидна — заработная плата одного человека — генерального директора — превышает заработную плату 95 работников — низкоквалифицированных и высококвалифицированных рабочих, инженеров и служащих, вместе взятых.

Ситуация напоминает описанную в известном рассказе о больнице, в которой десять больных, из них у девяти температура 40 °С, а один уже отмучился, лежит в морге с температурой 0 °С. Между тем средняя температура по больнице равна 36 °С — лучше не бывает!

Из сказанного ясно, что не всегда целесообразно использовать среднее арифметическое. Его можно порекомендовать лишь для достаточно однородных совокупностей (без больших выбросов в ту или иную сторону).

А какие средние стоит применять для описания заработной платы? Вполне естественно использовать медиану. Для данных табл. 9.2 медиана — среднее арифметическое 50-го и 51-го работника, если их

заработные платы расположены в порядке неубывания. Сначала идут зарплаты 40 низкоквалифицированных рабочих, а затем — с 41-го до 70-го работника — заработные платы высококвалифицированных рабочих. Следовательно, медиана попадает именно на них и равна 200. У 50 работников заработная плата не превосходит 200, и у 50 — не менее 200, поэтому медиана показывает «центр», около которого группируется основная масса исследуемых величин. Еще одна средняя величина — мода, наиболее часто встречающееся значение. В рассматриваемом случае это заработная плата низкоквалифицированных рабочих, т.е. 100. Таким образом, для описания зарплаты имеем три средние величины — моду (100 единиц), медиану (200 единиц) и среднее арифметическое (400 единиц). Для наблюдающихся в реальной экономике распределений доходов и заработной платы справедлива та же закономерность: мода меньше медианы, а медиана меньше среднего арифметического.

Для чего при разработке управленческих решений [87] используются средние величины? Обычно для того, чтобы заменить совокупность чисел одним числом, чтобы сравнивать совокупности с помощью средних.

Пусть, например, $Y_1, Y_2, ..., Y_n$ — 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 квадратичное. 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. Все перечисленные виды средних величин являются средними 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 условием инвариантности выводов, принятым как основное требование в ТИ). 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 ТИ 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 бы меньше среднего преобразованных значений для второй совокупности. Причем сформулированное условие должно быть выполнено для любых двух совокупностей $Y_1, Y_2, ..., Y_n$ и $Z_1, Z_2, ..., Z_n$ и, напомним, любого допустимого преобразования g.

Средние величины, удовлетворяющие сформулированному условию, назовем *допустимыми* (в соответствующей шкале). Согласно ТИ только допустимыми средними величинами можно пользоваться при анализе мнений экспертов и иных данных, измеренных в рассматриваемой шкале.

С помощью математической теории, развитой в монографии [88], удается описать вид допустимых средних величин в основных шкалах. Сразу ясно, что для данных, измеренных в шкале наименований, допустимых средних нет, поскольку допустимые в этой шкале преобразования — а ими являются все взаимно однозначные преобразования — могут как угодно перемешать значения усредняемых величин.

9.3. СРЕДНИЕ ВЕЛИЧИНЫ В ПОРЯДКОВОЙ ШКАЛЕ

Рассмотрим сначала обработку мнений экспертов, измеренных в порядковой шкале. Справедливо следующее утверждение.

Теорема 9.1. *Из всех средних по Коши допустимыми средними в порядковой шкале являются только члены вариационного ряда (порядковые статистики).*

Теорема 9.1, впервые полученная в статье [88], справедлива при условии, что средняя величина $f(X_1, X_2, ..., X_n)$ является непрерывной (по совокупности переменных) и симметрической функцией. Последнее означает, что при перестановке аргументов значение функции $f(X_1, X_2, ..., X_n)$ не меняется. Это условие является вполне естественным, ибо среднюю величину находим для *совокупности (множества)* чисел, а не для *их последовательности*. Множество не меняется в зависимости от того, в какой последовательности мы перечисляем его элементы.

Согласно теореме 9.1 в качестве среднего для данных, измеренных в порядковой шкале, можно использовать, в частности, медиану (при нечетном объеме выборки). При четном же объеме целесообразно применять один из двух центральных членов вариационного ряда — левую медиану или правую медиану, как их иногда называют. Моду тоже можно использовать — она всегда является членом вариационного ряда. Можно применять выборочные квартили, минимум и максимум, децили и т.п. Но теорема 9.1 запрещает использовать при анализе порядковых данных (т.е. измеренных в порядковой шкале) среднее арифметическое, среднее геометрическое и т.д. Таким образом, не рекомендуется разрабатывать управленческое решение на основе среднего арифметического или среднего геометрического мнений экспертов, поскольку такие мнения, как разъяснено ранее, обычно измерены в порядковой шкале.

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 $g(x) = x$ для 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.

Таким образом, ТИ выносит жесткий приговор среднему арифметическому: использовать его в порядковой шкале нельзя. Однако же те, кто не знает ТИ, используют его. Всегда ли они ошибаются? Оказывается, можно в какой-то мере (но отнюдь не полностью!) реабилитировать среднее арифметическое, если перейти к вероятностной постановке и к тому же удовлетвориться результатами для больших объемов выборок. В монографии [88], кроме теоремы 9.1, доказано также следующее утверждение.

Теорема 9.2. *Пусть $Y_1, Y_2, ..., Y_m$ — независимые одинаково распределенные случайные величины с функцией распределения $F(x)$, а $Z_1, Z_2, ..., Z_n$ — независимые одинаково распределенные случайные величины с функцией распределения $H(x)$, причем выборки $Y_1, Y_2, ..., Y_m$ и $Z_1, Z_2, ..., Z_n$ независимы между собой и $MY_1 > MZ_1$. Для того чтобы вероятность события*

$$\left\{\omega : \frac{g(Y_1)+g(Y_2)+\ldots+g(Y_m)}{m} > \frac{g(Z_1)+g(Z_2)+\ldots+g(Z_n)}{n}\right\}$$

стремилась к 1 *при* $\min(m, n) \to \infty$ *для любой строго возрастающей непрерывной функции g, удовлетворяющей условию*

$$\overline{\lim_{|x|\to\infty}}\left|\frac{g(x)}{x}\right| < \infty,$$

необходимо и достаточно, чтобы при всех x выполнялось неравенство $F(x) \le H(x)$, *причем существовало число* x_0, *для которого* $F(x_0) < H(x_0)$.

Примечание. Условие с верхним пределом носит чисто внутриматематический характер. Фактически функция g — произвольное допустимое преобразование в порядковой шкале.

Согласно теореме 9.2 средним арифметическим можно пользоваться и в порядковой шкале, если сравниваются выборки из двух распределений, удовлетворяющих приведенному в теореме неравенству. Проще говоря, одна из функций распределения должна всегда лежать над другой. Функции распределения не могут пересекаться, им разрешается только касаться друг друга. Это условие выполнено, например, если функции распределения отличаются только сдвигом, т.е.

$$F(x) = H(x + b)$$

при некотором b. Последнее условие выполняется, если два значения некоторой величины измеряются с помощью одного и того же средства измерения, у которого распределение погрешностей не меняется при переходе от измерения одного значения рассматриваемой величины к измерению другого.

9.4. СРЕДНИЕ ПО КОЛМОГОРОВУ

Естественная система аксиом (требований к средним величинам) приводит к так называемым ассоциативным средним. Их общий вид нашел в 1930 г. А.Н. Колмогоров [88]. Теперь их называют «средними по Колмогорову».

Для чисел $X_1, X_2, ..., X_n$ средним по Колмогорову является

$$G\{[F(X_1) + F(X_2) + ...+ F(X_n)]/n\},$$

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, если $F(x) = \ln x$, то это среднее геометрическое, если $F(x) = 1/x$, то это среднее гармоническое, если $F(x) = x^2$, то это среднее квадратичное и т.д. (в последних трех случаях усредняются положительные величины).

Средние по Колмогорову — частный случай средних по Коши. С другой стороны, такие популярные средние, как медиана и мода, нельзя представить в виде средних по Колмогорову.

Теорема 9.3. *При справедливости некоторых внутриматематических условий регулярности в шкале интервалов из всех средних по Колмогорову допустимым является только среднее арифметическое.*

Таким образом, среднее геометрическое или среднее квадратичное температур (в шкале Цельсия), потенциальных энергий или координат точек не имеют смысла. В качестве среднего надо применять среднее арифметическое. Также можно использовать медиану или моду — они не входят в число средних по Колмогорову.

Теорема 9.4. *При справедливости некоторых внутриматематических условий регулярности в шкале отношений из всех средних по Колмогорову допустимыми являются только степенные средние с* $F(x) = x^c$, $c \ne 0$, *и среднее геометрическое.*

Есть ли средние по Колмогорову, которыми нельзя пользоваться в шкале отношений? Конечно, есть, например, с $F(x) = e^x$.

Замечание 1. Среднее геометрическое является пределом степенных средних при $c \to 0$.

Замечание 2. Подробное описание «внутриматематических условий регулярности», упомянутых в формулировках теорем 9.3 и 9.4, можно найти в [80; 88]. Доказательства теорем 9.1—9.4 приведены в монографии [88]. Перенос на случай взвешенных средних дан в статье [80].

Аналогично средним величинам могут быть изучены и другие статистические характеристики — показатели разброса, связи, расстояния и др. (см. [88, 122]). Нетрудно показать, например, что коэффициент корреляции не меняется при любом допустимом преобразовании в шкале интервалов, как и отношение дисперсий. Дисперсия не меняется в шкале разностей, коэффициент вариации — в шкале отношений и т.д.

Приведенные результаты о средних величинах широко применяются, причем не только в экспертных исследованиях, но и в теории принятия решений, экономике, менеджменте, социологии, медицине, инженерном деле, например, для анализа методов агрегирования датчиков в автоматизированных системах управления технологическими процессами (АСУ ТП) доменных печей. Велико прикладное значение теории измерений в задачах стандартизации и управления качеством, в частности в квалиметрии [88]. В обзоре [6] проанализированы многочисленные работы последних десятилетий, посвященные связи теории измерений и теории средних величин.

При подготовке и принятии инженерных, технико-экономических и иных решений необходимо использовать только инвариантные алгоритмы обработки данных. Требование инвариантности выделяет из многих алгоритмов усреднения лишь некоторые, соответствующие используемым шкалам измерения. Инвариантные алгоритмы в общем случае рассматриваются в математической теории измерений [102]. Нацеленное на прикладные исследования изложение ряда вопросов теории измерений дается в монографиях [88; 89; 95; 122].

Контрольные вопросы

1. Всегда ли имеет смысл складывать числа, используемые в той или иной области человеческой деятельности?
2. Какие величины измеряют в шкале наименований?
3. Какие величины измеряют в порядковой шкале?
4. Какие величины измеряют в шкале интервалов?
5. Какие величины измеряют в шкале отношений?
6. Какие средние величины целесообразно использовать при расчете средней заработной платы (или среднего дохода)?
7. Как соотносятся средние по Коши и средние по Колмогорову?

Темы докладов и рефератов

1. Теория измерений как научная дисциплина, посвященная гомоморфизмам эмпирических систем с отношениями в числовые системы с отношениями.
2. Показатели разброса, связи, показатели различия (в том числе метрики) в порядковой шкале.
3. Ранговые методы математической статистики как инвариантные методы анализа порядковых данных.
4. Показатели разброса, связи, показатели различия (в том числе метрики) в шкале интервалов.
5. Показатели разброса, связи, показатели различия (в том числе метрики) в шкале отношений.
6. Теорема В.В. Подиновского: любое изменение коэффициентов весомости единичных показателей качества продукции приводит к изменению упорядочения изделий по средневзвешенному показателю (доказательство и прикладное значение).

ГЛАВА 10
МЕТОДЫ СРЕДНИХ РАНГОВ

10.1. ЭКСПЕРТНЫЕ РАНЖИРОВКИ

Современная теория измерений и экспертные оценки. Как проводить анализ собранных рабочей группой ответов экспертов? Для более углубленного рассмотрения проблем экспертных оценок понадобятся некоторые понятия ТИ (см. главу 9), служащей основой теории экспертных оценок, прежде всего той ее части, которая связана с анализом заключений экспертов, выраженных в качественном (а не в количественном) виде. Теория измерений интересует нас, в частности, в связи с агрегированием мнений экспертов, построением обобщенных показателей (их называют также рейтингами).

Получаемые от экспертов мнения часто выражены в порядковой шкале, т.е. эксперт может сказать (и обосновать), что один тип продукции будет более привлекателен для потребителей, что один показатель качества продукции более важен, чем другой, первый технологический объект более опасен, чем второй, и т.д. Но эксперт не в состоянии сказать, *во сколько раз* или *на сколько* объект важнее, соответственно, опаснее. Поэтому экспертов часто просят дать ранжировку (упорядочение) объектов экспертизы, т.е. расположить их в порядке возрастания (или, точнее, неубывания) интенсивности интересующей организаторов экспертизы характеристики.

Ранжировки определяются и изучаются с помощью рангов. Поэтому очевидно, что для анализа подобного рода качественных данных необходима не обычная арифметика, а другая теория, дающая базу для разработки, изучения и применения конкретных методов расчета. Эта другая теория и есть ТИ. Основы ТИ рассмотрены в главе 9.

Рассмотрим в качестве примера применения результатов ТИ, касающихся средних величин в порядковой шкале, один сюжет, связанный с ранжировками и рейтингами.

Сравнение на основе средних баллов. В настоящее время распространены экспертные, маркетинговые, квалиметрические, социологические и иные опросы, в которых используются балльные оценки. В таких исследованиях опрашиваемых просят выставить баллы объектам, изделиям, технологическим процессам, предприятиям, проектам или же заявкам на выполнение научно-исследовательских работ, идеям, проблемам, программам, политикам и т.п. Затем рассчитывают средние баллы и рассматривают их как *интегральные (т.е. обобщенные, итоговые) оценки,* выставленные объектам экспертизы коллективом опрошенных экспертов. Какими формулами пользоваться для вычисления средних величин? Ведь средних величин существует, как мы знаем, весьма много разных видов.

По традиции обычно применяют *среднее арифметическое*. Специалисты по ТИ уже более 30 лет знают, что такой способ некорректен, поскольку баллы обычно измерены в порядковой шкале. Обоснованным является использование медиан в качестве средних баллов. Однако полностью игнорировать средние арифметические нецелесообразно из-за их привычности и распространенности. Поэтому *представляется рациональным использовать одновременно оба метода — и метод средних арифметических баллов, и метод медиан баллов*. Такая рекомендация находится в согласии с общенаучной концепцией устойчивости [88], рекомендующей применять 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. В качестве баллов будем использовать ранги, присвоенные проектам в соответствии с их упорядочениями, полученными в результате работы экспертов.

В рассматриваемом примере по заданию руководства фирмы анализировались восемь проектов, предлагаемых для включения в план стратегического развития фирмы. Они 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, включенным в экспертную комиссию, организованную по решению правления фирмы. В таблице 10.1 приведены ранги восьми проектов, присвоенные им каждым из 12 экспертов.

Ранги присваивались в соответствии с представлениями экспертов о целесообразности включения проектов 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 в последнюю очередь.

Таблица 10.1

Ранги восьми проектов по степени привлекательности для включения в план стратегического развития фирмы

№ эксперта	Проект							
	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 2-м и 3-м местах и получить баллы 2 и 3. Поскольку они равноценны, то получают средний балл (2 + 3)/ 2 = 5/2 = 2,5.

Анализируя результаты работы экспертов (т.е. табл. 10.1), члены аналитического подразделения рабочей группы, анализировавшие ответы экспертов по заданию правления фирмы, были вынуждены констатировать, что полного согласия между экспертами нет, а потому данные, приведенные в табл. 10.1, следует подвергнуть более тщательному математическому анализу.

10.2. МЕТОДЫ СРЕДНИХ АРИФМЕТИЧЕСКИХ И МЕДИАН РАНГОВ

Метод средних арифметических рангов. Сначала для получения группового мнения экспертов был применен метод средних арифметических рангов. Для этого прежде всего была подсчитана сумма рангов, присвоенных проектам (см. табл. 10.1). Затем эта сумма была разделена на число экспертов, в результате рассчитан средний арифметический ранг (именно эта операция дала название методу). По средним рангам построена итоговая ранжировка (в другой терминологии — упорядочение), исходя из принципа — чем меньше средний ранг, тем 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-м и 4-м местах и получают средний балл (3 + 4)/2 = 3,5. Дальнейшие результаты приведены в табл. 10.2.

Таблица 10.2

Результаты расчетов по методу средних арифметических и методу медиан для данных, приведенных в табл. 10.1

Показатель	Проект							
	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*

Итак, ранжировка по суммам рангов (или, что то же самое, по средним арифметическим рангам) имеет вид:

$$Б < М\text{-}К < \{Л, Сол\} < Д < Стеф < Г\text{-}Б < К. \qquad (10.1)$$

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 (10.1) имеет одну связь.

Метод медиан рангов. Значит, наука сказала свое слово, итог расчетов — ранжировка (10.1), и на ее основе предстоит принимать решение? Так был поставлен вопрос при обсуждении полученных результатов на заседании правления фирмы. Но тут наиболее знакомый с современной эконометрикой член правления вспомнил то, о чем шла речь ранее. Он понял, 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 несколько непривычный термин «неубывание»). Получим последовательность: 1, 1, 5, 5, 5, 5, 5, 6, 6, 6, 7, 8. На центральных местах — 6-м и 7-м — стоят 5 и 5. Следовательно, медиана равна их среднему арифметическому, т.е. 5.

Медианы совокупностей из 12 рангов, соответствующих определенным проектам, приведены в предпоследней строке табл. 10.2. При этом медианы вычислены по обычным правилам статистики — как среднее арифметическое центральных членов вариационного ряда. Если бы число экспертов было нечетным, в качестве медианы надо было бы взять центральный член вариационного ряда. Итоговое упорядочение комиссии экспертов по методу медиан приведено в последней строке табл. 10.2. Ранжировка (т.е. упорядочение — итоговое мнение комиссии экспертов) по медианам имеет вид:

$$Б < \{М\text{-}К, Л\} < Сол < Д < Стеф < К < Г\text{-}Б. \qquad (10.2)$$

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 (10.2) имеет одну связь.

Сравнение ранжировок по методу средних арифметических и методу медиан. Сравнение ранжировок (10.1) и (10.2) 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 (10.1), а в другом — проекты М-К и Л — ранжировка (10.2). Существенным является только расхождение, касающееся упорядочения проектов К и Г-Б: в ранжировке (10.1) Г-Б < К, а в ранжировке (10.2), наоборот, К < Г-Б. Однако эти проекты наименее привлекательные из восьми рассматриваемых, и при выборе наиболее привлекательных проектов для дальнейшего обсуждения и использования на указанное расхождение можно не обращать внимания.

Рассмотренный пример демонстрирует сходство и различие ранжировок, полученных по методу средних арифметических рангов и по методу медиан, а также пользу от их совместного применения.

10.3. МЕТОД СОГЛАСОВАНИЯ КЛАСТЕРИЗОВАННЫХ РАНЖИРОВОК

Только что проведенное сравнение ранжировок по методу средних арифметических и методу медиан оставляет ощущение недостаточно строгого подхода. Обсудим проблему согласования кластеризованных ранжировок в общем виде и разберем математический алгоритм такого согласования.

Постановка задачи. 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. *Предлагается применять метод согласования кластеризованных ранжировок, позволяющий «загнать» противоречия внутрь специальным образом построенных кластеров (групп), в то время как упорядочение кластеров соответствует одновременно всем исходным упорядочениям.*

В различных прикладных областях возникает необходимость анализа нескольких кластеризованных ранжировок объектов. К таким областям относятся, прежде всего, инженерный бизнес, менеджмент, экономика, социология, экология, прогнозирование, научные и технические исследования и т.д. Особенно те их разделы, что связаны с экспертными оценками (см., например, [16; 89]). 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 далее метод был разработан в связи с проблемами химической безопасности биосферы и экологического страхования [16].

Рассмотрим ***метод построения кластеризованной ранжировки, согласованной*** (в раскрытом далее 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, которые могут быть как формально-математическими (например, вычисление медианы Кемени (о ней — далее), упорядочения внутри группы по средним рангам или по медианам с привлечением новых экспертов 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 их совокупность «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 (весь носитель).

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 записи изображать без фигурных скобок. Тогда кластеризованную ранжировку на основе введенных ранее 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, состоящему не менее чем из двух элементов, то можно сказать, что в кластеризованную ранжировку A входят два кластера {2, 3} и {5, 6, 7} и пять отдельных элементов.

Кластеризованная ранжировка, введенная описанным образом, является бинарным отношением на носителе — множестве {1, 2, 3, ..., 10}. Его структура такова. Задано отношение эквивалентности с семью классами эквивалентности, а именно {2, 3}, {5, 6, 7}, остальные пять классов состоят из оставшихся пяти отдельных элементов. Затем введен строгий линейный порядок между классами эквивалентности.

Рассматриваемый математический объект известен в литературе как «ранжировка со связями» (М. Холлендер, Д. Вулф [131]), «упорядочение» (Дж. Кемени, Дж. Снелл [25]), «квазисерия» (Б.Г. Миркин [53]), «совершенный квазипорядок» (Ю.А. Шрейдер [135, с. 127; 130]). Учитывая разнобой в терминологии, было признано полезным ввести 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 Ю.А. Шрейдера [135, глава 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 кластеризованных ранжировок 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, поскольку эквивалентность $a = b$ не образует «противоречия» ни с $a < b$, ни с $a > b$. Это свойство оказывается полезным при выделении противоречивых пар.

В качестве примера рассмотрим, кроме A, еще 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 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 последней пары (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, а именно {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 $k \times k$. 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 всегда 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 (двух или более) кластеризованных ранжировок состоит из трех этапов:

- на первом этапе 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 имеется между выбранными объектами в любой из рассматриваемых кластеризованных ранжировок (если в одной из исходных кластеризованных ранжировок имеется равенство, а в другой неравенство, то при построении итоговой кластеризованной ранжировки используется неравенство).

Корректность подобного упорядочивания, т.е. его независимость от выбора той или иной пары объектов при упорядочении двух кластеров и транзитивность такого упорядочения вытекает из соответствующих теорем, доказанных в статье [16].

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 третьему этапу 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

$$f(A, B) = [1 < 2 < 3 < 4 < 5 < 6 < 7 < \{8, 9\} < 10],$$

$$f(A, C) = [\{1, 3\} < \{2, 4\} < 6 < \{5, 7\} < 8 < 9 < 10],$$

$$f(B, C) = [\{1, 2, 3, 4\} < \{5, 6\} < 7 < \{8, 9\} < 10],$$

$$f(A, B, C) = f(B, C) = [\{1, 2, 3, 4\} < \{5, 6\} < 7 < \{8, 9\} < 10].$$

Итак, в случае $f(A, B)$ дополнительного изучения в целях упорядочения требуют только объекты 8 и 9. В случае $f(A, C)$ кластер $\{5, 7\}$ появился не потому, что относительно объектов 5 и 7 имеется противоречие, а потому, что в обеих исходных ранжировках эти объекты не различаются. В случае $f(B, C)$ четыре объекта 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.

Рассмотрим некоторые свойства алгоритмов согласования.

1. 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, ..., причем хотя бы в одной из них справедливо строгое неравенство.

2. 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*.

3. 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 ранее, 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.

4. 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 [89; 50]. Однако из ТИ известно (см. главу 9), что более обоснованным является использование не средних арифметических, а медиан. Вместе с тем метод средних арифметических рангов весьма известен и широко применяется, так что просто отбросить его нецелесообразно. Участвующие в исследовании и привыкшие к методу средних арифметических рангов специалисты не поймут и не примут такого решения рабочей группы. Поэтому было решено одновременно применить оба метода. Реализация этого решения потребовала разработки приведенной ранее методики согласования двух указанных кластеризованных ранжировок. Практическая апробация метода продемонстрировала правильность принятого решения об одновременном использовании метода средних арифметических рангов и метода медиан рангов.

Отметим, что во многих случаях кластеризованные ранжировки, полученные двумя методами, совпадали или были весьма близки, как в примере, рассмотренном в подразделе 10.2. Теоретическое объяснение этому экспериментальному факту дает теорема 9.2. Можно сказать, что в случае, когда объекты реально упорядочены, этот порядок выявит любой способ анализа данных. Проблема в том, что мы не знаем заранее, упорядочены ли объекты в действительности или нет. И одновременное применение двух (или более) методов позволяет найти ответ на этот вопрос. Если результаты анализа данных совпадают или почти совпадают, повышается уверенность в том, что они отражают действительность. Если результаты, полученные с помощью двух методов анализа данных, весьма различаются значит они не отражают реальность. Выводы, зависящие от субъективного выбора исследователем того или иного метода анализа данных, не могут использоваться для принятия объективного решения.

5. Область применения рассматриваемого метода не ограничивается экспертными оценками. Он может быть использован, например, для сравнения качества математических моделей процесса испарения жидкости. Имелись данные экспериментов и результаты расчетов по восьми математическим моделям. Сравнивать модели можно по различным критериям качества. Например, по сумме модулей относительных отклонений расчетных и экспериментальных значений. Можно действовать и по-другому, например в каждой экспериментальной точке упорядочить модели по качеству, а потом получать единые оценки методами средних рангов и медиан. Использовались и иные методы. Затем применялись методы согласования полученных различными способами кластеризованных ранжировок. В результате оказалось возможным упорядочить модели по качеству и использовать это упорядочение при разработке банка математических моделей, используемого в задачах химической безопасности биосферы.

6. Рассматриваемый метод согласования кластеризованных ранжировок построен в соответствии с методологией теории устойчивости [88], 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.

Контрольные вопросы

1. Чем метод средних арифметических рангов отличается от метода медиан рангов?
2. Почему метод средних арифметических рангов неприемлем с точки зрения теории измерений?
3. Дайте определение понятия «кластеризованная ранжировка».
4. Почему необходимо согласование кластеризованных ранжировок и как оно проводится?

Темы докладов и рефератов

1. Метод согласования кластеризированных ранжировок.
2. Варианты метода средних баллов.
3. Методы определения итогового мнения комиссии экспертов.
4. Теория измерений и традиции — источники подходов к нахождению группового мнения.

ГЛАВА 11
МАТЕМАТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ЭКСПЕРТНЫХ ОЦЕНОК

11.1. ОСНОВНЫЕ МАТЕМАТИЧЕСКИЕ ЗАДАЧИ АНАЛИЗА ЭКСПЕРТНЫХ ОЦЕНОК

Ясно, что при анализе мнений экспертов можно применять самые разнообразные статистические методы, описывать их — значит описывать практически 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, т.е. разделение их на группы сходных по мнению, 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 следует из условного (т.е. определяемого произвольно принятыми соглашениями, обычно оформленными в виде инструкций) характера балансовой прибыли, амортизационных отчислений и других экономических показателей. В этом одна из причин того, что 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 [87], а более подробно, со всеми доказательствами — в монографии [88]).

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 подобной обработки данных, могут не иметь никакого отношения к реальности.

В связи с «оцифровкой» уместно вспомнить классическую притчу о человеке, который ищет потерянные ключи под фонарем, хотя потерял их в кустах. На вопрос, почему он так делает, отвечает: «Под фонарем светлее». Это, конечно, верно. Но, к сожалению, весьма малы шансы найти потерянные ключи под фонарем. Так и с «оцифровкой» нечисловых данных. Она дает возможность имитации научной деятельности, но не возможность найти истину.

В соответствии с теорией измерений выводы, полученные на основе анализа мнений экспертов, должны быть инвариантны относительно допустимых преобразований шкал измерений, в случае порядковой шкалы — относительно любого строго возрастающего преобразования.

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. Как известно, в рамках методологии математической статистики 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)* [77] *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 разделить 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. далее) нашла многочисленных продолжателей. Однако методы кластер-анализа обычно являются эвристическими. В частности, обычно 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 [87].

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 арифметическим 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.

Напомним, что каждую кластеризованную ранжировку, как и любое бинарное отношение, можно задать квадратной матрицей $\|x(a, b)\|$ из 0 и 1 порядка $k \times k$. 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.

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 11.1. *Расстоянием Кемени между бинарными отношениями A и B, описываемыми матрицами $\|a(i, j)\|$ и $\|b(i, j)\|$ соответственно, называется число*

$$D(A, B) = \sum_{i,\, j=1}^{k} |a(i, j) - b(i, j)|,$$

т.е. расстояние Кемени между бинарными отношениями равно сумме модулей разностей элементов, стоящих на одних и тех же местах в матрицах, соответствующих этим бинарным отношениям.

Легко видеть, что расстояние Кемени — это число несовпадающих элементов в матрицах $\|a(i, j)\|$ и $\|b(i, j)\|$.

Вид расстояния Кемени не выбран произвольно. Он основан на некоторой системе аксиом. Эта система аксиом и вывод из нее формулы для расстояния Кемени между упорядочениями содержится в книге [25], которая сыграла большую роль в развитии в нашей стране такого научного направления, как анализ нечисловой информации [4; 5]. 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 [88].

Медиана Кемени и законы больших чисел. С помощью расстояния Кемени находят итоговое мнение комиссии экспертов. Пусть A_1, A_2, A_3, ..., A_p — ответы p экспертов, представленные в виде бинарных отношений. Для их усреднения используют ***медиану Кемени***:

$$\operatorname{Arg}\min_{\{A\}} \sum_{i=1}^{p} D(A_i, A),$$

где Arg min — 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,

$$\sum_{i=1}^{p} D(A_i, A) = D(A_1, A) + D(A_2, A) + D(A_3, A) + \ldots + D(A_p, A).$$

Кроме медианы Кемени, используют и другие средние величины в пространстве бинарных отношений, например введенное в [25] ***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 [77]. Для нее справедлив закон больших чисел, т.е. эмпирическое среднее приближается при росте числа составляющих (p — числа слагаемых в сумме) к теоретическому среднему:

$$\operatorname{Arg}\min_{\{A\}}\sum_{i=1}^{p}D(A_i, A)\to\operatorname{Arg}\min_{\{A\}}M\left[D(A_1, A)\right],$$

где M — символ математического ожидания.

Предполагается, что ответы p экспертов $A_1, A_2, A_3, ..., A_p$ есть основания рассматривать как независимые одинаково распределенные случайные элементы (т.е. как случайную выборку) в соответствующем пространстве бинарных отношений, например в пространстве упорядочений или отношений эквивалентности. Систематически эмпирические и теоретические средние и соответствующие различные варианты законов больших чисел изучены в ряде работ (см., например, [77; 89]).

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 пространствах произвольной природы), 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. Для ее нахождения используются 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 пример вычисления медианы Кемени. Пусть дана квадратная матрица (порядка 9) попарных расстояний для множества бинарных отношений из 9 элементов $A_1, A_2, A_3, ..., A_9$ (табл. 11.1). Найдем в этом множестве медиану для множества из 5 элементов $\{A_2, A_4, A_5, A_8, A_9\}$.

Таблица 11.1

Матрица попарных расстояний

Элементы	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
A_1	0	2	13	1	7	4	10	3	11
A_2	2	0	5	6	1	3	2	5	1
A_3	13	5	0	2	2	7	6	5	7
A_4	1	6	2	0	5	4	3	8	8
A_5	7	1	2	5	0	10	1	3	7
A_6	4	3	7	4	10	0	2	1	5
A_7	10	2	6	3	1	2	0	6	3
A_8	3	5	5	8	3	1	6	0	9
A_9	11	1	7	8	7	5	3	9	0

В соответствии с определением медианы Кемени следует ввести в рассмотрение функцию

$$C(A)=\sum_{i\in\{2,4,5,8,9\}}D(A_i, A)=$$
$$=D(A_2, A)+D(A_4, A)+D(A_5, A)+D(A_8, A)+D(A_9, A),$$

рассчитать ее значения для всех $A = A_1, A_2, A_3, ..., A_9$ 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.

В данном случае медиана Кемени — одно из исходных экспертных мнений. В общем случае медиана Кемени может не совпадать ни с одним из мнений экспертов. Последнее обстоятельство является поводом для критики рассматриваемого способа усреднения. Действительно, если представить себе, что ответы экспертов равномерно распределены по поверхности бублика (в математической терминологии — тора), то медиана Кемени — центр бублика — лежит в пустоте, следовательно, далека от мнений кого-либо из экспертов.

Выход из этого парадокса может быть найден путем изменения области минимизации $\{A\}$ в определении медианы Кемени. Действительно, если положить $\{A\} = \{A_1, A_2, A_3, ..., A_p\}$, то, очевидно, решением задачи минимизации будет одно из экспертных мнений. Такое среднее назовем «модифицированной медианой Кемени».

Преимуществом модифицированной медианы Кемени является значительно меньшая вычислительная трудоемкость. Если для расчета медианы Кемени необходимо применять специальные алгоритмы дискретной оптимизации (см., например, [41]), то модифицированную медиану Кемени можно найти без привлечения компьютера, как это и продемонстрировано ранее.

11.2. ЭКСПЕРТНЫЕ МНЕНИЯ И РАССТОЯНИЯ МЕЖДУ НИМИ

Как показано ранее, мнения экспертов могут иметь разнообразную математическую природу, являться элементами разнообразных пространств — конечномерных, функциональных, бинарных отношений, множеств, нечетких множеств и т.д. Следовательно, центральной частью математического аппарата теории экспертных оценок является статистика в пространствах произвольной природы [77]. Эта область прикладной статистики сама по себе не используется при анализе конкретных данных, поскольку конкретные данные всегда имеют вполне определенную природу. Однако общие подходы, методы, результаты статистики в пространствах произвольной природы представляют собой научный инструментарий, готовый к применению в каждой конкретной области.

Статистика в пространствах произвольной природы. Много ли общего у статистических методов анализа данных различной природы? На этот естественный вопрос можно сразу же однозначно ответить — да, очень много.

Прежде всего отметим, что понятия случайного события, вероятности, независимости событий и случайных величин являются общими для любых конечных вероятностных пространств и любых конечных областей значений случайных величин (см., например, [77, гл. 2]). Поскольку все реальные явления и процессы можно описывать с помощью математических объектов, являющихся элементами конечных множеств, сказанное ранее означает, что конечных вероятностных пространств и дискретных случайных величин (точнее, величин, принимающих значения в конечном множестве) вполне достаточно для всех практических применений. Переход к непрерывным моделям реальных явлений и процессов оправдан только тогда, когда этот переход облегчает проведение рассуждений и выкладок. Например, находить определенные интегралы зачастую проще, чем вычислять значения сумм. Нельзя не отметить, что приведенные соображения о взаимном соотнесении дискретных и непрерывных математических моделей автор услышал более 30 лет назад от академика А.Н. Колмогорова (ясно, что за конкретную формулировку несет ответственность автор данного учебника).

Основные проблемы прикладной статистики — описание данных, оценивание, проверка гипотез — также в своей существенной части могут быть рассмотрены в рамках статистики в пространствах произвольной природы. Например, для описания данных могут быть использованы эмпирические и теоретические средние, плотности вероятностей и их непараметрические оценки, регрессионные зависимости. Правда, для этого пространства произвольной природы должны быть снабжены соответствующим математическим инструментарием — расстояниями (показателями близости, мерами различия) между элементами рассматриваемых пространств.

Популярный в настоящее время метод оценивания параметров распределений — метод максимального правдоподобия — не накладывает каких-либо ограничений на конкретный вид элементов выборки. Они могут лежать в пространстве произвольной природы. Математические условия касаются только свойств плотностей вероятности и их производных по параметрам. Аналогично положение с методом одношаговых оценок, идущим на смену методу максимального правдоподобия [77, глава 6]. Асимптотику решений экстремальных статистических задач достаточно изучить для пространств произвольной природы, а затем применять в каждом конкретном случае, когда задачу прикладной статистики удается представить в оптимизационном виде [77]. Общая теория проверки статистических гипотез также не требует конкретизации математической природы рассматриваемых элементов выборок. Это относится, например, к лемме Неймана — Пирсона или теории статистических решений. Более того, естественная область построения теории статистик интегрального типа — это не числовая прямая, а пространства произвольной природы [77, подраздел 7.3].

Совершенно ясно, что в конкретных областях прикладной статистики накоплено большое число результатов, относящихся именно к этим областям. Особенно это касается областей, исследования в которых ведутся сотни лет, в частности статистики случайных величин (одномерной статистики). Однако принципиально важно указать на «ядро» прикладной статистики — статистику в пространствах произвольной природы. Если постоянно «держать в уме» это ядро, то становится ясно, что, например, многие методы непараметрической оценки плотности вероятности или кластер-анализа, использующие только расстояния между объектами и элементами выборки, относятся именно к статистике объектов произвольной природы, а не к статистике случайных величин или многомерному статистическому анализу. Следовательно, и применяться они могут во всех областях прикладной статистики, а не только в тех, в которых «родились».

Расстояния (метрики). В пространствах произвольной природы нет операции сложения, поэтому статистические процедуры не могут быть основаны на использовании сумм и применяется другой математический инструментарий, использующий такие понятия, как расстояния.

Как известно, ***расстоянием*** в пространстве X называется числовая функция двух переменных $d(x, y)$, $x \in X$, $y \in X$, определенная на этом пространстве, т.е. в стандартных обозначениях $d: X^2 \to R^1$, где R^1 — прямая, т.е. множество всех действительных чисел. Эта функция должна удовлетворять трем условиям (иногда их называют аксиомами):

1) неотрицательности: $d(x, y) \geq 0$, причем $d(x, x) = 0$ для любых значений $x \in X$, $y \in X$;
2) симметричности: $d(x, y) = d(y, x)$ для любых $x \in X$, $y \in X$;
3) неравенства треугольника: $d(x, y) + d(y, z) \geq d(x, z)$ для любых значений $x \in X$, $y \in X$, $z \in X$.

Для термина «расстояние» часто используется синоним «метрика».

Пример 11.1. Если $d(x, x) = 0$ и $d(x, y) = 1$ при $x \neq y$ для любых значений $x \in X$, $y \in X$, то, как легко проверить, функция $d(x, y)$ — расстояние (*метрика*). Такое расстояние естественно использовать в пространстве X значений номинального признака: если два значения (например, названные двумя экспертами) совпадают, то расстояние равно 0, а если различны — то 1.

Пример 11.2. *Расстояние, используемое в геометрии*, очевидно, удовлетворяет трем приведенным ранее аксиомам. Если X — это плоскость, а $x(1)$ и $x(2)$ — координаты точки $x \in X$ в некоторой прямоугольной системе координат, то эту точку естественно отождествить с двумерным вектором $(x(1), x(2))$. Тогда расстояние между точками $x = (x(1), x(2))$ и $y = (y(1), y(2))$, согласно известной формуле аналитической геометрии, равно

$$d(x, y) = \sqrt{\left[x(1) - y(1)\right]^2 + \left[x(2) - y(2)\right]^2}.$$

Пример 11.3. *Евклидовым расстоянием* в пространстве R^k векторов вида $x = (x(1), x(2), ..., x(k))$ и $y = (y(1), y(2), ..., y(k))$ размерности k называется

$$d(x, y) = \left(\sum_{j=1}^{k}\left[x(j) - y(j)\right]^2\right)^{1/2}.$$

В примере 11.2 рассмотрен частный случай с $k = 2$.

Пример 11.4. В пространстве R^k векторов размерности k используют также так называемое *блочное расстояние*, имеющее вид

$$d(x, y) = \sum_{j=1}^{k}\left|x(j) - y(j)\right|.$$

Блочное расстояние соответствует передвижению по городу, разбитому на кварталы горизонтальными и вертикальными улицами. В результате можно передвигаться только параллельно одной из осей координат.

Пример 11.5. В пространстве функций, элементами которого являются функции $x = x(t)$, $y = y(t)$, $0 \leq t \leq 1$, часто используют *расстояние Колмогорова*

$$d(x, y) = \sup_{0 \leq t \leq 1}\left|x(j) - y(j)\right|.$$

Пример 11.6. Пространство функций, элементами которого являются функции $x = x(t)$, $y = y(t)$, $0 \leq t \leq 1$, превращают в *ме-*

трическое пространство (т.е. в пространство с метрикой), вводя расстояние

$$d_p(x, y) = \left(\int_0^1 [x(t) - y(t)]^p dt \right)^{1/p}.$$

Это пространство обычно обозначают L^p, где параметр $p \geq 1$ (при $p < 1$ не выполняются аксиомы метрического пространства, в частности аксиома треугольника).

Пример 11.7. Рассмотрим пространство квадратных матриц порядка k. Как ввести расстояние между матрицами $A = \|a(i, j)\|$ и $B = \|b(i, j)\|$? Можно сложить расстояния между соответствующими элементами матриц:

$$d(A, B) = \sum_{i=1}^{k} \sum_{j=1}^{k} |a(i, j) - b(i, j)|.$$

Пример 11.8. Предыдущий пример наводит на мысль о следующем полезном свойстве расстояний. Если на некотором пространстве определены два или больше расстояний, то их сумма — также расстояние.

Пример 11.9. Пусть A и B — множества. Расстояние между множествами можно определить формулой

$$d(A, B) = \mu(A \Delta B).$$

Здесь μ — мера на рассматриваемом пространстве множеств, Δ — символ симметрической разности множеств,

$$A \Delta B = (A / B) \cup (B / A).$$

Если мера — так называемая считающая, т.е. приписывающая единичный вес каждому элементу множества, то введенное расстояние есть число несовпадающих элементов в множествах A и B.

Замечание. Строго говоря, функция $d(A, B) = \mu(A \Delta B)$ не задает метрику, поскольку из $d(A, B) = 0$ не всегда следует, что $A = B$, так как мера некоторых непустых множеств может равняться 0. Для функций $d(A, B)$, имеющих все свойства расстояний (метрик), кроме одного: из $d(A, B) = 0$ не всегда следует, что $A = B$, используют термин «псевдометрика».

Пример 11.10. Между множествами можно ввести и другое расстояние (*псевдометрику*):

$$d_1(A, B) = \frac{\mu(A \Delta B)}{\mu(A \cup B)}.$$

В ряде задач анализа экспертных данных используются функции двух переменных, для которых выполнены не все три аксиомы расстояния, а только некоторые. Их обычно называют показателями различия, поскольку чем больше различаются объекты, тем больше значение функции. Иногда в том же смысле используют термин «мера близости». Он менее удачен, поскольку большее значение функции соответствует меньшей близости.

Чаще всего отказываются от аксиомы, требующей выполнения неравенства треугольника, поскольку это требование не всегда находит обоснование в конкретной прикладной ситуации.

Пример 11.11. В конечномерном векторном пространстве показателем различия является функция

$$d(x, y) = \sum_{j=1}^{k} [x(j) - y(j)]^2$$

(сравните с примером 11.3).

Показателями различия, но не расстояниями являются такие популярные в прикладной статистике показатели, как дисперсия или средний квадрат ошибки при оценивании.

Иногда отказываются также и от аксиомы симметричности.

Пример 11.12. Показателем различия чисел x и y является функция

$$d(x, y) = \left| \frac{x}{y} - 1 \right|.$$

Такой показатель различия используют в ряде процедур экспертного оценивания [10].

В различных постановках задач анализа экспертных данных обычно принимают первую аксиому расстояния. Вполне естественно, что должен достигаться наименьший показатель различия, причем именно на совпадающих объектах. Имеет ли смысл это наименьшее значение делать отличным от 0? Вряд ли, поскольку всегда можно добавить одну и ту же константу ко всем значениям показателя различия и тем самым добиться выполнения первой аксиомы.

11.3. АКСИОМАТИЧЕСКОЕ ВВЕДЕНИЕ РАССТОЯНИЙ

При анализе экспертных данных используют большое количество метрик и показателей различия. Как обоснованно выбрать то или иное расстояние для использования в конкретной задаче? В 1959 г. американский математик Джон Кемени предложил использовать аксиоматический подход, согласно которому следует сформулировать естественные для конкретной задачи аксиомы и вывести из них вид метрики. Этот подход получил большую популярность в нашей стране после выхода в 1972 г. перевода на русский язык книги Дж. Кемени и Дж. Снелла [25], в которой дана система аксиом для расстояния Кемени между упорядочениями. Последовала большая серия работ, в которых из тех или иных систем аксиом выводился вид метрики или показателя различия для различных видов данных, прежде всего для объектов нечисловой природы. Многие полученные результаты описаны в обзоре [103], содержащем 161 ссылку на предыдущие публикации, в том числе 69 на русском языке. Рассмотрим некоторые задачи аксиоматического введения расстояний.

Аксиоматическое введение расстояния между толерантностями. ***Толерантность*** — это бинарное отношение, являющееся рефлексивным и симметричным. Его обычно используют для описания отношения сходства между реальными объектами, отношений знакомства или дружбы между людьми. От отношения эквивалентности толерантность отличается тем, что свойство транзитивности не предполагается обязательно выполненным. Действительно, Иванов может быть знаком с Петровым, Петров — с Сидоровым, но при этом ничего необычного нет в том, что Иванов и Сидоров не знакомы между собой.

Пусть множество X, на котором определено отношение толерантности, состоит из конечного числа элементов: $X = \{x_1, x_2, ..., x_k\}$. Тогда толерантность описывается квадратной матрицей $A = \|a(i,j)\|$, $i, j = 1, 2, ..., k$, такой, что $a(i, j) = 1$, если x_i и x_j связаны отношением толерантности, и $a(i, j) = 0$ в противном случае. Матрица A симметрична: $a(i, j) = a(j, i)$, на главной диагонали стоят единицы: $a(i, i) = 1$. Любая матрица, удовлетворяющая приведенным в предыдущей фразе условиям, является матрицей, соответствующей некоторому отношению толерантности. Матрице A можно сопоставить неориентированный граф с вершинами в точках X: вершины x_i и x_j соединены ребром тогда и только тогда, когда $a(i, j) = 1$. Толерантности часто используются при проведении экспертных исследований.

Будем говорить, что толерантность A_3 лежит между толерантностями A_1 и A_2, если при всех i, j число $a_3(i, j)$ лежит между числами $a_1(i, j)$ и $a_2(i, j)$, т.е. выполнены либо неравенства $a_1(i, j) \le a_3(i, j) \le a_2(i, j)$, либо неравенства $a_1(i, j) \ge a_3(i, j) \ge a_2(i, j)$.

Теорема 11.1 [88]. Пусть

(I) $d(A_1, A_2)$ — метрика в пространстве толерантностей, определенных на конечном множестве $X = \{x_1, x_2, ..., x_k\}$;

(II) $d(A_1, A_3) + d(A_3, A_2) = d(A_1, A_2)$ тогда и только тогда, когда A_3 лежит между A_1 и A_2;

(III) если отношения толерантности A_1 и A_2 отличаются только на одной паре элементов, т.е. $a_1(i, j) = a_2(i, j)$ при $(i, j) \ne (i_0, j_0)$, $i < j$, $i_0 < j_0$, и $a_1(i_0, j_0) \ne a_2(i_0, j_0)$, то $d(A_1, A_2) = 1$.

Тогда

$$d(A_1, A_2) = \sum_{1 \le i \le j \le k} |a_1(i, j) - a_2(i, j)| = \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} |a_1(i, j) - a_2(i, j)|.$$

Таким образом, расстояние $d(A_1, A_2)$ только постоянным множителем $^1/_2$ отличается от расстояния Кемени, введенного в подразделе 11.1 в пространстве всех бинарных отношений как расстояние Хемминга между описывающими отношения матрицами из 0 и 1. Теорема 1 дает аксиоматическое введение расстояния в пространстве толерантностей. Оказалось, что оно является сужением расстояния Кемени на это пространство. Сам Дж. Кемени дал аналогичную систему аксиом для сужения на пространство упорядочений. Доказательство теоремы 11.1 вытекает из рассмотрений, связанных с аксиоматическим введением расстояний между множествами, и приводится далее.

Мера симметрической разности как расстояние между множествами. Как известно, бинарное отношение можно рассматривать как подмножество декартова квадрата X^2 того множества X, на котором оно определено. Поэтому теорему 11.1 можно рассматривать как аксиоматическое введение расстояния между множествами специального вида. Укажем систему аксиом для расстояния между множествами общего вида, описанного в примере 11.9.

Определение 11.2. Множество B находится между множествами A и C, если $(A \cap C) \subseteq B \subseteq (A \cup C)$.

С помощью определения 11.2 в совокупности множеств вводятся геометрические соотношения, использование которых полезно для восприятия рассматриваемых ситуаций.

Расстояние между двумя точками в евклидовом пространстве не изменится, если обе точки сдвинуть на один и тот же вектор. Аналогичное свойство расстояния между множествами сформулируем в виде аксиомы 11.1. Оно соответствует аксиоме 3 Кемени и Снелла [25, с. 22] для расстояний между упорядочениями.

Аксиома 11.1. Если $A \cap C = B \cap C = \varnothing$, то $d(A, B) = d(A \cup C, B \cup C)$.

Определение 11.3. Непустая система множеств называется *кольцом*, если для любых двух входящих в нее множеств в эту систему входят их объединение, пересечение и разность. Множество X называется *единицей системы множеств*, если оно входит в эту систему, а все остальные множества системы являются подмножествами X. Кольцо множеств, содержащее единицу, называется *алгеброй множеств* [31, с. 38].

Теорема 11.2. Пусть W — алгебра множеств, d: $W^2 \to R^1$. Тогда аксиома 11.1 эквивалентна следующему условию: $d(A, B) = d(A\backslash B, B\backslash A)$ для любых $A, B \in W$.

Доказательство. Поскольку

$$(A\backslash B)\cap(A\cap B) = \varnothing,\ (B\backslash A)\notin\cap(A\cap B) = \varnothing,$$

то равенство $d(A, B) = d(A\backslash B, B\backslash A)$ следует из аксиомы 11.1. Обратное утверждение вытекает из того, что в условиях аксиомы 11.1

$$(A\cup C)\backslash(B\cup C) = A\backslash B,\ (B\cup C)\backslash(A\cup C) = B\backslash A.$$

Теорема 11.2 доказана.

С целью внести в алгебру множеств W отношение «находиться между», аналогичное используемому при аксиоматическом введении расстояний в пространствах бинарных отношений — см. условие (II) в теореме 11.1, примем следующую аксиому.

Аксиома 11.2. Если B лежит между A и C, то $d(A, B) + d(B, C) = d(A, C)$.

Определение 11.4. Неотрицательная функция μ, определенная на алгебре множеств W, называется *мерой*, если для любых двух непересекающихся множеств A и B из W справедливо соотношение

$$\mu(A\cup B) = \mu(A) + \mu(B).$$

Понятие меры — это обобщение понятий длины линии, площади фигуры, объема тела.

Теорема 11.3. Пусть W — алгебра множеств, аксиомы 1 и 2 выполнены для функции d: $W^2 \to [0; +\infty]$. Функция d симметрична: $d(A, B) = d(B, A)$ для любых A и B из W. Тогда существует, и притом единственная, мера μ на W такая, что

$$d(A, B) = \mu(A\Delta B) \qquad (11.1)$$

при всех A и B из W, где $A\Delta B$ — симметрическая разность множеств A и B, т.е. $A\Delta B = (A\backslash B)\cup(B\backslash A)$.

Доказательство. Положим

$$\mu(B) = d(\varnothing, B),\ B\in W. \qquad (11.2)$$

Покажем, что определенная формулой (5.5.2) функция множества μ является мерой. Неотрицательность μ следует из неотрицательности d. Остается доказать аддитивность, т.е. что из $A\cap B = \varnothing$ следует, что

$$\mu(A\cup B) = \mu(A) + \mu(B),\ A\in W,\ B\in W. \qquad (11.3)$$

Поскольку A всегда лежит между $\varnothing$ и $A\cup B$, то по аксиоме 11.2

$$\begin{aligned}\mu(A\cup B) &= d(\varnothing, A\cup B) = d(\varnothing, A) + d(A, A\cup B) = \\ &= \mu(A) + d(A, A\cup B).\end{aligned} \qquad (11.4)$$

Если $A\cap B = \varnothing$, то по аксиоме 1 $d(\varnothing, B) = d(A, A\cup B)$, откуда с учетом (11.4) и следует (11.3).

Докажем соотношение (11.1). Поскольку $A\backslash B$ и $B\backslash A$ имеют пустое пересечение, то, согласно определению 11.2 пустое множество $\varnothing$ лежит между $A\backslash B$ и $B\backslash A$. Поэтому по аксиоме 11.2

$$d(A\backslash B, B\backslash A) = d(A\backslash B, \varnothing) + d(\varnothing, B\backslash A).$$

Из симметричности и соотношения (11.2) следует, что

$$d(A\backslash B, \varnothing) = d(\varnothing, A\backslash B) = \mu(A\backslash B),$$

откуда $d(A\backslash B, B\backslash A) = \mu(A\backslash B) + \mu(B\backslash A)$. Из соотношения (11.3) следует, что $\mu(A\backslash B) + \mu(B\backslash A) = \mu(A\Delta B)$. С другой стороны, по аксиоме 11.2

$$d(A\backslash B, B\backslash A) = d[(A\backslash B)\cup(A\cap B), (B\backslash A)\cup(A\cap B)] = d(A, B).$$

Из трех последних равенств вытекает справедливость равенства (11.1).

Остается доказать единственность меры μ в соотношении (11.1). Поскольку $A\Delta B = B$ при $A = \varnothing$, то из формулы (11.1) следует формула (11.2), т.е. однозначность определения меры $\mu = \mu(d)$ по расстоянию d. Теорема 11.3 доказана.

Теорема 11.4 (обратная). Пусть μ — мера, определенная на алгебре множеств W. Тогда функция $d(A, B) = \mu(A\Delta B)$ является псевдометрикой, для нее выполнены аксиомы 11.1 и 11.2.

Доказательство. То, что функция $d(A, B)$ из равенства (11.1) задает псевдометрику, хорошо известно (см., например, [65, с. 79]). Доказательство аксиомы 11.2 содержится в [45, с. 181—183]. Аксиома 11.1 следует из того, что условия $A\cap C = B\cap C = \varnothing$ обеспечивают справедливость соотношений

$$\begin{aligned}(A\cup C)\Delta(B\cup C) &= [(A\cup C)\backslash(B\cup C)]\cup[(B\cup C)\backslash(A\cup C)] = \\ &= (A\backslash B)\cup(B\backslash A) = A\Delta B.\end{aligned}$$

Замечание. Полагая в аксиоме 11.2 $A = B = C$, получаем, что $d(A, A) + d(A, A) = d(A, A)$, т.е. $d(A, A) = 0$. Согласно теоремам 11.3 и 11.4 из условий теоремы 11.3 следует неравенство треугольника. Таким образом, в теореме 11.3 действительно приведена система аксиом, определяющая семейство псевдометрик в пространстве множеств.

Обсудим независимость (друг от друга) условий теоремы 11.3. Отбрасывание неотрицательности функции d приводит к тому, что слово «мера» в теоремах 11.3 и 11.4 необходимо заменить на «заряд» [31, с. 328]. Этот термин обозначает аддитивную функцию множеств, не обладающую свойством неотрицательности. Заряд можно представить как разность двух мер.

Функция $d_1(A, B) = \sqrt{\mu(A\Delta B)}$ является псевдометрикой, для нее выполнена аксиома 11.1, но не выполнена аксиома 11.2, следовательно, ее нельзя представить в виде (11.1).

Приведем пример системы множеств W и метрики в ней, для которых верна аксиома 11.2, но не верна аксиома 11.1, а потому эту метрику нельзя представить в виде (11.1). Пусть W состоит из множеств $\varnothing$, A, B, $A \cup B$, причем $A \cap B = \varnothing$, а расстояния таковы:

$$d(\varnothing, A) = d(\varnothing, B) = 1,\ d(A, A \cup B) = d(B, A \cup B) = d(A, B) = \\ = 2,\ d(\varnothing, A \cup B) = 3.$$

Если единица X алгебры множеств W конечна, т.е. $X = \{x_1, x_2, ..., x_k\}$, то расстояние, определенное формулой (11.1), принимает вид

$$d(A,B)\sum_{i=1}^{k}\mu_i\left|\chi_A(x_i)-\chi_B(x_i)\right|, \qquad (11.5)$$

где $\chi_A(x_i)$ — индикатор (индикаторная функция) множества $A(B)$, т.е. $\chi_A(x) = 1$, если $x \in A$, и $\chi_A(x) = 0$ — в противном случае. Аналогично — для χ_B. Как следует из теоремы 11.3, неотрицательный коэффициент μ_i — это мера одноэлементного множества $\{x_i\}$, а также расстояние этого множества от пустого множества, т.е.

$$\mu_i = \mu(\{x_i\}) = d(\varnothing, \{x_i\}).$$

Если все коэффициенты μ_i положительны, то формула (11.5) определяет метрику, если хотя бы один равен 0, то — псевдометрику, поскольку в таком случае найдутся два различающиеся между собой множества A и B такие, что $d(A, B) = 0$.

Расстояние определяется однозначно, если априори известны коэффициенты μ_i. В частности, равноправность объектов (элементов единицы алгебры множеств X) приводит к $\mu_i \equiv 1$. Требование равноправности содержится в аксиомах 2 и 4 Кемени [25, с. 21—22].

Применим полученные результаты к толерантностям и докажем теорему 11.1. Совокупность всех толерантностей, определенных на конечном множестве Y, естественным образом ассоциируется с совокупностью всех подмножеств множества $X = \{(y_i, y_j),\ 1 \le i < j \le k\}$. Именно пара (y_i, y_j) входит в подмножество тогда и только тогда, когда y_i и y_j связаны отношением толерантности. Указанная совокупность подмножеств является алгеброй множеств с единицей X. Определение 11.2 понятия «находиться между» для множеств полностью соответствует ранее данному определению понятия «находиться между» для толерантностей.

Теорема 11.5. Пусть выполнены условия (I) и (II) теоремы 11.1 и аксиома 11.1. Тогда существуют числа $\mu_{ij} > 0$ такие, что

$$d(A,B)=\sum_{1\le ii<j\le k}\mu_{ij}\left|a(i,j)-b(i,j)\right|. \qquad (11.6)$$

Для доказательства достаточно сослаться на теорему 11.3. Поскольку в условии (I) теоремы 11.1 требуется, чтобы функция $d(A, B)$ являлась метрикой, то необходимо $\mu_{ij} > 0$.

Теорема 11.6. Пусть выполнены условия теоремы 11.1 и, кроме того, аксиома 11.1. Тогда верно заключение теоремы 11.1.

Доказательство. Рассмотрим толерантность A, для которой $a(i, j) = 1$ при $(i, j) = (i_0, j_0)$ и $a(i, j) = 0$ в противном случае. Согласно условию (III) теоремы 11.1 $d(\varnothing, A) = 1$, а согласно формуле (11.6) имеем $d(\varnothing, A) = \mu_{i_0 j_0}$. Следовательно, коэффициент $\mu_{i_0 j_0} = 1$, что и требовалось доказать.

Для окончательного доказательства теоремы 11.1 осталось избавиться от требования справедливости аксиомы 11.1.

Доказательство теоремы 11.1. Рассмотрим две толерантности A и B такие, что при представлении их в виде множеств $A \subseteq B$, это означает, что $a(i, j) \le b(i, j)$ при всех i, j. Поскольку X — конечное множество, то существует конечная последовательность толерантностей A_1, A_2, ..., A_m, ..., A_t такая, что $A_1 = A$, $A_t = B$, $A_1 \subseteq A_2 \subseteq ... \subseteq A_m \subseteq ... \subseteq A_t$, причем A_{m+1} получается из A_m заменой ровно одного значения $a_m(i_m, j_m) = 0$ на $a_{m+1}(i_m, j_m) = 1$, для $(i, j) \ne (i_m, j_m)$, при этом $a_m(i, j) = a_{m+1}(i, j)$. Тогда A_m находится между A_{m-1} и A_{m+1}, следовательно, по условию (II)

$$d(A, B) = d(A_1, A_2) + d(A_2, A_3) + ... + d(A_m, A_{m+1}) + ... + d(A_{t-1}, A_t).$$

По условию (III) $d(A_m, A_{m+1}) = 1$ при всех m, а потому заключение теоремы 11.1 верно для любых A и B таких, что $A \subseteq B$.

Поскольку $A \cap B$ лежит между A и B, то по условию (II)

$$d(A, B) = d(A \cap B, A) + d(A \cap B, B).$$

При этом $A \cap B \subseteq A$ и $A \cap B \subseteq B$. Применяя результат предыдущей формулы, получаем, что заключение теоремы 11.1 верно всегда.

Замечания. 1. Условие (III) не только дает нормировку, но и заменяет аксиому 11.1.

2. Условие (I) теоремы 11.1 не использовалось в доказательстве, но было приведено в первоначальной публикации [88], чтобы подчеркнуть цель рассуждения. По той же причине оно сохранено в формулировке теоремы 11.1, хотя в доказательстве удалось без него обойтись. Понадобилась только симметричность функции d.

Аксиоматическое введение метрики в пространстве неотрицательных суммируемых функций. Рассмотрим пространство $L(E, \mu)$ неотрицательных суммируемых функций на множестве E с мерой μ. Далее до конца настоящего раздела будем рассматривать только функции из пространства $L(E, \mu)$. Интегрирование всюду проводится по множеству (пространству) E и по мере μ. Будем писать $g = h$ или $g \le h$, если указанные соотношения справедливы почти всюду по μ на E (т.е. могут нарушаться лишь на множестве нулевой меры).

Аксиоматически введем расстояние в пространстве $L(E, \mu)$ (изложение следует работе [77]). Обозначим $M(g, h) = \max(g, h)$

и $m(g, h) = \min(g, h)$. Пусть функция D: $L(E, \mu) \times L(E, \mu) \to R^1$ — тот основной объект изучения, аксиомы для которого будут сейчас сформулированы.

Аксиома 11.3. Если $gh = 0$, $g + h \neq 0$, то $D(g, h) = 1$.

Аксиома 11.4. Если $h \le g$, то $D(g, h) = C\int(g - h)d\mu$, где множитель C не зависит от h, т.е. $C = C(g)$.

Лемма. Из аксиом 11.3, 11.4 следует, что для $h \le g \neq 0$

$$D(g,h) = \frac{\int (g-h)d\mu}{\int g d\mu}.$$

Для доказательства заметим, что по аксиоме 11.3 $D(g, 0) = 1$, а по аксиоме 11.4 $D(g, 0) = C\int g d\mu$, откуда $C = (\int g d\mu)^{-1}$. Подставляя это соотношение в аксиому 11.4, получаем заключение леммы.

Требование согласованности расстояния в пространстве $L(E, \mu)$ с отношением «находиться между» приводит, как и ранее для расстояния $d(A, B)$, к следующей аксиоме.

Аксиома 11.5. Для любых g и h справедливо равенство $D(g, h) = D[M(g, h), g] + D[M(g, h), h]$.

Замечание. В ряде реальных ситуаций естественно считать, что наибольшее расстояние между элементами пространства множеств (которое без ограничения общности можно положить равным 1), т.е. наибольшее несходство, соответствует множествам, не имеющим общих элементов. Расстояние, введенное в теореме 11.3 (формула (11.1)), этому условию не удовлетворяет. Поэтому в пространстве множеств была аксиоматически введена [103] так называемая D-метрика (от англ. *dissimilarity* — несходство), для которого это условие выполнено. Она имеет вид:

$$D(A,B)\begin{cases} \dfrac{\mu(A\Delta B)}{\mu(A\cup B)}, \mu(A\cup B)>0, \\ 0, \mu(A)=\mu(B)=0. \end{cases} \tag{11.7}$$

Приведенные ранее аксиомы являются обобщениями соответствующих аксиом для D-метрики в пространстве множеств.

Теорема 11.7. Из аксиом 11.3—11.5 следует, что

$$D(g,h) = \begin{cases} \dfrac{\int |g-h| d\mu}{\int M(g,h)d\mu}, g+h\neq 0, \\ 0, g=h=0. \end{cases} \tag{11.8}$$

Доказательство. Поскольку

$$M(g, h) - g + M(g, h) - h = |g - h|,$$

то заключение теоремы 11.7 при $g + h \neq 0$ вытекает из леммы и аксиомы 11.5. Из аксиомы 11.4 при $g = 0$ следует, что $D(0, 0) = 0$. Легко видеть, что функция D, заданная формулой (11.8), удовлетворяет аксиомам 11.3—11.5 и, кроме того, $D(g, h) \le 1$ при любых g и h.

Замечание. Если g и h — индикаторные функции множеств, то формула (11.8) переходит в формулу (11.7). Если g и h — функции принадлежности нечетких множеств, то формула (8) задает метрику в пространстве нечетких множеств, а именно D-метрику в этом пространстве [103].

Теорема 11.8. Функция $D(g,h)$, определенная формулой (11.8), является метрикой в $L(E, \mu)$ (при отождествлении функций, отличающихся лишь на множестве нулевой меры), причем $D(g, f) + D(f, h) = D(g, h)$ тогда и только тогда, когда $f = g$, $f = h$ или $f = M(g, h)$.

Доказательство. Обратимся к определению метрики. Для рассматриваемой функции непосредственно очевидна справедливость условий неотрицательности и симметричности. Очевидна и эквивалентность условия $D(g, h) = 0$ равенству $g = h$. Остается доказать неравенство треугольника и установить, когда оно обращается в равенство.

Без ограничения общности можно считать, что рассматриваемые расстояния задаются верхней строкой формулы (11.8) и, кроме того,

$$R = \int M(g, f)\, d\mu - \int M(f, h)\, d\mu \ge 0,$$

— частные случаи с использованием нижней строки формулы (11.8) рассматриваются элементарно, а справедливости последнего неравенства можно добиться заменой обозначений функций — элементов пространства $L(E, \mu)$. Тогда

$$D(g,f) + D(f,h) \ge \frac{\int (|g-f|+|f-h|)d\mu}{\int M(g,f)d\mu}, \tag{11.9}$$

причем равенство имеет место тогда и только тогда, когда $R = 0$ или $f = h$. Положим

$$P = \int(|g - f| + |f - h| - |g - h|)\, d\mu,\ Q = \int(M(g, f) - M(g, h))d\mu.$$

Ясно, что $P \ge 0$ и

$$\frac{\int (|g-f|+|f-h|)d\mu}{\int M(9g,f)d\mu} = \frac{\int |g-h|d\mu + P}{\int M(g,h)d\mu + Q}. \tag{11.10}$$

Если $Q < 0$, то, очевидно, неравенство треугольника выполнено, причем неравенство является строгим. Рассмотрим случай $Q > 0$.

Воспользуемся следующим элементарным фактом: если $y \geq x$, $y > 0$, $P > Q > 0$, то

$$\frac{x+P}{y+Q} > \frac{x}{y}. \qquad (11.11)$$

Из соотношений (11.10) и (11.11) вытекает, что для доказательства неравенства треугольника достаточно показать, что $P - Q > 0$.

Рассмотрим

$$k = \{|g - f| + |f - h| - |g - h|\} - M(g, f) + M(g, h).$$

Применяя равенство $M(g, h) - g + M(g, h) - h = |g - h|$ к слагаемым, заключенным в фигурные скобки, получаем, что

$$k = M(f, h) + [M(g, f) + M(f, h) - M(g, h) - 2f].$$

Применяя соотношение

$$M(g, h) = g + h - m(g, h) \qquad (11.12)$$

к слагаемым, заключенным в квадратные скобки, получаем, что

$$k = M(f, h) - m(f, h) - m(g, f) + m(g, h).$$

Так как $M(f, h) - m(f, h) = |f - h|$, то

$$k = |f - h| - [m(g, f) - m(g, h)] \geq (f - h) - [m(g, f) - m(g, h)]. \quad (11.13)$$

В соответствии с формулой (11.12) правая часть формулы (11.13) есть $M(g, f) - M(g, h)$, а потому

$$P - Q = \int k d\mu \geq Q > 0,$$

что завершает доказательство для случая $Q > 0$. При этом неравенство треугольника является строгим.

Осталось рассмотреть случай $Q = 0$. В силу соотношений (11.9) и (11.10) неравенство треугольника выполнено. Когда оно обращается в равенство? Тривиальные случаи: $f = g$ или $f = h$. Если же f отлично от g и h, то необходимо, чтобы $R = 0$ и $P = 0$. Как легко проверить, последнее условие эквивалентно неравенствам

$$m(g, h) \leq f \leq M(g, h). \qquad (11.14)$$

Из правого неравенства в (11.14) следует, что $M(g, f) \leq M[g, M(g, h)] = M(g, h)$. Так как $Q = 0$, то $M(g, f) = M(g, h)$. Аналогичным образом из соотношений

$$M(h, f) \leq M[h, M(g, h)] = M(g, h) = M(g, f)$$

и $R = 0$ следует, что $M(f, h) = M(g, h)$.

Рассмотрим измеримое множество $X = \{x \in E: h(x) < g(x)\}$. Тогда $M(g, h)(x) = M(f, h)(x) = g(x) > h(x)$, т.е. $h(x) < f(x) = M(g, h)(x)$ для почти всех $x \in X$. Для почти всех $y \in \{x \in E: h(x) > g(x)\}$ точно так же получаем

$f(y) = M(g, h)(y)$. Для почти всех $z \in \{x \in E: h(x) = g(x)\}$ в силу формулы (11.14) $f(z) = M(g, h)(z)$, что и завершает доказательство теоремы.

Замечание. Назовем функции g и h подобными, если существует число $b > 0$ такое, что $g = bh$. Тогда при $0 < b \leq 1$ имеем $D(g, h) = 1 - b$, т.е. расстояние между подобными функциями линейно зависит от коэффициента подобия. Далее, пусть $a > 0$, тогда $D(ag, ah) = D(g, h)$. Таким образом, метрика (11.8) инвариантна по отношению к преобразованиям подобия, которые образуют группу допустимых преобразований в шкале отношений. Это дает основания именовать метрику (11.8) *метрикой подобия* [77].

11.4. СВОЙСТВА МЕДИАНЫ КЕМЕНИ

Иногда пытаются противопоставить дискретные и вероятностно-статистические методы анализа экспертных оценок. Исходят из того, что во втором случае используются те или иные вероятностно-статистические модели, а в первом — только детерминированные. Мы полагаем, что речь идет о двух разных этапах изучения ситуации. Начать естественно с детерминированного анализа конкретных экспертных данных, разработать методы расчетов и получения выводов (заключений о данных), а затем изучить свойства этих методов расчета и получения выводов, используя вероятностно-статистические модели. Если мы хотим перенести выводы с конкретной выборки на генеральную совокупность, нам не обойтись без вероятностно-статистических моделей (подробнее см. [77; 89]).

Компьютерное изучение свойств медианы Кемени при конечных объемах выборок. С помощью специально разработанной программной системы В.Н. Жихаревым был проведен ряд серий численных экспериментов по изучению свойств выборочных медиан Кемени. Представление о полученных результатах дает табл. 11.2 [77].

Таблица 11.2

Вычислительный эксперимент по изучению свойств медианы Кемени

Характеристика	Номер серии					
	1	2	3	4	5	6
Число испытаний	100	1 000	50	50	1 000	1 000
Число объектов	5	5	7	7	5	5
Количество экспертов	10	30	10	30	10	10
Частота непустого пересечения	0,85	0,58	0,52	0,2	0,786	0,911

Окончание

Характеристика	Номер серии					
	1	2	3	4	5	6
Среднее отношение диаметров	0,283	0,124	0,191	0,0892	0,202	0,0437
Средняя мощность медианы	5,04	2,41	6,4	2,88	3,51	1,35
Максимальная мощность медианы	30	14	19	11	40	12

В каждой из шести серий методом статистических испытаний определенное число раз моделировался случайный и независимый выбор экспертных ранжировок, а затем находились все медианы Кемени для смоделированного набора мнений экспертов. При этом в сериях 1—5 распределение ответа эксперта предполагалось равномерным на множестве всех ранжировок.

В серии 6 это распределение являлось монотонным относительно расстояния Кемени с некоторым центром, т.е. вероятность выбора определенной ранжировки убывала с увеличением расстояния Кемени этой ранжировки от центра.

Определение 11.5. Распределение бинарного отношения X называется монотонным с центром в C_0 относительно расстояния (показателя различия) d, если из $d(C, C_0) < d(D, C_0)$ следует, что $P(X = C) > P(X = D)$.

Это определение впервые дано в монографии [88, с. 196]. Оно может использоваться в любых пространствах бинарных отношений и, более того, в любых пространствах из конечного числа элементов, лишь бы в них была введена функция $d(C, D)$ — показатель различия элементов C и D этого пространства. Монотонное распределение унимодально, мода находится в C_0.

Таким образом, серии 1—5 соответствуют ситуации, когда у экспертов нет почвы для согласия, нет группировки их мнений относительно некоторого единого среднего группового мнения, в то время как в серии 6 есть единое мнение — описанный ранее центр, к которому тяготеют ответы экспертов.

Обсуждение результатов. Результаты, приведенные в табл. 11.2, можно комментировать разными способами. Неожиданным явилось большое число элементов в выборочной медиане Кемени — как среднее, так и максимальное. Одновременно обращает на себя внимание убывание этих чисел при росте числа экспертов и особенно при переходе к ситуации реального существования группового мнения (серия 6).

Достаточно часто один из ответов экспертов входит в медиану Кемени (т.е. пересечение множества ответов экспертов и медианы Кемени является непустым множеством), а диаметр медианы как множества в пространстве ранжировок заметно меньше диаметра множества ответов экспертов. По этим показателям наилучшее положение в серии 5. Грубо говоря, всяческие «патологии» в поведении медианы Кемени наиболее резко проявляются в ситуации, когда ее применение не имеет содержательного обоснования, т.е. когда у экспертов нет основы для согласия, их ответы равномерно распределены на множестве ранжировок.

Увеличение числа испытаний в 10 раз серии 5 по сравнению с серией 1 не очень сильно повлияло на приведенные в табл. 11.2 характеристики, поэтому представляется, что суть дела выявляется при числе испытаний (в методе Монте-Карло), равном 100 или даже 50. Увеличение числа объектов или экспертов увеличивает число элементов в рассматриваемом пространстве ранжировок, а потому уменьшается частота попадания какого-либо из мнений экспертов внутрь медианы Кемени. Также уменьшается отношение диаметра медианы к диаметру множества экспертов и число элементов медианы Кемени (среднее и максимальное). Можно сказать, что увеличение числа объектов или экспертов уменьшает степень дискретности задачи, приближает ее к непрерывному случаю, а потому уменьшает выраженность различных «патологий».

Есть много интересных результатов, которые здесь не рассмотрены. Они связаны, в частности, со сравнением медианы Кемени с другими методами усреднения мнений экспертов, например с нахождением итогового упорядочения по методу средних рангов, а также с использованием малых окрестностей ответов экспертов для поиска входящих в медиану ранжировок, с теоретической и численной оценкой скорости сходимости в законах больших чисел.

11.5. КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ И КОНКОРДАЦИИ

Термин «корреляция» означает «связь». Применительно к анализу данных этот термин обычно используется в сочетании «коэффициенты корреляции». В подразделе 5.3 рассмотрены линейный коэффициент парной корреляции К. Пирсона и непараметрический ранговый коэффициент парной корреляции Спирмена (часто употребляют краткие формы терминов — коэффициент корреляции Пирсона, коэффициент ранговой корреляции Спирмена, а также иные варианты сокращений полных наименований).

Широко используется также коэффициент ранговой корреляции τ Кендалла, коэффициент ранговой конкордации Кендалла и Б. Смита и др. Наиболее подробное обсуждение этой тематики содержится в монографии [28]. Дискуссии о выборе вида коэффициентов корреляции продолжаются до настоящего времени.

Коэффициент ранговой корреляции τ Кендалла определяется так [28]. Пусть N — количество тех упорядоченных пар индексов (i, j), $i < j$, для которых эксперты одинаково упорядочивают объекты, т.е. для которых либо одновременно $r_i < r_j$, $q_i < q_j$, либо одновременно $r_i > r_j$, $q_i > q_j$. Тогда

$$\tau = \frac{4N}{n(n-1)} - 1.$$

Если экспертные упорядочения совпадают, то коэффициент ранговой корреляции Кендалла принимает максимальное значение $\tau = 1$. Именно так обстоит дело для данных, приведенных в табл. 11.3. Если эксперты дают прямо противоположные упорядочения, их мнения противоречат друг другу для любой пары объектов, то коэффициент ранговой корреляции Кендалла минимален, $\tau = -1$.

Если число экспертов $m > 2$, то данные ими m упорядочений можно записать в виде матрицы, i-я строка которой содержит ранжировку, полученную от i-го эксперта, а столбцы соответствуют n объектам экспертизы, рассматриваемым в данном исследовании:

$$\begin{Vmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,n} \\ r_{2,1} & r_{2,2} & \dots & r_{2,n} \\ \dots & \dots & \dots & \dots \\ r_{m,1} & r_{m,2} & \dots & r_{m,n} \end{Vmatrix}. \tag{11.15}$$

Иногда используется более общая терминология. Вместо ранжировки, полученной от i-го эксперта, рассматривается ранжировка по i-му признаку.

В качестве единой выборочной меры связи m признаков Кендалл и Б. Смит предложили коэффициент согласованности W, называемый также *коэффициентом конкордации* (от лат. *concordare* — привести в соответствие, упорядочить):

$$W = \frac{12 S_W}{m^2 (n^3 - n)},$$

где

$$S_W = \sum_{i=1}^{n} \left[\sum_{j=1}^{m} r_{i,j} - \frac{m(n+1)}{2} \right]^2.$$

Можно показать, что среднее арифметическое коэффициентов ранговой корреляции Спирмена ρ для $m(m-1)/2$ пар признаков равно $(mW-1)/(m-1)$. В частности, если $m = 2$, то $\rho = 2W - 1$.

Все три коэффициента — $|\rho|$, $|\tau|$ и W — принимают значения из отрезка $[0; 1]$ и используются для проверки нулевой гипотезы H_0 о независимости признаков. Признаки называются независимыми, если для наугад выбранного столбца матрицы (11.15) ранги (порядковые номера) $r_{1,j}$, $r_{2,j}$, ..., $r_{m,j}$ являются взаимно независимыми случайными величинами. В терминах теории экспертных оценок гипотеза H_0 — это гипотеза о том, что случайные ранжировки независимы и равномерно распределены на множестве всех ранжировок (без связей).

Если рассматриваемый коэффициент ($|\rho|$, $|\tau|$ или W) не превосходит заданного граничного значения, то гипотеза H_0 принимается, если превосходит — отклоняется в пользу альтернативной гипотезы общего вида, т.е. гипотезы о том, что совместное распределение ранжировок отличается от совместного распределения независимых одинаково распределенных ранжировок. При этом остается неизвестным, нарушается ли предположение независимости, или предположение равномерности распределения, или и то и другое вместе. Например, нулевая гипотеза отклоняется, если все эксперты повторяют ответ первого из них, но сам этот ответ равномерно распределен. Или тогда, когда половина экспертов выбирает одну определенную ранжировку или похожие на нее, а вторая половина экспертов — другую определенную ранжировку (или похожую на нее). В этом случае нет равномерности распределения и нулевая гипотеза отклоняется, хотя говорить о согласованности экспертов не приходится. Если же нулевая гипотеза принимается, то ни о какой согласованности мнений экспертов говорить нельзя.

Распределения коэффициентов ($|\rho|$, $|\tau|$ или W) — дискретные, граничные значения зависят от числа объектов экспертизы n, числа экспертов m и уровня значимости α. Распределения коэффициентов ранговой корреляции $|\rho|$ и $|\tau|$ и коэффициента согласованности (конкордации) W приведены в [28].

Если гипотеза H_0 верна, то

$$M(\rho) = 0,\ M(\tau) = 0,\ M(W) = \frac{1}{m},$$

$$D(\rho) = \frac{1}{n-1},\ D(\tau) = \frac{2(2n+5)}{0n(n-1)},\ D(W) = \frac{2(m-1)}{m^3(n-1)}.$$

Распределения коэффициентов ранговой корреляции ρ и τ и коэффициента согласованности (конкордации) W являются асимптотически нормальными, причем с приведенными ранее значениями математических ожиданий и дисперсий. Асимптотической нормальностью распределений коэффициентов ранговой корреляции ρ и τ можно пользоваться для вычисления их критических значений при $n > 10$. В то же время коэффициент согласованности (конкордации) W распределен асимметрично, для него сходимость распределения к нормальному закону медленнее, чем для коэффициентов ранговой корреляции ρ и τ, и в рекомендуется использовать аппроксимацию бета-распределением (β-распределением).

Подробнее о ранговой корреляции и ее применениях, о мощности критериев некоррелированности признаков, о предельных теоремах можно узнать в [28]. Полезная информация собрана в [106], хотя эта статья и содержит некоторые неаккуратные (с математической точки зрения) формулировки.

Пример 11.13. Необходимо определить степень согласованности мнения пяти экспертов ($m = 5$), результаты ранжирования которыми семи объектов ($n = 7$) приведены в табл. 11.3.

Таблица 11.3

Данные для оценки согласованности мнений пяти экспертов

Номер объекта экспертизы	Оценка эксперта					Сумма рангов	Отклонение от среднего	Квадрат отклонения
	1	2	3	4	5			
1	4	6	4	4	3	21	1	1
2	3	3	2	3	4	15	–5	25
3	2	2	1	2	2	9	11	121
4	6	5	6	5	6	28	8	64
5	1	1	3	1	1	7	–13	169
6	5	4	5	6	5	25	5	25
7	7	7	7	7	7	35	15	225

Рассчитаем среднее арифметическое рангов:

$$\frac{m(n+1)}{2} = \frac{5(7+1)}{2} = 20.$$

Затем рассчитаем сумму квадратов отклонений сумм рангов по объектам экспертизы от их среднего арифметического:

$$S_W = \sum_{i=1}^{7}\left[\sum_{j=1}^{5} r_{i,j} - 20\right]^2 = 630.$$

Определим величину коэффициента конкордации:

$$W = \frac{12 \times 630}{5^2(7^3 - 7)} = 0{,}9.$$

Много это или мало? Если проведем соответствующие вычисления с помощью программного продукта *Statistica*, то получим значение достигаемого уровня значимости 0,00014. Это значит, что нулевая гипотеза отклоняется на любом из реально используемых в социально-экономических и технических исследованиях уровней значимости (т.е. 0,05, 0,01 или 0,1), поскольку все они много больше достигаемого уровня значимости.

Напомним, что достигаемый уровень значимости — это случайная величина, равная вероятности попадания статистики критерия в критическую область, заданную рассчитанным по выборке значением статистики критерия. Для критической области вида $\{x: x > a\}$ достигаемый уровень значимости есть $F(X_n)$, где X_n — рассчитанное по выборке значение статистики критерия X, а $F(a) = P(X > a)$ — дополнение до единицы функции распределения статистики критерия X. Достигаемый уровень значимости — это вероятность того, что статистика критерия X в новом независимом эксперименте примет значение большее, чем при расчете по конкретной выборке, т.е. большее, чем X_n [89, приложение 1].

Нормированная и центрированная величина коэффициента конкордации W такова:

$$\frac{W - M(W)}{\sqrt{D(W)}} = \frac{W - \frac{1}{m}}{\sqrt{\frac{2(m-1)}{m^3(n-1)}}} = \frac{W - \frac{1}{5}}{\sqrt{\frac{2 \times 4}{125 \times 6}}} = \frac{W - 0{,}2}{0{,}1033} = 6{,}78.$$

Из асимптотической нормальности W вытекает тот же вывод, что и из расчетов с помощью пакета *Statistica*.

Расстояние Кемени и коэффициенты ранговой корреляции. Пусть A и B — две ранжировки (без связей). Рассмотрим относительное расстояние Кемени между ранжировками, т.е.

$$d(A,B) = \frac{D(A,B)}{\max\limits_{A,B} D(A,B)} = \frac{2\sum\limits_{1 \le i < j \le k} |a(i,j) - b(i,j)|}{k(k-1)}.$$

Относительное расстояние неотрицательно и не превосходит 1. Оно равно 1 только для пар противоположных упорядочений, для которых различны все элементы описывающих их матриц, кроме лежащих на главной диагонали.

Пусть $\tau(A, B)$ — коэффициент ранговой корреляции Кендалла между ранжировками A и B. Тогда

$$2d(A, B) + \tau(A, B) = 1.$$

Более того, единственная с точностью до постоянного множителя линейная функция от $\tau(A, B)$, задающая расстояние между ранжировками A и B, есть

$$d(A,B)=\frac{1-\tau(A,B)}{2}.$$

При этом никакая линейная функция от коэффициента ранговой корреляции Спирмена $\rho(A, B)$ не задает расстояние между ранжировками.

Сформулированные здесь результаты получены в работе [88]. Они позволяют установить связь между двумя, казалось бы, совсем различными подходами к анализу экспертных мнений, выраженных ранжировками.

Контрольные вопросы

1. В чем состоит проблема согласованности ответов экспертов?
2. Как бинарные отношения используются в экспертизах?
3. Как бинарные отношения описываются матрицами из 0 и 1?
4. Что такое расстояние Кемени и медиана Кемени?
5. Чем закон больших чисел для медианы Кемени отличается от «классического» закона больших чисел, известного в статистике?
6. Выпишите матрицу из 0 и 1, соответствующую бинарному отношению (кластеризованной ранжировке) $5 < \{1, 3\} < 4 < 2 < \{6, 7\}$.

Темы докладов и рефератов

1. Классификация мнений экспертов и проверка согласованности.
2. Формирование итогового мнения комиссии экспертов.
3. Расстояние по Кемени и медиана Кемени в экспертных оценках.
4. Законы больших чисел в пространствах нечисловой природы.
5. Методы теории люсианов в теории и практике экспертных оценок.
6. Центральная роль статистики объектов произвольной природы в математической теории анализа экспертных оценок.
7. Расстояния в пространствах функций.

ГЛАВА 12
БИНАРНЫЕ ДАННЫЕ И ПАРНЫЕ СРАВНЕНИЯ

12.1. ТЕОРЕТИЧЕСКОЕ ОБОСНОВАНИЕ «ТУРНИРНОГО» МЕТОДА РАНЖИРОВАНИЯ ВАРИАНТОВ

Парное сравнение — это сравнение двух объектов экспертизы, когда эксперт выбирает из них лучший. В таблице 12.1 приведены результаты попарных сравнений шести объектов одним экспертом. Результат сравнения i-го и j-го объектов кодируется символом 1, если i-й объект лучше j-го, и символом 0 — в противном случае.

Таблица 12.1

Ранжирование шести объектов путем попарного сравнения

Номер объекта	1	2	3	4	5	6	Итог
1	×	1	0	1	1	1	4
2	0	×	0	1	1	1	4
3	1	1	×	1	1	1	5
4	0	0	0	×	0	0	0
5	0	0	0	1	×	0	1
6	0	0	0	1	1	×	2

Бинарные данные — данные, которые могут принимать два значения.

На основе парных сравнений можно решить многие задачи анализа экспертных данных. Например, можно упорядочить объекты по рассматриваемому признаку. Для этого достаточно, например, подсчитать, сколько раз определенный объект доминирует над другими, т.е. рассмотреть число единиц в строке. Эти величины приведены в последнем столбце табл. 12.1. Затем упорядочиваем объекты по указанным значениям. Получаем кластеризованную ранжировку

$$4 < 5 < 6 < \{1, 2\} < 3,$$

отражающую мнение эксперта. Итак, самым хорошим является объект 3, а самым плохим — объект 4.

В главе 13 рассмотрены различные методы анализа результатов парных сравнений и иных видов бинарных экспертных данных, т.е. данных, принимающих одно из двух значений: 0 или 1.

В статье [126] предложен оригинальный метод ранжирования вариантов, названный «турнирным». Он напоминает метод построения кластеризованной ранжировки на основе данных табл. 12.1 и подробно описан далее. Но сразу можно сформулировать естественные вопросы, на которые необходимо получить ответы.

Каковы статистические свойства этого метода? Позволяет ли он выявить истинное упорядочение вариантов? Другими словами, является ли состоятельной оценка упорядочения (ранжировки) вариантов, рассчитываемая с помощью «игрового» метода?

Для ответа на эти вопросы необходимо изучить свойства расчетной процедуры анализа данных. Как известно [89], такое изучение, как правило, состоит из двух этапов:

1) построения вероятностно-статистической модели порождения данных;
2) математико-статистическое изучения свойств расчетной процедуры анализа данных.

Пусть рассматривается k вариантов технического решения. В соответствии с описанием процедуры в статье [126] будем считать, что влияние i-го варианта, $i = 1, 2, ..., k$, на изучаемый параметр описывается (числовой) случайной величиной X_i с функцией распределения $F_i(x)$. Таким образом, сравнение двух вариантов — это сравнение функций распределения. Такое сравнение можно проводить разными способами — по тем или иным характеристикам (математическим ожиданиям, медианам, дисперсиям, квантилям порядка 0,999999, коэффициентам вариации и др.) или непосредственно с целью обнаружения различия между функциями распределения. Выбор того или иного вероятностно-статистического способа сравнения зависит от решаемой задачи. На примере оценки рисков (аварий, загрязнения окружающей среды, дефектности и др.) в [89] продемонстрирован подобный выбор.

Согласно [126] сравнивать надо математические ожидания. Лучше тот вариант, у которого математическое ожидание больше. Тогда результаты сравнения k вариантов технического решения описываются кластеризованной ранжировкой. Другими словами, варианты разбиты на группы. В каждой группе математические ожидания совпадают, между группами — различаются. Группы упорядочены в порядке возрастания математических ожиданий. Теоретическую кластеризованную ранжировку, соответствующую математическим ожиданиям, необходимо оценить по эмпирическим данным.

Поскольку функции распределения и их математические ожидания при сравнении конкретных вариантов технического решения неизвестны, то сравнения приходится проводить на основе выборок.

Принимаем, что влияние i-го варианта, $i = 1, 2, ..., k$, на изучаемый параметр оценивается с помощью выборки объема n_i, т.е. набора реализаций n_i независимых случайных величин с общей функцией распределения $F_i(x)$. Выборки предполагаются независимыми. Могут использоваться как экспертные оценки, так и объективные результаты измерения. Итак, вероятностно-статистическая модель порождения данных описана.

В соответствии с «турнирным» методом ранжирования сравнение двух вариантов состоит в статистической проверке нулевой гипотезы о равенстве соответствующих математических ожиданий. Если нулевая гипотеза принимается, то каждому варианту присваивается по 0,5 очка. Если нулевая гипотеза отклоняется, то варианту с бóльшим выборочным средним арифметическим присваивается 1 очко, а с меньшим — 0 очков. Проводятся все $k(k-1)/2$ парных сравнений, полученные очки суммируются, варианты упорядочиваются в порядке возрастания набранных сумм. Получаем эмпирическую кластеризованную ранжировку.

В соответствии с рекомендациями [77, 89] для проверки равенства математических ожиданий в работе [126] применяется критерий Крамера — Уэлча. Граничное значение для модуля статистики принято равным 1,645, что соответствует уровню значимости 0,1 (точнее, асимптотическому уровню значимости при безграничном росте объемов выборок). При решении задачи выбора конструкции коллектора для трибоэлектрического генератора в [126] получена следующая эмпирическая кластеризованная ранжировка типов коллекторов:

{игольчатый} < {кисточкообразный; ленточный с изгибом} < {штыковой} < {пилообразный; Г-образный}.

Качество вариантов убывает при движении справа налево. Самыми лучшими являются такие варианты, как пилообразный и Г-образный (причем по данным [126], эти варианты надо считать эквивалентными, они образуют кластер). Хуже по качеству штыковой коллектор, и т.д.

Эмпирическая кластеризованная ранжировка используется как оценка теоретической. Каковы математико-статистические свойства этой оценки? Поскольку кластеризованная ранжировка — это объект нечисловой природы, то для изучения свойств процедуры, предложенной в статье [126], необходимо применить подходы и результаты статистики объектов нечисловой природы [77; 89].

Теорема 12.1. При безграничном росте объемов выборок (т.е. при $\min\{n_i, i = 1, 2, ..., k\} \to \infty$) и фиксированном числе k вариантов вероятность того, что эмпирическая кластеризованная ранжировка совпадает с теоретической, стремится к 1.

В соответствии с теоремой 12.1 предложенная в работе [126] оценка теоретической кластеризованной ранжировки является состоятельной. Доказательство теоремы 12.1 проводится методами, разработанными в главе 8 монографии [89].

Как измерить степень близости эмпирической и теоретической кластеризованных ранжировок? В соответствии с известным в статистике объектов нечисловой природы аксиоматическим подходом целесообразно использовать расстояние Кемени или, что эквивалентно, коэффициент ранговой корреляции Кендалла (см. главу 11, подраздел 11.5). Справедливы следующие теоремы.

Теорема 12.2. При справедливости условий теоремы 12.1 расстояние Кемени между эмпирической кластеризованной ранжировкой и теоретической стремится к 0.

Теорема 12.3. При справедливости условий теоремы 12.1 коэффициент ранговой корреляции Кендалла между эмпирической кластеризованной ранжировкой и теоретической стремится к 1.

Доказательства теорем 12.2 и 12.3 проводятся методами, развитыми в [77, 89]. Таким образом, с точки зрения асимптотической математической статистики предложенный в работе [126] «турнирный» метод ранжирования вариантов получил обоснование. Что же касается конечных объемов выборок, особенно столь малых, как в [126], где все $n_i = 3$, то необходимы дальнейшие исследования, прежде всего методом статистических испытаний. Различные методы оценки близости допредельных и предельных распределений статистик проанализированы в [89, глава 10.3]. Приходится констатировать, что простые рекомендации отсутствуют.

12.2. ТЕОРИЯ СЛУЧАЙНЫХ ТОЛЕРАНТНОСТЕЙ

Толерантность как вид бинарных отношений. В прикладных исследованиях обычно используют три конкретных вида бинарных отношений — ранжировки, разбиения и толерантности. Статистические теории ранжировок [28] и разбиений [76] достаточно сложны с математической точки зрения. Поэтому продвинуться удается не очень далеко. Теория случайных ранжировок, в частности, изучает в основном равномерные распределения на множестве ранжировок. Теория случайных толерантностей позволяет рассмотреть принципиально более общие ситуации. Это объясняется, грубо говоря, тем, что для теории толерантностей оказываются полезными суммы некоторых независимых случайных величин, а для теории ранжировок и разбиений аналогичные случайные величины зависимы, а потому изучение их сумм затруднено.

Теория случайных толерантностей является частным случаем теории люсианов, рассматриваемой в подразделе 12.4. Здесь приводим результаты, специфичные именно для толерантностей.

Пусть X — конечное множество из k элементов. Толерантность A на множестве X, как и любое бинарное отношение, однозначно описывается матрицей $\|a(i, j)\|$, $1 \le i, j \le k$, где $a(i, j) = 1$, если элементы с номерами i и j связаны отношением толерантности, и $a(i, j) = 0$ — в противном случае. Поскольку толерантность — это рефлексивное и симметричное бинарное отношение, то достаточно рассматривать часть матрицы, лежащую над главной диагональю: $\|a(i, j), 1 \le i < j \le k\|$. Между наборами $\|a(i, j), 1 \le i < j \le k\|$ из 0 и 1 и толерантностями на X имеется взаимнооднозначное соответствие.

Пусть $A = A(\omega)$ — случайная толерантность, равномерно распределенная на множестве всех толерантностей на X. Легко видеть, что в этом случае $a(i, j)$, $1 \le i < j \le k$ — независимые случайные величины, принимающие значения 0 и 1 с вероятностями 0,5. Этот факт, несмотря на свою математическую тривиальность, является решающим для построения базовой части теории толерантностей. Для аналогичных постановок в теории ранжировок и разбиений величины $a(i, j)$ оказываются зависимыми.

Следовательно, случайная величина

$$B(A) = \sum_{i=1}^{k} \sum_{j=1}^{k} a(i, j)$$

имеет биномиальное распределение с параметрами $k(k-1)/2$, S и асимптотически нормальна при $k \to \infty$.

Проверка гипотез о согласованности. Рассмотрим s независимых толерантностей $A_1, A_2, ..., A_s$, равномерно распределенных на множестве всех толерантностей на X. Рассмотрим вектор

$$\xi_{ks} = \left\{ d(A_p, A_q),\ 1 \le p < q \le s \right\} =$$
$$= \sum_{1 \le p < q \le k} \left\{ \left| a_p(i, j) - a_q(i, j) \right|,\ 1 \le p < q \le s \right\}, \qquad (12.1)$$

где $d(A_p, A_q)$ — расстояние между толерантностями A_p и A_q, аксиоматически введенное в главу 11. В подразделе 12.1 предполагается, что пары (p, q), $p < q$, располагаются в раз и навсегда установленном порядке, для определенности — в лексиграфическом (т.е. пары упорядочиваются в соответствии со значением p, а при одинаковых p — по значению q).

Вектор ξ_{ks} является суммой $k(k-1)/2$ независимых одинаково распределенных случайных векторов, а потому асимптотически нормален при $k \to \infty$. Координаты этого вектора независимы, поскольку, как нетрудно видеть, координаты каждого слагаемого независимы (это свойство не сохраняется при отклонении от равномерности распределения). Распределения случайных величин $a_p(i,j)$ и $|a_p(i,j) - a_q(i,j)|$ совпадают, поэтому распределения $B(A)$ и $d(A_p, A_q)$ также совпадают.

В силу многомерной центральной предельной теоремы распределение вектора

$$\eta_{ks} = \sqrt{\frac{2}{k(k-1)}}\left[\xi_{rs} - \frac{k(k-1)}{2}\left(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}\right)\right]$$

сходится при $k \to \infty$ к распределению многомерного нормального вектора η_s, ковариационная матрица которого совпадает с ковариационной матрицей вектора η_{ks}, а математическое ожидание равно 0. Таким образом, координаты случайного вектора η_s независимы и имеют стандартное нормальное распределение с математическим ожиданием 0 и дисперсией 1. В соответствии с теоремами о наследовании сходимости [77, подраздел 4.3] распределение $f(\eta_{ks})$ сходится при $k \to \infty$ к распределению $f(\eta_s)$ для достаточно широкого класса функций f, в частности для всех непрерывных функций. В качестве примеров рассмотрим статистики:

$$W = \sum_{1 \le p < q \le s} d(A_p, A_q); \quad N = \sum_{1 \le p < q \le s} \left(d(A_p, A_q - \frac{k(k-1)_2}{4})\right)^2.$$

При $k \to \infty$ распределения случайных величин

$$\frac{8W - s(s-1)k(k-1)}{2\sqrt{s(s-1)k(k-1)}}, \quad \frac{8N}{k(k-1)}$$

сходятся соответственно к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1 и распределению хи-квадрат с $s(s-1)/2$ степенями свободы. Статистики W и N могут быть использованы для проверки гипотезы о равномерности распределения толерантностей.

Как известно, в теории ранговой корреляции, т.е. в теории случайных ранжировок, в качестве единой выборочной меры связи нескольких признаков используется *коэффициент согласованности* $W = W(R)$, называемый также *коэффициентом конкордации*. Его распределение затабулировано в предположении равномерности распределения на пространстве ранжировок (без связей). Непосредственным аналогом коэффициента конкордации $W(R)$ в случае толерантностей является только что введенная статистика W. Статистики W и N играют ту же роль для толерантностей, что $W(R)$ для ранжировок, однако математико-статистическая теория в случае толерантностей гораздо проще, чем для ранжировок.

Обобщением равномерно распределенных толерантностей являются толерантности с независимыми связями. В этой постановке предполагается, что $a(i,j)$, $1 \le i < j \le k$ — независимые случайные величины, принимающие значения 0 и 1. Обозначим $P[a(i,j) = 1] = p(i,j)$. Тогда $P[a(i,j) = 0] = 1 - p(i,j)$. Таким образом, распределение толерантности с независимыми связями задается нечеткой толерантностью, т.е. вектором

$$P = \{p(i,j), 1 \le i < j \le k\}.$$

Нечеткая толерантность — частный случай нечеткого множества. В свою очередь, нечеткие множества — один из видов объектов нечисловой природы, рассматриваемых в статистике нечисловых данных [77, 89].

Пусть имеется s независимых случайных толерантностей A_1, A_2, ..., A_s с независимыми связями, распределения которых задаются векторами P_1, P_2, ..., P_s соответственно. Рассмотрим проверку гипотезы согласованности

$$H_0 : P_1 = P_2 = \ldots = P_s.$$

Она является более слабой, чем гипотеза равномерности

$$H'_0 : P_1 = P_2 = \ldots = P_s = (^1/_2, {}^1/_2, \ldots, {}^1/_2),$$

для проверки которой используют статистики W и N (см. ранее).

Пусть сначала $s = 2$. Тогда

$$P\{|a_1(i,j) - a_2(i,j)| = 1\} = q(i,j), P\{|a_1(i,j) - a_2(i,j)| = 0\} = 1 - q(i,j),$$

где

$$q(i,j) = p_1(i,j)[1 - p_2(i,j)] + p_2(i,j)[1 - p_1(i,j)].$$

Следовательно, расстояние $d(A_1, A_2)$ между двумя случайными толерантностями с независимыми связями есть сумма $k(k-1)/2$ независимых случайных величин, принимающих значения 0 и 1, причем математическое ожидание и дисперсия $d(A_1, A_2)$ таковы:

$$Md(A_1, A_2) = \sum_{1 \le i < j \le k} q(i,j), \; Dd(A_1, A_2) = \sum_{1 \le i < j \le k} q(i,j)(1 - q(i,j)). \quad (12.2)$$

Пусть $k \to \infty$. Если $Dd(A_1, A_2) \to \infty$, то условие Линденберга Центральной предельной теоремы теории вероятностей выполнено и распределение нормированного расстояния

$$\frac{d(A_1, A_2) - M[d(A_1, A_2)]}{\sqrt{Dd(A_1, A_2)}} \quad (12.3)$$

сходится к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Если существует число $\delta > 0$ такое, что при всех k, i, j, $1 \le i < j \le k$, вероятности $p_1(i, j)$ и $p_2(i, j)$ лежат внутри интервала $(\delta; 1 - \delta)$, то $Dd(A_1, A_2) \to \infty$.

Соотношения (12.2), (12.3) и им подобные позволяют рассчитать мощность критериев, основанных на статистиках W и N, при $k \to \infty$, подобно тому, как это сделано в [88, подраздел 4.5]. Поскольку подобные расчеты не требуют новых идей, не будем приводить их здесь.

Обычно P_1 и P_2 неизвестны. Для проверки гипотезы $P_1 = P_2$ в некоторых случаях можно порекомендовать отвергать гипотезу на уровне значимости α, если $d(A_1, A_2) \ge d_0$, где d_0 есть $(1 - \alpha)$ — квантиль распределения расстояния между двумя независимыми равномерно распределенными случайными толерантностями, т.е. квантиль биномиального распределения $B(A)$. Укажем достаточные условия такой рекомендации.

Пусть

$$p = [p_1(i, j) + p_2(i, j)]/2,\ p_1(i, j) = p + \Delta,$$

тогда

$$p_2(i, j) = p - \Delta,\ q = q(i, j) = 2p(1 - p) + 2\Delta^2. \tag{12.4}$$

Если существует число $\delta > 0$ такое, что

$$q - {}^1/_2 > \delta > 0 \tag{12.5}$$

при всех k, i, j, то гипотеза $P_1 = P_2$ будет отвергаться с вероятностью, стремящейся к 1 при $k \to \infty$. Из формулы (12.4) следует, что при фиксированном p существует Δ такое, что условие (12.5) будет выполнено тогда и только тогда, когда $0{,}25 < p < 0{,}75$.

Своеобразие постановки задачи проверки гипотезы состоит в том, что при росте k число неизвестных параметров, т.е. координат векторов P_i, растет пропорционально объему данных. Поэтому и столь далекая от оптимальности процедура, как описанная в двух предыдущих абзацах, представляет некоторый практический интерес. Для случая $s \ge 4$ в теории люсианов (см. подраздел 12.4) разработаны методы проверки гипотезы согласованности $H_0 : P_1 = P_2 = ... = P_s$.

Нахождение группового мнения. Пусть $A_1, A_2, ..., A_s$ — случайные толерантности, описывающие мнения s экспертов. Для нахождения группового мнения будем использовать медиану Кемени, т.е. эмпирическое среднее относительно расстояния Кемени, введенного в главе 11. Медианой Кемени является

$$A_{\text{ср}} = \operatorname{Arg}\min_{A} \sum_{p=1}^{s} d(A_p, A).$$

Легко видеть, что $A_{\text{ср}} = \|a_{cp}(i, j)\|$ удовлетворяет условию $a_{\text{ср}}(i, j) = 1$, если

$$\sum_{p=1}^{s} a_p(i, j) > \frac{s}{2},$$

и $a_{\text{ср}}(i, j) = 0$, если

$$\sum_{p=1}^{s} a_p(i, j) < \frac{s}{2}.$$

Следовательно, при нечетном s групповое мнение $A_{\text{ср}}$ определяется однозначно. При четном s неоднозначность возникает в случае

$$\sum_{p=1}^{s} a_p(i, j) = \frac{s}{2}.$$

Тогда медиана Кемени $A_{\text{ср}}$ — не одна толерантность, а множество толерантностей, минимум суммы расстояний достигается и при $a_{\text{ср}}(i, j) = 1$, и при $a_{\text{ср}}(i, j) = 0$.

Асимптотическое поведение группового мнения (медианы Кемени для толерантностей) вытекает из общих результатов о законах больших чисел в пространствах произвольной природы [77; 89], поэтому рассматривать его здесь нет необходимости.

Дихотомические (бинарные) признаки в классической асимптотике. Многое в предыдущем изложении определялось спецификой толерантностей. В частности, особая роль равномерности распределения на множестве всех толерантностей оправдывала специальное рассмотрение статистик W и N; аксиоматически введенное расстояние d между толерантностями играло важную роль в приведенных ранее результатах. Однако модель толерантностей с независимыми связями уже меньше связана со спецификой толерантностей. В ней толерантности можно рассматривать просто как частный случай таких популярных объектов нечисловой природы, как *люсианы*. Широко применяется следующая модель порождения данных.

Пусть $A_1, A_2, ..., A_s$ — независимые люсианы. Это значит, что статистические данные имеют вид:

$$(A_1, A_2, ..., A_s) = \|X_{ij},\ i = 1, 2, ..., s;\ j = 1, 2, ..., k\|, \tag{12.6}$$

где X_{ij} — независимые в совокупности испытания Бернулли с вероятностями успеха

$$(P_1, P_2, ..., P_s) = \|p_{ij},\ i = 1, 2, ..., s;\ j = 1, 2, ..., k\|, \tag{12.7}$$

где P_i — вектор вероятностей, описывающий распределение люсиана A_i. Особое значение имеют одинаково распределенные люсианы, для которых $P_1 = P_2 = ... = P_s = P$, где символом P обозначен общий вектор вероятностей.

Как обычно в математической статистике, содержательные результаты при изучении модели (12.6)—(12.7) можно получить в асимптотических постановках. При этом есть два принципиально разных предельных перехода: $s \to \infty$ и $k \to \infty$. Первый из них — традиционный: число неизвестных параметров постоянно, объем выборки s растет. Во втором число параметров растет, объем выборки остается постоянным, но общий объем данных ks растет пропорционально числу неизвестных параметров. Аналогом является асимптотическое изучение коэффициентов ранговой корреляции Кендалла и Спирмена: число ранжировок, т.е. объем выборки, постоянно (и равно 2), а число ранжируемых объектов растет.

Вторая постановка изучается в подразделе 12.4, посвященном люсианам. Некоторые задачи в первой постановке рассмотрим здесь.

Случайные толерантности используются, в частности, для оценки нечетких толерантностей [88]. Для описания результатов опроса группы экспертов о сходстве объектов строят нечеткую толерантность $M = \|\mu_{ij}\|$, $\mu_{ij} = l_{ij}/n_{ij}$, где n_{ij} — число ответов о сходстве i-го и j-го объектов, а l_{ij} — число положительных ответов из них. Если эксперты действуют в соответствии с единым вектором параметров P, то M — состоятельная оценка для P. Следующий вопрос при таком подходе: верно ли, что две группы экспертов «думают одинаково», т.е. используют совпадающие вектора P? Рассмотрим эту постановку на более общем языке люсианов.

Пусть $A_1, A_2, ..., A_m$ и $B_1, B_2, ..., B_n$ — две группы независимых в совокупности люсианов, одинаково распределенные в каждой группе с параметрами $P(A)$ и $P(B)$ соответственно. Требуется проверить гипотезу $P(A) = P(B)$. Естественным является переход к пределу при $\min(m, n) \to \infty$.

Пусть гипотеза справедлива. Предположим, что $p_i = p_i(A) = p_i(B) \neq 0$ при всех $i = 1, 2, ..., k$. (Разбор последствий нарушений этого условия оставляем читателю.) Пусть s_i — число единиц на i-м месте в первой группе люсианов, а t_i — во второй. Рассмотрим случайные величины

$$\xi_i = \sqrt{\frac{mn}{m+n}}\left(\frac{s_i}{m} - \frac{t_i}{n}\right)\frac{1}{\sqrt{p_i(1-p_i)}}. \tag{12.8}$$

Они независимы в совокупности. В соответствии с предельными теоремами [77, глава 4] распределения случайных величин ξ_i при $\min(m, n) \to \infty$ сходятся к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1. Эти свойства сохраняются при замене p_i в формуле (12.8) на состоятельные оценки, построенные по статистическим данным, соответствующим i-му месту. Будем использовать эффективную оценку [35, с. 529]

$$p_i^* = \frac{s_i + t_i}{m+n}. \tag{12.9}$$

Подставим формулу (12.9) в (12.8), получим статистики

$$\xi_i^* = \sqrt{\frac{mn(m+n)}{(s_i+t_i)(m+n-s_i-t_i)}}\left(\frac{s_i}{m} - \frac{t_i}{n}\right).$$

Полученные статистики можно использовать для проверки рассматриваемой гипотезы, например, с помощью критериев, основанных на статистиках

$$W = \frac{1}{\sqrt{k}}\sum_{i=1}^{k} a_i \xi_i^*, \quad T = \sum_{i=1}^{k} (\xi_i^*)^2, \quad \sum_{i=1}^{k} a_i^2 = 1.$$

С помощью результатов [77, глава 4] получаем, что W имеет в пределе при $\min(m, n) \to \infty$ стандартное нормальное распределение, а T — распределение хи-квадрат с k степенями свободы.

Рассмотрим распределение статистики W при альтернативных гипотезах. Положим

$$\eta_{1m}^i = \frac{\sqrt{m}\left[\frac{s_i}{m} - p_i(A)\right]}{\sqrt{p_i(A)\left[1-p_i(A)\right]}}, \quad \eta_{2n}^i = \frac{\sqrt{n}\left[\frac{t_i}{n} - p_i(B)\right]}{\sqrt{p_i(B)\left[1-p_i(B)\right]}}.$$

Эти случайные величины независимы, распределение каждой из них при $\min(m, n) \to \infty$ сходится к стандартному нормальному распределению. Поскольку

$$\frac{s_i}{m} = \frac{\eta_{1m}^i}{\sqrt{m}}\sqrt{p_i(A)\left[1-p_i(A)\right]} + p_i(A), \quad \frac{t_i}{n} = \frac{\eta_{2n}^i}{\sqrt{n}}\sqrt{p_i(B)\left[1-p_i(B)\right]} + p_i(B),$$

то

$$\sqrt{\frac{mn}{m+n}}\left(\frac{s_i}{m} - \frac{t_i}{n}\right) = F + G,$$

где

$$F = \sqrt{\frac{mn}{m+n}}\left\{\frac{\eta_{1m}^i}{\sqrt{m}}\sqrt{p_i(A)\left[1-p_i(A)\right]} - \frac{\eta_{2n}^i}{\sqrt{n}}\sqrt{p_i(B)\left[1-p_i(B)\right]}\right\}$$

и

$$G = \sqrt{\frac{mn}{m+n}}\left[p_i(A) - p_i(B)\right].$$

В силу результатов [77, глава 4] распределение F при $\min(m,n)\to\infty$ сближается с нормальным распределением, математическое ожидание которого равно 0, а дисперсия есть

$$\frac{n}{m+n}p_i(A)\left[1-p_i(A)\right]+\frac{m}{m+n}p_i(B)\left[1-p_i(B)\right]\le\frac{1}{4}.$$

Поэтому, чтобы получить собственное (т.е. невырожденное) распределение W при альтернативах, естественно рассмотреть модель

$$p_i(A)=p_i+\frac{\theta_i}{2}\sqrt{\frac{m+n}{mn}}\sqrt{p_i(1-p_i)},\ \ p_i(B)=$$

$$=p_i-\frac{\theta_i}{2}\sqrt{\frac{m+n}{mn}}\sqrt{p_i(1-p_i)},\ \ i=1,2,...,k,$$

где θ_i — некоторые фиксированные числа.

Тогда при $\min(m, n)\to\infty$ оценки p_i^* из (12.9) сходятся к p_i и ξ_i^* являются независимыми асимптотически нормальными случайными величинами с математическими ожиданиями θ_i и единичными дисперсиями. Опираясь на результаты [77, глава 4], заключаем, что распределение статистики W сходится к нормальному распределению с математическим ожиданием

$$\theta_0=\frac{1}{\sqrt{k}}\sum_{i=1}^{k}a_i\theta_i$$

и единичной дисперсией.

Если в последней формуле $\theta_0 = 0$, то асимптотическое распределение W таково же, как и в случае справедливости нулевой гипотезы. От указанного недостатка свободна статистика T. Тем же путем, как и для W, получаем, что при $\min(m, n)\to\infty$ распределение T сходится к нецентральному хи-квадрат- распределению с k степенями свободы и параметром нецентральности

$$\Theta=\sum_{i=1}^{k}\theta_i^2.$$

Можно рассматривать ряд других задач, например проверку совпадения параметров для нескольких групп люсианов (аналог дисперсионного анализа), установление зависимости $P(B)$ от $P(A)$ (аналог регрессионного анализа), отнесение вновь поступающего люсиана к одной из групп (речь идет о задаче диагностики — аналоге дискриминантного анализа; она представляет интерес, например, при применении тестов типа MMPI оценки психического состояния личности) и т.д. Однако принципиальных трудностей на пути развития соответствующих методов не видно, и мы не будем их здесь рассматривать.

Создание соответствующих алгоритмов проводится специалистами по прикладной статистике в соответствии с непосредственными заказами пользователей.

12.3. МЕТОД ПРОВЕРКИ ГИПОТЕЗ ПО СОВОКУПНОСТИ МАЛЫХ ВЫБОРОК

Испытания по двум альтернативным признакам. Одна из областей применения экспертных оценок связана с контролем качества продукции. Для многих видов пищевой продукции наиболее надежные методы контроля — органолептические, т.е основанные на использовании органов чувств человека. Примером является дегустация чая или кофе.

Обсудим статистический приемочный контроль, в котором по результатам испытаний элементов выборки делается вывод о качестве партии продукции [89, глава 13]. В простейшем варианте проводится контроль по альтернативному признаку, при котором возможны лишь два результата контроля конкретной единицы продукции — «соответствует требованиям» или «не соответствует требованиям», иными словами — «да» или «нет».

Рассмотрим статистический приемочный контроль по двум альтернативным признакам одновременно. В терминах 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.

При статистическом контроле качества продукции, в частности при сертификации, чаще всего используют контроль по альтернативным признакам. При этом устанавливают, 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 здесь 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 этой документации 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 таблице 12.2 сведены вместе введенные вероятности.

Таблица 12.2

Вероятности результатов испытаний при контроле по двум альтернативным признакам

Значение $Y(w)$	Значение $X(w)$		*Всего*
	0	1	
0	p_{00}	p_{10}	$1 - p_2$
1	p_{02}	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 $\{w\colon X(w) = 1\}$ и $\{w\colon Y(w) = 1\}$. 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 $p_1 p_2$ является 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 массе и длине. Пусть результаты контроля 1000 единиц продукции представлены в табл. 12.3

Таблица 12.3

Результаты 1000 испытаний по двум альтернативным признакам (случай поглощения)

Значение $Y(w)$	Значение $X(w)$		*Всего*
	0	1	
0	952	0	952
1	0	48	48
Всего	952	48	1 000

Судя по данным табл. 12.3, 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. 12.3, целесообразно перейти к контролю только одного показателя, а не двух. Какого именно? Видимо, того, контроль которого дешевле.

Совсем иная ситуация в случае несовместности дефектов (табл. 12.4).

Таблица 12.4

Результаты 1000 испытаний по двум альтернативным признакам (случай несовместности)

Значение $Y(\omega)$	Значение $X(\omega)$		*Всего*
	0	1	
0	904	48	952
1	48	0	48
Всего	952	48	1 000

Судя по данным табл. 12.4, 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. 12.5) лежит между крайними случаями поглощения и несовместности. Независимость альтернативных признаков обосновывается путем статистической проверки с помощью описанного ниже критерия $n^{1/2}V$.

Таблица 12.5

Результаты 1000 испытаний по двум альтернативным признакам (случай независимости)

Значение $Y(\omega)$	Значение $X(\omega)$		*Всего*
	0	1	
0	909	43	952
1	43	5	48
Всего	952	48	1 000

Согласно данным табл. 12.5 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 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

$$H_0 : p_{11} = p_1 p_2, \tag{12.10}$$

что эквивалентно проверке равенства $p_{00} = (1 - p_1)(1 - p_2)$. Нетрудно проверить, что гипотеза о справедливости равенства (12.10) эквивалентна гипотезе

$$H_0 : p_{00} p_{11} - p_{10} p_{01} = 0. \tag{12.11}$$

В простейшем случае предполагается, что проведено n независимых испытаний (X_i, 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. 12.6.

Таблица 12.6

Общий вид результатов контроля по двум альтернативным признакам

Значение $Y(\omega)$	Значение $X(\omega)$		*Всего*
	0	1	
0	a	b	$a + b$
1	c	d	$c + d$
Всего	$a + c$	$b + d$	n

В таблице 12.6 величина a — число испытаний, в которых $(X_i, Y_i) = (0,0)$, b — число испытаний, в которых $(X_i, Y_i) = (1,0)$, 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 (12.11) может быть основан на статистике

$$Z = ad - bc. \tag{12.12}$$

Как следует из известной формулы для ковариаций мультиномиального вектора (см., например, формулу (6.3.5) в учебнике С. Уилкса [125, с. 153]),

$$M(Z) = n\,(p_{10}p_{01} - p_{00}p_{11}), \tag{12.13}$$

что равно 0 при справедливости гипотезы независимости (12.11).

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 [27, с. 723]). При справедливости гипотезы H_0 и больших n случайная величина nV^2 имеет хи-квадрат-распределение с одной степенью свободы, а величина $n^{1/2}V$ имеет стандартное нормальное распределение с математическим ожиданием 0 и дисперсией 1 [27, с. 736]. Значение $n^{1/2}V$ для данных табл. 12.5 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. 12.7. Тогда

$$V = (50 \times 20 - 10 \times 20)(60 \times 70 \times 30 \times 40)^{-1/2} =$$
$$= (1000 - 200)5\,940\,000^{-1/2} = 800/2245 = 0{,}35635,$$
$$n^{1/2}V = 3{,}5635.$$

Таблица 12.7

Результаты 100 испытаний по двум альтернативным признакам

Значение $Y(w)$	Значение $X(w)$		*Всего*
	0	1	
0	50	10	60
1	20	20	40
Всего	70	30	100

Поскольку полученное значение $n^{1/2}V$ превышает критическое 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. производится 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

$$H_0 : p_{11j} = (p_{10j} + p_{11j})(p_{01j} + p_{11j}),\ j = 1, 2, ..., k, \tag{12.14}$$

или в эквивалентной формулировке

$$H_0 : p_{11j}p_{00j} - p_{10j}p_{01j},\ j = 1, 2, ..., k\ . \tag{12.15}$$

Основная идея состоит в нахождении асимптотического распределения статистики типа $n^{1/2}V$ при росте числа k малых выборок. А именно, будем использовать статистику

$$S = g_1Z_1 + g_2Z_2 + ... + g_kZ_k, \tag{12.16}$$

где $Z_1, Z_2, ..., Z_k$ — статистики, рассчитанные по формуле (12.12) для каждой из k выборок, т.е. $Z_j = a_jd_j - b_jc_j$, $j = 1, 2, ..., k$, а $g_1, g_2, ..., g_k$ — некоторые весовые коэффициенты, которые, в частности, могут совпадать.

Поскольку

$$M(S) = g_1M(Z_1) + g_2M(Z_2) + ... + g_k\,M(Z_k),$$

то при справедливости гипотезы независимости (12.14)—(12.15) имеем $M(S) = 0$, поскольку

$$M(Z_j) = 0,\ j = 1, 2, ..., k,$$

при всех возможных значениях вектора параметров $(p_{00j}, p_{10j}, p_{01j}, p_{11j})$ согласно соотношению (12.13). Поскольку слагаемые в сумме (12.16) независимы, то при росте k случайная величина S в силу Центральной предельной теоремы является асимптотически нормальной. Дисперсия этой величины равна сумме дисперсий слагаемых:

$$D(S) = g_1^2D(Z_1) + g_2^2D(Z_2) + ... + g_k^2D(Z_k). \tag{12.17}$$

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

$$M(T_j) = D(Z_j), j = 1, 2, ..., k. \tag{12.18}$$

Тогда при некоторых математических «условиях регулярности», на которых нет необходимости здесь останавливаться, несмещенная оценка дисперсии статистики S, имеющая, согласно формулам (12.17) и (12.18), 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 (12.14)—(12.15) 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 значения величины $M(S)D(S)^{-1/2}$ при альтернативной гипотезе.

Для реализации намеченного плана осталось научиться несмещенно оценивать $D(Z_j)$. К сожалению, в литературе по несмещенному оцениванию не рассматривают случай мультиномиального распределения, поэтому кратко опишем процедуру построения несмещенной оценки $D(Z_j)$. Поскольку, согласно формулам (12.12) и (12.13),

$$\begin{aligned} D(Z_j) &= M(Z_j^2) - (M(Z_j))^2 = \\ &= M(a_j^2 d_j^2) - 2M(a_j b_j c_j d_j) + M(b_j^2 c_j^2) + \\ &+ n_j^2 (p_{00j} p_{11j} - p_{01j} p_{10j})^2, \end{aligned} \tag{12.19}$$

то для вычисления $D(Z_j)$ достаточно найти входящие в правую часть формулы (12.19) начальные смешанные моменты мультиномиального распределения (четвертого порядка). Теоретически это просто — известен вид характеристической функции мультиномиального распределения (см., например, формулу (6.3.4) в монографии [125, с. 152]), а начальные смешанные моменты равны значениям ее соответствующих производных в точке 0, деленным на нужную степень мнимой единицы, — формула (5.2.3) в монографии [125, с. 131]. Например, с помощью описанной процедуры после некоторых вычислений получаем, что (для упрощения записи здесь и далее опустим индекс j)

$$\begin{aligned} M(a^2 d^2) &= n(n-1)(n-2)(n-3) p_{11}^2 p_{00}^2 + \\ &+ n(n-1)(n-2)(p_{11}^2 p_{00} + p_{11} p_{00}^2) + n(n-1) p_{11} p_{00}. \end{aligned} \tag{12.20}$$

Формула (12.20) 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 (12.20) в качестве образца.

Как следует из формул (12.19) и (12.20), 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 r и m. Известно, что для ковариации мультиномиального вектора

$$M(ad) = -n\, p_{00} p_{11} \tag{12.21}$$

(см., например, формулу (6.3.5) в монографии [125, с. 153]), а потому несмещенной оценкой для $p_{00}p_{11}$ является $(-ad/n)$. Далее, поскольку справедлива аналогичная (12.20) формула

$$M(a^2 d) = n(n-1)(n-2) p_{11} p_{00}^2 + n(n-1) p_{11} p_{00}, \tag{12.22}$$

то с помощью формулы (12.21) преобразуем формулу (12.22) к виду

$$M(a^2 d + (n-1)ad) = n(n-1)(n-2) p_{11} p_{00}^2, \tag{12.23}$$

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 (12.21) и (12.23) 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 увидеть,

$$\frac{Z_j}{\sqrt{T_j}} = V_j \sqrt{n_j - 1},$$

т.е. в случае одной выборки предлагаемый метод проверки независимости совпадает с классическим.

Таким образом, 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 [88, подраздел 4.5]. Нестандартность постановки состоит в том, что число неизвестных параметров растет пропорционально объему данных, т.е. имеет место так называемая «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.

12.4. ТЕОРИЯ ЛЮСИАНОВ

Асимптотика растущей размерности и проверяемые гипотезы. Продолжим изучение модели порождения данных — формулы (12.6)—(12.7) подраздела 12.2. Будем использовать асимптотику $s = \text{const}$, $k \to \infty$. При этом число неизвестных параметров растет пропорционально объему данных.

В последние десятилетия (с начала 1970-х гг.) 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» [76], перенося терминологию исследований по дискриминантному анализу на общий случай. Как известно, в задаче дискриминации в две совокупности (т.е. отнесения вновь появляющегося объекта к одному из двух классов) академик АН СССР А.Н. Колмогоров (1903—1987) предложил рассматривать асимптотику

$$A \to \infty, \quad N_i \to \infty, \quad \frac{A}{N_i} \to \lambda_i > 0, \quad i = 1, 2,$$

где A — размерность пространства (число признаков), N_i — объемы обучающих выборок, λ_i — константы.

Эта асимптотика естественна при обработке многих видов технических, организационно-экономических, социологических, медицинских данных, поскольку число признаков, определяемых для каждого изучаемого объекта, респондента или пациента, обычно имеет тот же порядок, что и объем выборки.

Пусть $A_1, A_2, ..., A_s$ — независимые (между собой) люсианы с векторами параметров $P_1, P_2, ..., P_s$ соответственно. *Гипотезой согласованности* будем называть гипотезу

$$P_1 = P_2 = ... = P_s. \tag{12.24}$$

Для 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 проверки гипотезы (12.24) на проверку гипотезы

$$P_1 = P_2 = ... = P_s = (1/2, 1/2, ..., 1/2). \tag{12.25}$$

Как разъяснено в [83; 88], гипотеза (12.24) более адекватна конкретным задачам обработки реальных данных, например экспертных оценок, чем (12.25). Поэтому полученные от экспертов данные, содержащие противоречия, целесообразно рассматривать как люсианы и проверять гипотезу (12.24), а не подбирать ближайшие ранжировки или разбиения, после чего проверять согласованность 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* называется гипотеза

$$P(A) = P(B).$$

В асимптотике растущей размерности принимаем, что m и n постоянны, а $k \to \infty$.

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. внутри пар. В ранее введенных обозначениях *гипотеза независимости* — это гипотеза

$$P[X_{ij}(A) = 1, X_{ij}(B) = 1] = P[X_{ij}(A) = 1]P[X_{ij}(B) = 1],$$

$$i = 1, 2, ..., s; \quad j = 1, 2, ..., k,$$

проверяемая в предположении

$$P_1(A) = P_2(A) = ... = P_s(A), \quad P_1(B) = P_2(B) = ... = P_s(B).$$

В настоящем разделе излагается метод проверки гипотез о люсианах в асимптотике растущей размерности на примере гипотезы согласованности. Эти результаты получены в [82, 83, 88]. Дальнейшее изучение проведено Г.В. Рыдановой, Т.Н. Дылько, Г.В. Раушенбахом, О.В. Филипповым, А.М. Никифоровым и др. Гипотеза однородности рассмотрена, например, в [82]. Методы проверки гипотезы однородности люсианов развиты и изучены Г.В. Рыдановой [108] на основе описанного далее подхода. Она помимо доказательства предельных теорем провела подробное изучение скорости сходимости методом статистических испытаний.

Методы проверки согласованности люсианов нашли практическое применение, в частности, при анализе медицинских экспертных данных. Они были использованы в кардиологии при анализе данных кинетотопографии [76; 82].

Метод проверки гипотез о люсианах в асимптотике растущей размерности. Будем использовать дальнейшее развитие метода, описанного в подразделе 12.3. Почему нельзя использовать иные подходы, имеющиеся в математической статистике? Поскольку число неизвестных параметров растет вместе с объемом выборки и пропорционально ему, эти параметры не являются мешающими (в том смысле, как этот термин понимается в теории математической статистики). Отметим, что согласно [38] не существует равномерно наиболее мощных критериев, поскольку параметров много (больше одного). Не останавливаясь на других подходах математической статистики, констатируем необходимость применения метода проверки гипотез по совокупности малых выборок.

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 $p \geq 1$, $i = 1, 2, ..., k$. 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-ой выборке вычислить статистику ξ_i такую, что

$$M\xi_i = 0 \tag{12.26}$$

при всех $\theta_i \in \Theta_{0i}$. Очевидно, $\xi_i \equiv 0$ удовлетворяют (12.24). Однако для рассматриваемого метода необходимо, чтобы при всех $\theta_i \in \Theta_{0i}$ ковариационная матрица вектора ξ_i была ненулевой:

$$\operatorname{cov}(\xi_i) = M(\xi_i^T \xi_i) \neq 0. \tag{12.27}$$

В теории математической статистики иногда используют понятие полноты параметрического семейства распределений. Если рассматриваемое семейство является полным — а так и есть для люсианов, — то не существует достаточной статистики, удовлетворяющей одновременно условиям (12.24) и (12.25) (см., например, [8, § 2.12—2.14]). Поэтому будем использовать статистики, не являющиеся достаточными.

Следующее предположение — ковариационные матрицы статистик ξ_i, т.е. $\operatorname{cov}(\xi_i)$, также допускают несмещенные оценки S_i по тем же выборкам:

$$M(S_i) = \operatorname{cov}(\xi_i) \tag{12.28}$$

при всех $\theta_i \in \Theta_{0i}$.

Рассматриваемый метод основан на том, что поскольку случайные векторы ξ_i определяются по независимым между собой выборкам, то ξ_i независимы в совокупности, а потому случайный вектор

$$\xi = \sum_{i=1}^{k} \xi_i \tag{12.29}$$

является суммой независимых случайных векторов, имеет в силу (12.26) нулевое математическое ожидание, а его ковариационная матрица

$$C_k = \sum_{i=1}^{k} \operatorname{cov}(\xi_i).$$

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 [77, глава 4]) 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. Исходя из формулы (12.28) в качестве оценки C_k естественно 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 $k \to \infty$

$$\frac{1}{k}\left(C_k^* - C_k\right) \to 0, \quad \frac{1}{k} C_k \to C. \tag{12.30}$$

В силу результатов [77, глава 4] из асимптотической нормальности ξ и соотношений (12.30) 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 квадратичные функции q от координат вектора η.

Отклонения от нулевой гипотезы приводят, как правило, к нарушению равенств (12.26) и (12.27). Случайный вектор η при этом обычно остается асимптотически нормальным, но с другими параметрами, что может быть обычным образом использовано для построения оптимального решающего правила, соответствующего заданной альтернативе (например, согласно лемме Неймана — Пирсона). Поведение при альтернативах для некоторых гипотез изучено в [82, 108], здесь его не будем рассматривать, поскольку вычисление мощности не требует новых идей.

Несмещенные оценки параметров асимптотического распределения вектора попарных расстояний. Применим описанный метод для проверки гипотезы согласованности люсианов. Исходные данные — люсианы

$$A_j = (X_{1j}, X_{2j}, ..., X_{kj}), j = 1, 2, ..., s.$$

В качестве i-й выборки возьмем совокупность испытаний Бернулли, стоящих на i-м месте в рассматриваемых люсианах:

$$X_{i1}, X_{i2}, ..., X_{is}. \qquad (12.31)$$

При справедливости нулевой гипотезы в (12.31) 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

$$\xi = \{d(A_p, A_q), 1 \le p < q \le s\}, \qquad (12.32)$$

в котором пары (p, q) упорядочены лексикографически,

$$d(A_p, A_q) = \sum_{i=1}^{k} \mu_i \left|X_{ip} - X_{iq}\right|, \quad \mu_i > 0. \qquad (12.33)$$

В главе 6 это расстояние выведено из некоторой системы аксиом (напомним, что совокупность векторов из 0 и 1 размерности k находится во взаимнооднозначном соответствии с совокупностью подмножеств множества из k элементов; при этом 1 соответствует тому, что элемент входит в подмножество, а 0 — что не входит).

Из вида расстояния в формуле (12.33) следует, что введенный в формуле (12.32) вектор ξ имеет вид (12.29) с

$$\xi_i = \mu_i\{|X_{ip} - X_{iq}|, 1 \le p < q \le s\}. \qquad (12.34)$$

Следовательно, для применения описанного ранее метода проверки гипотез о люсианах в асимптотике растущей размерности достаточно построить на основе вектора ξ_i из (12.34) 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 [43, с. 56—57], 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:

$$f^*(p) = \sum_{h=0}^{m} a_h \frac{\gamma^{[h]}}{m^{[h]}}, \qquad (12.35)$$

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 (12.35) получаем несмещенную оценку

$$\beta = \frac{2}{m-1}\left(\gamma - \frac{\gamma^2}{m}\right). \qquad (12.36)$$

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 (12.36) 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:

$$\xi = \left\{d(A_p, A_q),\ 1 \le p < q \le s\right\} - \sum_{i=1}^{k} \mu_i \beta_i e.$$

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 $\mathrm{cov}(\xi_i)$. Для этого достаточно найти математические ожидания элементов матрицы $M(\xi_i^T \xi_i)$ как функции (многочлены) от p_i, а затем использовать формулу (12.35) для получения несмещенных оценок.

Вычисление матрицы $M(\xi_i^T \xi_i)$ хотя и трудоемко, но не содержит каких-либо принципиальных трудностей. В работе [82] вычислены диагональные элементы рассматриваемой матрицы. Вычисление занимает около двух с половиной книжных страниц (с. 299—301). Поэтому здесь приведен только окончательный итог.

Обозначим для краткости $p_i = p$. В работе [82] показано, что

$$D = \left(|X_{ip} - X_{iq}| - \beta_i\right) = \left(2 - \frac{4}{s}\right) p(1-p) - 4\frac{(s-2)(s-3)}{s(s-1)} p^2 (1-p)^2.$$

Если двухэлементные множества $\{p, q\}$ и $\{r, t\}$ не имеют ни одного общего элемента, то

$$C_1 = M\left(|X_{ip} - X_{iq}| - \beta_i\right)\left(|X_{ir} - X_{it}| - \beta_i\right) =$$

$$= -\frac{4}{s} p(1-p) + \frac{8(2s-3)}{s(s-1)} p^2 (1-p)^2,$$

а если имеют ровно один общий элемент, то

$$C_2 = M\left(|X_{ip} - X_{iq}| - \beta_i\right)\left(|X_{ir} - X_{it}| - \beta_i\right) =$$

$$= \left(1 - \frac{4}{s}\right) p(1-p) - 4\frac{(s-2)(s-3)}{s(s-1)} p^2 (1-p)^2.$$

С помощью формулы (12.35) получаем несмещенные оценки для D, C_1 и C_2 как многочленов от p:

$$D^* = \frac{2\gamma_i(s-\gamma_i)}{s^2(s-1)^2}\left\{(s-2)(s-1) - 2(\gamma_i - 1)(s - \gamma_i - 1)\right\};$$

$$C_1^* = \frac{4\gamma_i(s-\gamma_i)}{s^2(s-1)}\left\{\frac{2(2s-3)(\gamma_i - 1)(s - \gamma_i - 1)}{(s-1)(s-2)(s-3)} - 1\right\};$$

$$C_2^* = \frac{\gamma_i(s-\gamma_i)}{s^2(s-1)^2}\left\{(s-4)(s-1) - 4(\gamma_i - 1)(s - \gamma_i - 1)\right\}.$$

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 (12.30), оцениваемую с помощью

$$\mathrm{cov}(\xi)^* = \sum_{i=1}^{k} \mu_i^2 B_i. \tag{12.37}$$

Асимптотическая нормальность доказывается, естественно, в схеме серий. Достаточным условием является существование положительной константы ε такой, что

$$\mu_i \ge \varepsilon,\ \frac{1}{\mu_i} \ge \varepsilon,\ p_i \ge \varepsilon,\ 1 - p_i \ge \varepsilon \tag{12.38}$$

при всех k и i $1 \le i \le k$.

Поскольку D, C_1 и C_2 являются многочленами четвертой степени от p, то несмещенные оценки для них существуют при $s \ge 4$. Если же $s < 4$, то несмещенных оценок не существует. Поэтому указанным методом проверять согласованность можно лишь при числе люсианов $s \ge 4$.

Проверка согласованности люсианов. Пусть α — нормально распределенный случайный вектор размерности $s(s-1)/2$ с нулевым математическим ожиданием и ковариационной матрицей, определенной формулой (12.37). Согласно результатам [77, глава 4] для любой действительнозначной функции f, интегрируемой по Риману по любому гиперкубу, распределения случайных величин $f(\xi)$ и $f(\alpha)$ сближаются при $k \to \infty$. Это означает, что вместо распределения случайной величины $f(\xi)$ для построения критериев проверки гипотез можно использовать распределение случайной величины $f(\alpha)$. 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.

$$(\xi, a)=\sum_{i=1}^{k}\left(\mu_i \sum_{1\le j<t\le s} a_{jt}\left(\left|X_{ij}-X_{it}\right|-\beta_i\right)\right). \tag{12.39}$$

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 выписанных ранее оценок для D, C_1 и C_2.

Отметим, что $(\xi, a) = 0$ при $a_{ij} \equiv 1$; $1 \le j < t \le s$. 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 (12.39) при альтернативах изучено в работе [108].

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 (12.39), 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 (12.39), в которой

$$a_{jt} = 1,\ t = j + 1, j + 2, ..., s;\quad a_{tj} = 1;\quad t = 1, 2, ..., j - 1;$$

$$a_{qt} = 0,\ q \ne j,\ t \ne j,\ 1 \le q < t \le s.$$

Другими словами, она имеет вид:

$$W=\sum_{t=1}^{s} d(A_j, A_t)-(s-1)\sum_{i=1}^{k}\mu_i\beta_i, \tag{12.40}$$

где расстояние d между люсианами определено в (12.33), а β_i — в (12.36) 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 (12.30) — оценка дисперсии W имеет вид:

$$D^*(W)=\sum_{i=1}^{k}\mu_i^2\frac{\gamma_i(s-\gamma_i)}{s^2}\left\{(s-2)^2-4\gamma_i-1)(s-\gamma_i-1)\right\}.$$

Тогда при выполнении некоторых внутриматематических условий регулярности, например условий (12.38), распределение статистики

$$\frac{1}{\sqrt{D^*(W)}}W$$

сходится при $k\to\infty$, $s = \mathrm{const}$ к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1 (при справедливости гипотезы (12.24) согласованности люсианов).

Различные подходы к понятию согласованности. Обсудим условия, при выполнении которых люсианы естественно считать согласованными (а экспертов, чьи мнения отражают люсианы, — имеющими единое мнение, искаженное случайными ошибками), т.е. обсудим различные методы проверки гипотезы (12.24).

Полное индивидуальное согласие имеет место, если никакие два эксперта не являются «несогласованными». Уровень значимости определяется описанным ранее способом (первый частный случай). Однако наличие одной или нескольких пар экспертов, чьи мнения нельзя считать согласованными, не свидетельствует о необходимости отклонения гипотезы (12.24), поскольку парных проверок проводится много, а именно, $s(s-1) \ge 6$, а способы установления уровня значимости при множественных проверках, зависимых между собой, к настоящему времени плохо разработаны [89, подраздел 11.1]. Проблема множественных проверок для количественных признаков обсуждается А.А. Любищевым, выход дается дисперсионным анализом. Можно брать не все попарные проверки, а только для $[s/2]$ пар люсианов, причем разбиение на пары проводить независимо от принятых люсианами значений, как это делает Т.Н. Дылько [20]. Тогда для проверки гипотезы (12.24) на уровне значимости α надо брать для проверки в каждой паре уровень значимости β, где β рассчитывается приближенно: $\beta = \alpha/[s/2]$.

Полное согласие в целом означает, что для любого эксперта мнения всех остальных оказываются с ним согласованными при использовании статистики (12.40) — второй частный случай. Отсутствие подобного согласия для одного или нескольких экспертов не означает отклонения гипотезы согласованности люсианов (12.24) по тем же причинам, что и в предыдущем случае.

Минимальное согласие имеют мнения экспертов, если хотя бы для одного из них гипотеза согласованности не отвергается с помощью статистики (12.40). В этом случае групповое мнение целесообразно строить, выделяя «ядро», о чем подробнее сказано далее.

Расстояние d между люсианами — см. формулу (12.33) — введено аксиоматически в главе 11 (напомним, что реализацию люсиана можно рассматривать как подмножество конечного множества). Там же из иной системы аксиом выведено другое расстояние — D-метрика. Рассмотрим проверку согласованности люсианов с использованием D-метрики. В этом случае расстояние между люсианами A_1 и A_2 имеет вид

$$D(A_1, A_2) = \begin{cases} \dfrac{d(A_1, A_2)}{T(A_1, A_2)}, \ T(A_1, A_2) \neq 0, \\ 0, \ T(A_1, A_2) = 0, \end{cases}$$

где

$$T(A_1, A_2) = \sum_{i=1}^{k} \mu_i \max(X_{i1}, X_{i2}).$$

Ясно, что теория, основанная на D-метрике, из-за наличия знаменателя в только что приведенной формуле существенно сложнее теории, основанной на метрике d. Ясно также, что описанный ранее метод проверки гипотез о люсианах в асимптотике растущей размерности применить не удается. Чтобы продемонстрировать существенное усложнение ситуации, опишем лишь асимптотическое поведение расстояния $D(A_1, A_2)$ между двумя люсианами.

Теорема 12.4 [94]. Пусть p_{1i} и p_{2i} отделены от 0 и 1, а μ_i отделены от 0 и $+\infty$. Тогда расстояние $D(A_1, A_2)$ между люсианами A_1 и A_2 асимптотически нормально при $k \to \infty$ с параметрами

$$t_k = \frac{N_1}{N_2}, \quad q_k = \frac{N_1}{N_2} \sqrt{\frac{N_3}{N_1^2} + \frac{N_4}{N_2^2} - 2\frac{N_5}{N_1 N_2}},$$

т.е. для любого числа x справедливо предельное соотношение

$$\lim_{k \to \infty} P\left\{ \frac{D(A_1, A_2) - t_k}{q_k} \leq x \right\} = \Phi(x),$$

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

Величины N_j, $j = 1, 2, 2, 4, 5$, выражаются через μ_i и величины

$$p_{3i} = p_{1i} + p_{2i} - 2p_{1i}p_{2i}, \quad p_{4i} = p_{1i} + p_{2i} - p_{1i}p_{2i}$$

следующим образом:

$$N_1 = \sum_{i=1}^{k} \mu_i p_{3i}, \quad N_2 = \sum_{i=1}^{k} \mu_i p_{4i}, \quad N_3 = \sum_{i=1}^{k} \mu_i^2 p_{3i}(1 - p_{3i}),$$

$$N_4 = \sum_{i=1}^{k} \mu_i^2 p_{4i}(1 - p_{4i}), \quad N_5 = \sum_{i=1}^{k} \mu_i^2 p_{3i}(1 - p_{4i}).$$

Следствие 12.1. Пусть $p_{1i} = p_1$ и $p_{2i} = p_2$ при всех i, k, причем p_1 и p_2 лежат внутри отрезка (0; 1). Пусть μ_i отделены от 0 и $+\infty$. Тогда расстояние $D(A_1, A_2)$ между люсианами A_1 и A_2 асимптотически нормально при $k \to \infty$ с параметрами

$$t_k = \frac{p_3}{p_4}, \quad q_k^2 = \frac{p_1 p_2 p_3}{p_4^3} \frac{\sum_{i=1}^{k} \mu_i^2}{\left(\sum_{i=1}^{k} \mu_i \right)^2},$$

где

$$p_3 = p_1 + p_2 - 2p_1p_2, \quad p_4 = p_1 + p_2 - p_1p_2.$$

Следствие 12.2. Пусть в предположениях следствия 12.1 $p_1 = p_2 = p$ и $\mu_i = 1$ при всех i, k. Тогда

$$t_k = \frac{2(1-p)}{2-p}, \quad q_k = \frac{2(1-p)}{k(2-p)^3}.$$

Замечание. Пусть в следствии 12.2 $p = {}^1/_2$. Тогда A_1 и A_2 — люсианы, равномерно распределенные на множестве всех последовательностей из 0 и 1 длины k. В частности, эти люсианы могут соответствовать независимым случайным множествам, равномерно распределенным на совокупности всех подмножеств конечного множества из k элементов, или независимым толерантностям, равномерно распределенным на множестве всех толерантностей, определенных на множества из m элементов, где $m(m - 1)/2 = k$. По следствию 12.2, расстояние между люсианами $D(A_1, A_2)$ асимптотически нормально с математическим ожиданием 0,667 и дисперсией 0,296 k^{-1}. Напомним, что распределения коэффициентов ранговой корреляции Кендалла и Спирмена изучены (в основном) лишь при условии равномерности распределения случайных ранжировок на множестве всех возможных ранжировок фиксированного числа объектов. Для теории люсианов случай равномерности распределения — весьма частный, а для теории ранжировок — основной.

Как уже говорилось, отказ от равномерности — привлекательная черта теории люсианов.

Классификация люсианов. Отсутствие согласованности в одном из перечисленных вариантов позволяет сделать заключение о целесообразности разбиения всех люсианов (например, если они выражают мнения экспертов) на группы близких между собой, т.е. о целесообразности классификации люсианов, точнее их кластер-анализа. Поскольку введена мера близости между люсианами $d(A_1, A_2)$ или $D(A_1, A_2)$, то напрашивается следующий способ действий: провести разбиение на кластеры с помощью одного из алгоритмов, основанных на использовании меры близости, а затем проверить мнения в каждом классе на согласованность. Однако применение того или иного алгоритма кластер-анализа, вообще говоря, может нарушить предпосылки описанных способов проверки согласованности (сравните с обсуждением похожей проблемы, связанной с применением регрессионного анализа после кластер-анализа, в [89, глава 11]). Поэтому опишем методы классификации, опирающиеся на результаты проверки согласованности.

Разбиение на кластеры, внутри каждого из которых имеет место «полное индивидуальное согласие», может быть проведено с помощью агломеративного иерархического алгоритма «дальнего соседа», дополненного ограничением сверху на диаметр кластера. Это ограничение строится из статистических соображений, в отличие от методов, обычно используемых в кластер-анализе [89, глава 5]. При этом в качестве меры близости между люсианами используют не расстояния d или D, а модуль статистики, применяемой для проверки согласованности двух люсианов, т.е. статистики (12.39), в которой только одно из чисел a_{ij} отлично от 0. Упомянутое ограничение таково: диаметр кластера не должен превосходить процентной точки предельного распределения, соответствующей используемому при анализе рассматриваемых данных уровню значимости (можно порекомендовать 5%-ный уровень значимости). В результате работы алгоритма получим кластеры, в которых имеется «полное индивидуальное согласие», причем объединение любых двух кластеров приведет к исчезновению этого свойства у объединения. Поскольку способ выделения итогового разбиения из иерархического дерева разбиений имеет вероятностно-статистическое обоснование, изложенное ранее, то описанный метод классификации люсианов следует считать — в терминологии [104] — не методом анализа данных, а вероятностно-статистическим методом.

Кластеры «с полным согласием в целом» могут быть получены с помощью агломеративного иерархического алгоритма, в котором мерой близости двух кластеров является максимальное значение модуля статистики (12.40), когда j пробегает номера мнений (люсианов), вошедших в объединение рассматриваемых кластеров, а суммирование в формуле (12.40) проводится по всем люсианам в этом объединении. Ограничение сверху на меру близости кластеров определяется процентной точкой предельного распределения статистики W, заданной формулой (12.40).

Кластеры «с минимальным согласием» можно получить, при фиксированном j выделяя совокупность люсианов, согласованных с A_j в смысле статистики W из формулы (12.40).

На основе двух рассмотренных ранее частных случаев линейной статистики — формула (12.39) — можно строить и другие способы классификации. Например, для каждого люсиана A_m можно выделить кластер «типа шара» (см. [89, глава 5]) из люсианов, попарно согласованных с A_m. Все такие способы имеют вероятностно-статистическое обоснование, и потому к ним относится сказанное ранее относительно выделения кластеров «с полным индивидуальным согласием».

Замечание. Проверка согласованности приведенными ранее критериями может привести к отрицательному результату в двух случаях — либо значение статистики окажется слишком большим, либо слишком малым. Первый означает, что гипотеза согласованности люсианов (12.24) неверна, второй означает, что неверна вероятностная модель реального явления или процесса, основанная на люсианах. С необходимостью учета второй возможности мы столкнулись при применении теории люсианов для анализа данных, полученных при проведении кинетокардиографии у больных инфарктом миокарда [76].

Нахождение среднего. В результате классификации получаем согласованные (в одном из указанных ранее смыслов) группы люсианов. Для каждой из них полезно рассмотреть среднее. В зависимости от конкретных приложений в прикладных исследованиях применяют либо среднее в виде последовательностей 0 и 1, т.е. в виде реализации люсиана, либо среднее в виде последовательности оценок вероятностей $(p_1, p_2, ..., p_k)$. Кроме того, оно может находиться либо с помощью методов, подавляющих «засорения» («выбросы»), либо без учета возможности засорения. Рассмотрим все четыре возможности.

В соответствии с подходом главы 11 при отсутствии засорения эмпирическое среднее ищется как решение задачи

$$\sum_{j=1}^{m} d(A_j, A) \to \min_{A \in X}, \qquad (12.41)$$

где $A_1, A_2, ..., A_m$ — люсианы, входящие в рассматриваемый кластер, X — множество, которому принадлежит среднее. Если X — совокупность последовательностей из 0 и 1, то формула (12.41) дает решение по правилу большинства.

Если X — пространство последовательностей вероятностей, то решением задачи (12.41) является та же последовательность 0 и 1, что и в первом случае. Поэтому в качестве среднего вместо решения задачи (12.41) целесообразно рассматривать просто последовательность частот.

Асимптотическое поведение средних при $m \to \infty$ вытекает из законов больших чисел, теорем, описывающих асимптотику решений экстремальных статистических задач [77, подраздел 6.3], и теоремы Муавра — Лапласа соответственно.

В работе [125] при анализе результатов эксперимента показано, что ответы реальных экспертов формируют многочисленную группу, составляющую «ядро», расположенное вокруг истинного мнения, и отдельных «диссидентов», разбросанных по периферии. Причем оценка истинного мнения по «ядру» является более точной, чем по всей совокупности, поскольку мнения «диссидентов» не отражают истинного мнения. Поэтому для построения группового мнения, в том числе среднего для совокупности люсианов, отражающих мнения экспертов, естественно применять методы, подавляющие мнения «диссидентов», что соответствует методологии робастности.

«Ядро» может быть построено следующим образом. Решается задача (12.41) с конечным множеством X, состоящим из всех исходных люсианов: $X = \{A_1, A_2, ..., A_m\}$, т.е. из результатов наблюдений выбирается тот, который находится «в центре» совокупности результатов наблюдений. Пусть A_j является решением этой задачи. В качестве ядра предлагается рассматривать совокупность всех люсианов, которые попарно согласованы с A_j. Другой вариант: рассматривается кластер с «полным внутренним согласием», куда входит A_j. При этом, очевидно, должно быть изменено (уменьшено) критическое значение по сравнению с процедурой, приведшей к выделению рассматриваемой группы. Затем групповое мнение ищется лишь для элементов «ядра». Описанная процедура особенно необходима в случае, когда не было предварительного разбиения совокупности люсианов на группы согласованных друг с другом. Новым по сравнению со [120] является придание вероятностного смысла порогу, выделяющему «ядро».

Обобщая идею выделения «ядра», приходим к «взвешенным итеративным методам оценивания среднего», введенным и изученным в работе [72]. Их применение для люсианов не требует специальных рассмотрений.

Таким образом, в настоящем разделе представлен ряд методов обработки специального вида объектов нечисловой природы — люсианов.

При этом для решения одной и той же задачи, например задачи классификации, предлагается ряд методов, точно так же, как для решения классической задачи проверки однородности двух независимых выборок имеется большое число методов [89, глава 4].

12.5. МЕТОД ПАРНЫХ СРАВНЕНИЙ

Пример практического применения метода парных сравнений. Деятельность предприятия по реализации товаров и услуг всегда сопряжена с рядом проблем, от качества решения которых зависит его будущее. Руководителю службы маркетинга необходимо знать факторы, сдерживающие продажи, и оценить степень важности каждого из них. При кажущейся очевидности и простоте решения далеко не вся управленческая команда дает однозначный ответ: какая из проблем на текущий момент является наиболее важной. Необходим экспертный опрос на эту тему.

Целью исследования факторов, влияющих на объемы продаж, является их ранжирование по степени важности. Для этого среди 25 сотрудников отдела сбыта, а также 10 руководителей завода ГАРО (Великий Новгород) одним из топ-менеджеров предприятия А.А. Пивнем был проведен опрос, в котором предлагалось сравнить попарно факторы, определив более важный среди двух. Итог определялся как среднее арифметическое сумм баллов, набранных каждым фактором у всех опрашиваемых.

Были проанализированы следующие 15 факторов:

- потребительские свойства изделий (качество, надежность, показатели назначения и т.д.);
- уровень цен;
- срок поставки продукции;
- информация о предлагаемых к продаже изделиях;
- уровень гарантийного и сервисного обслуживания;
- работа дилеров, представительств;
- рекламная деятельность;
- численность персонала;
- мотивация труда;
- инициативность персонала;
- маркетинговая деятельность;
- оснащенность техническими средствами;
- квалификация персонала;
- корпоративная культура;
- репутация компании.

На основе анализа результатов парных сравнений построена структурная схема, показывающих степень влияния факторов на объемы продаж (рис. 12.1).

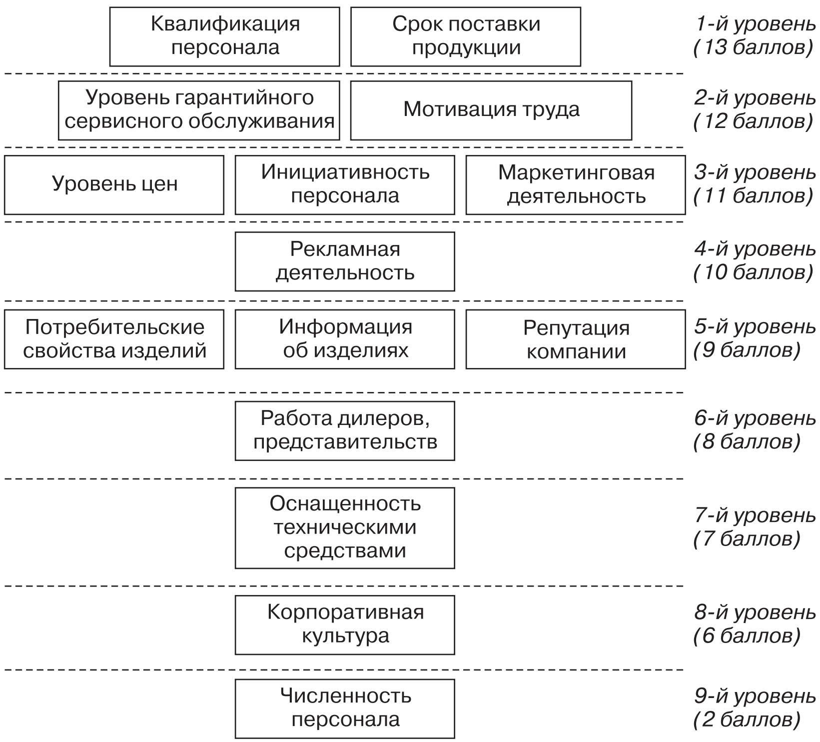


Рис. 12.1. Распределение факторов по их значимости

Наибольшую значимость на сегодняшний день имеет срок поставки продукции и квалификация персонала. Меняются подходы к продвижению товаров на рынке. Ранее успешно применяемые способы продаж (почтовая рассылка рекламы, участие в специализированных выставках, публикации в газетах и специализированных изданиях, конференции и т.д.) сегодня требуют иного качественного подхода. Срок поставки продукции, как правило, связан с производственно-технологическим циклом изготовления и настройки изделий. Мотивация труда, равно как и уровень гарантийного и сервисного обслуживания, имеют также большое значение. Разрабатывается и утверждается новая система оплаты труда, которая позволяет устранить возникающие противоречия. Отдел сервисного обслуживания гаражного оборудования должен

разработать концепцию развития сервисной сети с целью наиболее полного удовлетворения потребителя, а значит и завоевания преимуществ в конкурентной борьбе.

Среди проблем более низкого уровня значимости необходимо отметить место корпоративной культуры. Понимание и осознание себя как части сплоченного коллектива — сложный процесс. Достижение синергетического эффекта возможно только в коллективе, в котором отдельный сотрудник понимает и делает свою работу через понимание целей и задач всей компании. Формированию корпоративной культуры следует уделить особое внимание.

Проведенный анализ дает возможность компании сосредоточить свои усилия на наиболее важных на данный момент обозначенных проблемах. Выбор пути решения каждой из них определяется возможностями компании и опытом руководителей.

Вероятностное моделирование парных сравнений. Опишем общую модель парных сравнений (см., например, [21; 77, глава 5]).

Пусть t объектов $A_1, A_2, ..., A_t$ сравниваются попарно каждым из n экспертов. Следовательно, возможных пар для сравнения имеется $s = t(t-1)/2$. Эксперт с номером γ делает r_γ повторных сравнений для каждой из s возможностей. Пусть $X(i, j, \gamma, \delta)$, $i, j = 1, 2, ..., t$, $i \neq j$, $\gamma = 1, 2, ..., n$; $\delta = 1, 2, ..., r_\gamma$, — случайная величина, принимающая значения 1 или 0 в зависимости от того, предпочитает ли эксперт γ объект A_i или объект A_j в δ-м сравнении двух объектов. Обычно принимают, что все сравнения проводятся независимо друг от друга, так что случайные величины $X(i, j, \gamma, \delta)$ независимы в совокупности, если не считать того, что $X(i, j, \gamma, \delta) + X(j, i, \gamma, \delta) = 1$. Положим

$$P[X(i, j, \gamma, \delta) = 1] = \pi(i, j, \gamma, \delta).$$

Ясно, что описанная модель парных сравнений представляет собой частный случай люсиана (в другой терминологии — бернуллиевского вектора). В этой модели число наблюдений равно числу неизвестных параметров, поэтому для получения статистических выводов необходимо наложить те или иные априорные условия на $\pi(i, j, \gamma, \delta)$, например:

$\pi(i, j, \gamma, \delta) = \pi(i, j, \gamma)$ (нет эффекта от повторений);

$\pi(i, j, \gamma, \delta) = \pi(i, j)$ (нет эффекта от повторений и от экспертов).

Теорию независимых парных сравнений целесообразно разделить на две части — непараметрическую, в которой статистические задачи ставятся непосредственно в терминах $\pi(i, j, \gamma, \delta)$, и параметрическую, в которой вероятности $\pi(i, j, \gamma, \delta)$ выражаются через меньшее число иных параметров. Ряд результатов непараметрической теории парных сравнений непосредственно вытекает из теории люсианов.

В параметрической теории парных сравнений наиболее популярна линейная модель, в которой предполагается, что каждому объекту A_i можно сопоставить некоторую «ценность» V_i так, что вероятность предпочтения $\pi(i, j)$ (т.е. предполагается дополнительно, что эффект от повторений и от экспертов отсутствует) выражается следующим образом:

$$\pi(i, j) = H(V_i - V_j), \tag{12.42}$$

где $H(x)$ — функция распределения, симметричная относительно 0, т.е.

$$H(-x) = 1 - H(x) \tag{12.43}$$

при всех x.

Широко применяются модели Терстоуна — Мостеллера и Брэдли —Терри, в которых $H(x)$ — соответственно функции нормального и логистического распределений. С прикладной точки зрения эти две модели практически совпадают. Действительно, поскольку функция $\Phi(x)$ стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1 и функция

$$\Psi(x) = e^x (1 + e^x)^{-1}$$

стандартного логистического распределения удовлетворяют соотношению [77, глава 5]

$$\sup_{x \in R^1} |\Phi(x) - \Psi(1{,}7x)| < 0{,}01,$$

то для обоснованного выбора по статистическим данным между моделями Терстоуна — Мостеллера и Брэдли — Терри необходимо не менее тысячи наблюдений. Ясно, что при реальном проведении экспертного опроса число наблюдений, по крайней мере, на порядок меньше.

Соотношение (12.42) вытекает из следующей модели поведения эксперта: он измеряет «ценность» V_i и V_j объектов A_i и A_j, но с ошибками ε_i и ε_j соответственно, а затем сравнивает свои оценки ценности объектов $y_i = V_i + \varepsilon_i$ и $y_j = V_j + \varepsilon_j$. Если $y_i > y_j$, то он предпочитает A_i, в противном случае — A_j. Тогда

$$\pi(i, j) = P(\varepsilon_i - \varepsilon_j < V_i - V_j) = H(V_i - V_j). \tag{12.44}$$

Обычно предполагают, что субъективные ошибки эксперта ε_i и ε_j независимы и имеют одно и то же непрерывное распределение. Тогда функция распределения $H(x)$ из соотношения (12.44) непрерывна и удовлетворяет функциональному уравнению (12.43).

Пример 12.1. При опросе экспертов (август 2001 г.) попарно сравнивались четыре компании — ТНК, ЛУКОЙЛ, ЮКОС, Татнефть, продающие автомобильное топливо ($t = 4$). Сравнение проводилось по качеству бензина. Для сравнения имеется $s = t(t - 1)/2 = 6$ пар. Результаты парных сравнений приведены в таблице. По ним необходимо определить взаимное положение четырех компаний на оси «качество бензина», т.е. найти их «ценности»: V_1, V_2, V_3, V_4.

Сравнение компаний по качеству бензина

Пары	Частота выбора первого элемента пары	Частота выбора второго элемента пары
ТНК—ЛУКОЙЛ	$\pi(1,2) = 0{,}508$	$\pi(2,1) = 0{,}492$
ТНК—ЮКОС	$\pi(1,3) = 0{,}331$	$\pi(3,1) = 0{,}669$
ТНК—Татнефть	$\pi(1,4) = 0{,}990$	$\pi(4,1) = 0{,}010$
ЛУКОЙЛ—ЮКОС	$\pi(2,3) = 0{,}338$	$\pi(3,2) = 0{,}662$
ЛУКОЙЛ—Татнефть	$\pi(2,4) = 0{,}990$	$\pi(4,2) = 0{,}010$
ЮКОС—Татнефть	$\pi(3,4) = 0{,}997$	$\pi(4,3) = 0{,}003$

Применим модель Терстоуна — Мостеллера, согласно которой погрешности мнений экспертов ε_i являются независимыми нормально распределенными случайными величинами с нулевым математическим ожиданием и дисперсией σ^2.

Легко видеть, что «ценности» V_1, V_2, V_3, V_4 измерены в шкале интервалов. Начало координат можно выбрать произвольно, поскольку вероятности результатов сравнения зависят только от попарных разностей «ценностей» V_1, V_2, V_3, V_4. Например, можно положить $V_4 = 0$. Единицу измерения также можно выбрать произвольно. При изменении единицы измерения меняется σ^2, точнее, единица измерения однозначно связана с величиной σ. Дисперсия разности $\varepsilon_i - \varepsilon_j$ равна $2\sigma^2$. В соответствии с формулой (12.44) удобно выбрать единицу измерения так, чтобы $2\sigma^2 = 1$, т.е. $\sigma = 1/\sqrt{2}$. Тогда H в формуле (12.44) — это функция Φ стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1.

В соответствии с формулой (12.44) имеем систему шести уравнений с тремя неизвестными:

$$\Phi(V_1 - V_2) = \pi(1, 2) = 0{,}508,$$

$$\Phi(V_1 - V_3) = \pi(1, 3) = 0{,}331,$$

$$\Phi(V_1) = \pi(1, 4) = 0{,}990,$$

$$\Phi(V_2 - V_3) = \pi(2, 3) = 0{,}338,$$

$$\Phi(V_2) = \pi(2, 4) = 0{,}990,$$

$$\Phi(V_3) = \pi(3, 4) = 0{,}997.$$

Применяя к каждому их этих уравнений преобразование Φ^{-1}, получаем систему шести линейных уравнений с тремя неизвестными:

$$V_1 - V_2 = a_1 = \Phi^{-1}(0{,}508) = 0{,}020054,$$

$$V_1 - V_3 = a_2 = \Phi^{-1}(0{,}331) = -0{,}437154,$$

$$V_1 = a_3 = \Phi^{-1}(0{,}990) = 2{,}326348,$$

$$V_2 - V_3 = a_4 = \Phi^{-1}(0{,}338) = -0{,}417928,$$

$$V_2 = a_5 = \Phi^{-1}(0{,}990) = 2{,}326348,$$

$$V_3 = a_6 = \Phi^{-1}(0{,}997) = 2{,}747781.$$

Значения Φ^{-1} взяты из таблиц сборника математической статистики.

В полученной системе число уравнений больше числа неизвестных, т.е. система переопределена. Дальнейшие расчеты могут проводиться разными способами. Простейшие из них состоят в том, чтобы выбрать три уравнения, а именно третье, пятое и шестое, которые и дают искомые значения:

$$V_1 = V_2 = 2{,}326348,\ V_3 = 2{,}747781.$$

Таким образом, качество бензина лучше всего у ЮКОСа, оно несколько хуже у ТНК и ЛУКОЙЛа, одинаковых по этому показателю, а Татнефть значительно хуже тройки лидеров. Можно показать, что если модель Терстоуна — Мостеллера верна и число экспертов достаточно велико, то отбрасывание «лишних» уравнений является корректным способом обработки экспертных данных, поскольку дает состоятельные оценки «ценностей» $V_1, V_2, ..., V_n$.

Однако ясно, что при отбрасывании трех уравнений из шести часть информации теряется. Например, первое уравнение показывает, что по мнению экспертов качество бензина у ТНК несколько лучше, чем у ЛУКОЙЛа. Поэтому целесообразно применить метод наименьших квадратов для оценивания V_1, V_2, V_3, V_4. А именно, рассмотрим функцию трех переменных

$$f(V_1, V_2, V_3) = (V_1 - V_2 - a_1)^2 + (V_1 - V_3 - a_2)^2 + (V_1 - a_3)^2 + \\ +(V_2 - V_3 - a_4)^2 + (V_2 - a_5)^2 + (V_3 - a_6)^2.$$

Оценки по методу наименьших квадратов — это результат минимизации функции $f(V_1, V_2, V_3)$ по совокупности переменных V_1, V_2, V_3. Для минимизации этой функции достаточно приравнять 0 частные производные этой функции по V_1, V_2, V_3. Имеем:

$$\frac{\partial f}{\partial V_1} = 2(V_1 - V_2 - a_1) + 2(V_1 - V_3 - a_2) + 2(V_1 - a_3),$$

$$\frac{\partial f}{\partial V_2} = -2(V_1 - V_2 - a_1) + 2(V_2 - V_3 - a_4) + 2(V_2 - a_5),$$

$$\frac{\partial f}{\partial V_3} = -2(V_1 - V_3 - a_2) - 2(V_2 - V_3 - a_4) + 2(V_3 - a_6).$$

Приравнивая частные производные 0, деля на 2, раскрывая скобки и перенося свободные члены в правую часть, получаем систему трех линейных уравнений с тремя неизвестными:

$$3V_1 - V_2 - V_3 = a_1 + a_2 + a_3,$$
$$-V_1 + 3V_2 - V_3 = -a_1 + a_4 + a_5,$$
$$-V_1 - V_2 + 3V_3 = -a_2 - a_4 + a_6.$$

Решение этой системы не представляет трудностей.

Вообще говоря, не всегда сравниваемые объекты можно представить точками на прямой, т.е. не всегда их можно линейно упорядочить. Возможно, более соответствует данным опроса экспертов представление объектов точками на плоскости или в пространстве большей размерности. В статистике парных сравнений [21] разработаны методы проверки адекватности модели Терстоуна — Мостеллера и других параметрических моделей. Для этого обычно используются критерии типа хи-квадрат.

Разработано много интересных и полезных методов анализа результатов парных сравнений [109]. Во многих областях прикладной статистики, и в частности при анализе мнений экспертов, полезна теория несмещенных оценок [11].

Контрольные вопросы

1. В чем состоит «турнирный» метод ранжирования вариантов?
2. Как связаны случайные толерантности и нечеткие толерантности?
3. Какие задачи проверки статистических гипотез рассматривают в теории случайных толерантностей?
4. Как проверяют гипотезы согласованности, однородности и независимости в теории люсианов?
5. В чем заключаются вероятностно-статистические методы классификации люсианов?

Темы докладов и рефератов

1. Рассчет мощности статистик W и N, рассматриваемых в теории равномерно распределенных случайных толерантностей.
2. Распределение при альтернативах статистики T, используемой для проверки однородности двух групп люсианов (при безграничном росте объемов групп).
3. Несмещенные оценки в прикладной статистике.
4. Применение метода проверки гипотез по совокупности малых выборок в задачах обнаружения эффекта и проверки однородности.
5. Классификация мнений экспертов и проверка согласованности мнений экспертов, выраженных люсианами.
6. Использование люсианов в теории и практике экспертных оценок.

ГЛАВА 13
РЕЙТИНГИ (ОБОБЩЕННЫЕ ПОКАЗАТЕЛИ)

13.1. БИНАРНЫЕ РЕЙТИНГИ

Слово «рейтинг» происходит от английского *to rate* (оценивать) и *rating* (оценка, оценивание). Рейтинги строят обычно на основе анализа многих показателей, как объективных, так и оцениваемых экспертно. Технологии объединения оценок единичных показателей в групповые и обобщенные также обычно бывают экспертными. Примером достаточно сложного рейтинга является оценка вероятности успешного выполнения инновационного проекта. Рейтинги используются в различных процедурах принятия решений, прежде всего для оценивания, выбора, планирования.

Определение бинарного рейтинга. Обсудим наиболее простой случай, когда рейтинговая оценка принимает два значения, для простоты изложения — 0 и 1. Такие рейтинги будем называть ***бинарными***. Например, потенциальный клиент банка может быть кредитоспособным или нет, сам банк — надежным или нет, больной — тяжелым или нет. Для выбора одного из двух возможных решений достаточно, чтобы рейтинговая оценка принимала два значения.

Иногда проводят избыточную работу, строя рейтинг с большим числом значений, например в виде функции $f(x_1, x_2, ..., x_m)$ от единичных показателей (факторов) $x_1, x_2, ..., x_m$. В таких случаях для принятия решения используют некоторое граничное значение K, принимают одно решение, если

$$f(x_1, x_2, ..., x_m) < K,$$

и альтернативное, если

$$f(x_1, x_2, ..., x_m) \geq K.$$

Можно сказать, что в этом случае для принятия решения используется бинарный рейтинг вида $g[f(x_1, x_2, ..., x_m)]$, где функция g принимает два значения, а именно $g(z) = 0$ при $z < K$ и $g(z) = 1$ при $z \geq K$.

На основе бинарных рейтингов можно сконструировать рейтинг с большим числом градаций. Пусть рейтинговая оценка h принимает одно из трех значений $A < B < C$. С ней можно связать два бинарных рейтинга p и q таких, что для первого из них $p = 0$ при $h < C$ и $p = 1$ при

$h = C$, для второго $q = 0$ при $h < B$ и $q = 1$ при $h \geq B$. Ясно, что $h = A$ тогда и только тогда, когда $p = q = 0$, и $h = C$ тогда и только тогда, когда $p = q = 1$, в то время как $h = B$ тогда и только тогда, когда $p = 0$, $q = 1$. Таким образом, использование рейтинга h с тремя возможными значениями эквивалентно использованию двух бинарных рейтингов p и q.

Бинарные рейтинги и дискриминантный анализ. Объект оценки с помощью бинарного рейтинга относится к одному из двух классов. Следовательно, теория бинарных рейтингов — часть теории классификации.

Математическая теория классификации — обширная область прикладной статистики и эконометрики [77; 89]. 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 их авторы и не упоминали термин «классификация». Это предполагает ее сложную внутреннюю структуру.

Следует иметь в виду, что в литературных источниках наряду с термином «классификация» в близких смыслах используются термины «группирование», «распознавание образов», «диагностирование», «дискриминация», «сортирование» 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 (УДК), публикаций, классификатор товаров на основе штрих-кодов.

Другое направление опирается на опыт технических исследований, экономики, маркетинговых исследований, социологии, медицины. Типичные задачи — техническое и медицинское диагностирование, в том числе построение бинарных рейтингов, 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 используется ЭВМ. 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 1960-х г. внутри прикладной статистики достаточно четко оформилась область, посвященная методам классификации. Несколько модифицируя формулировки М. Дж. Кендалла и А. Стьюарта 1966 г. (см. русский перевод [26, с. 437]), в теории классификации выделим три подобласти: дискриминация (дискриминантный анализ), кластеризация (кластер-анализ), группирование. 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 *группировании* 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» (от англ. *cluster* — 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 группировании, наоборот, «...мы хотим разбить элементы на группы независимо от того, естественны ли границы разбиения или нет» [26, 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. Поскольку бинарная рейтинговая оценка принимает только два значения, то может случиться так, что близкие по своим параметрам (т.е. похожие) объекты будут иметь разные рейтинги — если две группы, соответствующие определенному значению рейтинга, не имеют резких различий.

Задачи кластеризации и группирования 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 возможно только их группирование, поскольку совокупность объектов достаточно однородна и не разбивается на резко разделяющиеся между собой кластеры.

Как правило, в математических задачах кластеризации и группирования основное — выбор метрики, расстояния между объектами, меры близости, сходства, различия. Хорошо известно, что для любого заданного разбиения объектов на группы и любого числа $\varepsilon > 0$ 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.

Для построения системы диагностических классов используют разнообразные методы кластерного анализа и группирования объектов. Наименее известен второй член триады (отсутствующий у Кендалла и Стьюарта [26]) — изучение отношений эквивалентности, полученных в результате построения системы диагностических классов. Статистический анализ, отношений эквивалентности полученных, в частности, экспертами, — часть статистики бинарных отношений и тем самым — статистики объектов нечисловой природы [76; 77; 89].

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. 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 группирования). 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 два 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: при одном подходе — исходя из заданного числа кластеров, при другом — из соображений предметной области, при третьем — 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* (в этом алгоритме 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 группирования.

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 группирование.

Часто применяется 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 [88]), 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 группирования 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.

Подходы к построению рейтинговых оценок (правил диагностики, прогностических правил). Для решения задач диагностики используют два подхода — параметрический и непараметрический. Первый из них обычно основан на использовании того или иного индекса (рейтинга) и сравнения его с порогом. Индекс может быть построен по статистическим данным, например как в классическом линейном дискриминантном анализе Фишера [26; 138]. 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 без обоснования принимают гипотезу о том, что ковариационные матрицы для различных классов совпадают. Именно в таких предположениях сформулирован классический дискриминантный анализ Фишера. Как известно, обычно не только нет теоретических оснований считать, что наблюдения извлечены из нормального распределения, но и проверка статистических гипотез согласия с нормальным законом дает отрицательный результат [77; 89]. Известно также, что по выборкам, объем которых не превосходит 50, нельзя сделать обоснованный вывод о принадлежности к нормальному закону [113].

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 подробно рассмотрены в литературе по прикладной статистике и эконометрике [77; 89]. На основе этих оценок могут быть построены непараметрические бинарные рейтинги. Достоинством таких рейтингом является их универсальность, возможность применения без необходимости обоснования трудно проверяемых условий, например нормальности распределения характеристик объектов оценки. Недостатком является отсутствие явных формул, задающих рейтинг в виде некоторой конкретной функции $f(x_1, x_2, ..., x_m)$ от единичных показателей (факторов) $x_1, x_2, ..., x_m$, описывающих объект оценки.

Кроме того, для построения непараметрического бинарного рейтинга нужны обучающие выборки, например выборка описаний (объективных и экспертных данных) кредитоспособных потенциальных клиентов банка и аналогичная выборка некредитоспособных — для построения рейтинга кредитоспособности.

13.2. СРАВНЕНИЕ РЕЙТИНГОВ И ЛИНЕЙНЫЕ РЕЙТИНГИ

Из-за своей простоты популярны линейные рейтинги

$$f(x_1, x_2, ..., x_m) = a_1x_1 + a_2x_2 + ... + a_mx_m$$

в виде линейной функции от единичных показателей (факторов) x_1, $x_2, ..., x_m$.

Коэффициенты $a_1, a_2, ..., a_m$ называют *коэффициентами важности* (весомости, значимости). Их определяют либо экспертным путем, либо оценивают по статистическим данным, используя обучающие выборки. Например, строят рейтинг (интегральный показатель, оценку) финансового положения предприятия в виде линейной функции от некоторого числа переменных (показателей, факторов). Эта функция находится с помощью линейного дискриминантного анализа Фишера [26, 138] и используется для принятия решения о финансовом положении предприятия.

Такой подход хорошо известен в эконометрике. В частности, с критических позиций он описан в главе 5 учебника [89]. Подход является устаревшим. Обоснованность его сомнительна, поскольку он основывается на модели многомерного нормального распределения. В настоящее время, как разъяснено в предыдущем разделе, рекомендуется применять непараметрический дискриминантный анализ, основанный на непараметрических ядерных оценках плотностей классов по обучающим же выборкам.

Однако и устаревший подход может дать практически полезные выводы. Обычно его применение разбито на этапы. Первый — построение системы показателей. Сначала составляют возможно более полный исходный перечень (специалисты по финансово-хозяйственной деятельности предприятия выделяют сотни и тысячи показателей). Затем список показателей сокращают. Например, проводят кластер-анализ показателей, оставляя из каждого кластера по одному представителю. Отбор информативного подмножества признаков в дискриминантном анализе — самостоятельный раздел прикладной статистики. Следующий этап — непосредственное построение линейного рейтинга на основе отобранных показателей с помощью алгоритмов дискриминантного анализа Фишера.

По одним и тем же данным могут быть построены различные рейтинги. Например, с помощью обучающих выборок можно построить непараметрический бинарный рейтинг (заданный алгоритмически) и линейный рейтинг. В той же прикладной задаче может оказаться полезным также и линейный рейтинг на основе экспертных оценок коэффициентов.

Обсудим два вопроса. Как сравнивать рейтинги, какой из них лучше? Можно ли вообще использовать линейный рейтинг?

О сравнении алгоритмов диагностики по результатам обработки реальных данных. Из трех этапов развития теории классификации в конкретной области рассмотрим этап 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 такого 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 далее мы покажем, что этот показатель качества некорректен, а потому пользоваться им не рекомендуется. Целесообразно применять другой показатель качества алгоритма диагностики — описанную далее оценку специального вида так называемого «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. Естественно 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 [72], что позволяет использовать приводимые далее 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 группой ученых под руководством академика АН СССР И.М. Гельфанда. Применение этого алгоритма с медицинской точки зрения вполне оправдано.

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 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.

Итак, 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.

Для выявления информативного набора признаков целесообразно использовать *метод пересчета на модель линейного дискриминантного анализа,* согласно которому статистической оценкой прогностической «силы» δ является так называемая «эмпирическая прогностическая сила»

$$\delta^* = \Phi(d^*/2),\ d^* = \Phi^{-1}(\kappa) + \Phi^{-1}(\lambda),$$

где $\Phi(x)$ — функция стандартного нормального распределения вероятностей с математическим ожиданием 0 и дисперсией 1, а $\Phi^{-1}(y)$ — обратная ей функция.

Пример 13.1. 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 эмпирическая 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 (т.е. понятие истинной прогностической «силы» δ не используется).

Пусть алгоритм классификации применялся к совокупности, состоящей из m объектов первого класса и n объектов второго класса.

Теорема 13.1. Пусть $m, n \to \infty$. Тогда для всех x

$$P\left\{\frac{\delta^* - \delta}{A(\kappa, \lambda)} < x\right\} \to \Phi(x),$$

где δ^* — эмпирическая оценка истинной «прогностической силы» δ алгоритма диагностики; асимптотическое среднеквадратичное отклонение

$$A(\kappa,\lambda) = \sqrt{\frac{1}{4}\left\{\left[\frac{\varphi(d^*/2)}{\varphi\left[\Phi^{-1}(\kappa)\right]}\right]^2 \frac{\kappa(1-\kappa)}{m} + \left[\frac{\varphi(d^*/2)}{\varphi\left[\Phi^{-1}(\lambda)\right]}\right]^2 \frac{\lambda(1-\lambda)}{n}\right\}};$$

здесь $\varphi(x) = \Phi'(x)$ — плотность стандартного нормального распределения вероятностей с математическим ожиданием 0 и дисперсией 1.

С помощью теоремы 13 по κ и λ обычным образом определяют доверительные границы для «прогностической силы» δ.

Пример 13.2. В условиях примера 13.1 при $m = n = 100$ найдем асимптотическое среднее квадратичное 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, имеются в справочниках).

При $m = n = 100$ имеем $A(0{,}90;\ 0{,}80) = 0{,}0252$. При доверительной вероятности $\gamma = 0{,}95$ имеем $\Phi^{-1}(0{,}975) = 1{,}96$, а потому нижняя доверительная граница для прогностической силы δ есть $\delta_Н = 0{,}86 - 1{,}96 \times 0{,}0252 = 0{,}81$, а верхняя доверительная граница такова: $\delta_В = 0{,}86 + 1{,}96 \times 0{,}0252 = 0{,}91$. Аналогичный расчет при $m = n = 1000$ дает $\delta_Н = 0{,}845$, $\delta_В = 0{,}875$.

Можно ли использовать линейный рейтинг? 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 $y \le c$, 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. Укажем способ проверки, т.е. опишем соответствующий критерий проверки статистической гипотезы.

Пусть κ_1 — доля объектов первого класса, для которых $y \le c_1$, а κ_2 — доля объектов первого класса, для которых $c_1 < y \le c_2$. Аналогично пусть λ_2 — доля объектов второго класса, для которых $c_1 < y \le c_2$, 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:

$$d^*(c_1) = \Phi^{-1}(\kappa_1) + \Phi^{-1}(\lambda_2 + \lambda_3),\ d^*(c_2) = \\ = \Phi^{-1}(\kappa_1 + \kappa_2) + \Phi^{-1}(\lambda_3).$$

Теорема 13.2. Если истинные прогностические силы двух правил диагностики совпадают, т.е. $\delta(c_1) = \delta(c_2)$, 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

$$B^2 = \frac{1}{m}T(\kappa_1; \kappa_2) + \frac{1}{n}T(\lambda_2; \lambda_2);$$

$$T(x; y) = \frac{x(1-x)}{\varphi^2\left[\Phi^{-1}(x)\right]} + \frac{(x+y)(1-x-y)}{\varphi^2\left[\Phi^{-1}(x+y)\right]} - \frac{2x(1-x-y)}{\varphi\left[\Phi^{-1}(x)\right]\varphi\left[\Phi^{-1}(x+y)\right]}.$$

Из теоремы 13.2 вытекает метод проверки рассматриваемой гипотезы: при выполнении неравенства

$$\left|\frac{d^*(c_1) - d^*(c_2)}{B}\right| \le \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

она принимается на уровне значимости, асимптотически равном α, в противном случае — отвергается.

Пример 13.3. Пусть данные примеров 13.1 и 13.2 соответствуют порогу c_1. Пусть порогу c_2 соответствуют $\kappa' = 0{,}95$ и $\lambda' = 0{,}70$. Тогда в обозначениях теоремы 13.2 $\kappa_1 = 0{,}90$, $\kappa_2 = 0{,}05$, $\lambda_2 = 0{,}10$, $\lambda_3 = 0{,}70$. Далее $d^*(c_1) = 2{,}12$ (см. пример 13.1), $d^*(c_2) = 2{,}17$, $T(\kappa_1, \kappa_2) = 2{,}22$, $T(\lambda_3, \lambda_2) = 0{,}89$. 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 $m = n = 1000$ и отвергается при $m = n = 5000$.

Экспертно-статистический метод. Оценивание экспертами коэффициентов линейного рейтинга не всегда надежно. Особенно в ситуации, когда экспертов мало, а разброс мнений экспертов велик. Тогда представляется целесообразным не оценивать коэффициенты, а привлечь высококвалифицированных экспертов для глобальной оценки, т.е. оценки непосредственно рейтинга Y. Предположим, что

$$Y = f(x_1, x_2, ..., x_m) = a_1x_1 + a_2x_2 + ... + a_mx_m,$$

и будем использовать глобальные рейтинговые оценки экспертов для расчета коэффициентов (т.е. не будем привлекать экспертов для непосредственной оценки коэффициентов при переменных).

Имеем сначала, что рейтинговые оценки высококвалифицированных экспертов являются числовыми. Тогда в качестве данных, исходных для статистического анализа, имеем выборку $(Y_i; x_{1i}, x_{2i}, ..., x_{mi})$, $i = 1, 2, ..., n$, где n — число ответов высококвалифицированных экспертов, содержащих глобальные оценки рейтинга для n ситуаций. С точки зрения прикладной статистики имеем задачу линейного регрессионного анализа, которая решается стандартными методами (с помощью непараметрического метода наименьших квадратов [77; 89]).

Нет необходимости обязательно требовать, чтобы оценки высококвалифицированных экспертов являлись числами. Можно ограничиться результатами парных сравнений или ранжировками. Ясно, что такого рода глобальные оценки гораздо легче получить, и они будут более надежными (исходя из ранее обоснованного общего утверждения, что нечисловые ответы более естественны для экспертов, чем числовые). Затем по глобальным экспертным оценкам для n ситуаций можно состоятельно оценить коэффициенты линейного рейтинга. Математический аппарат необходим иной, не тот, что в ранее рассмотренном случае глобальных числовых оценок высококвалифицированных экспертов.

В настоящее время теория рейтингов продолжает бурно развиваться. Так, проблемам обоснованного выбора коэффициентов важности посвящена работа В.В. Подиновского [99]. Сравнительный анализ пяти традиционных и четырех относительно новых методов нахождения коэффициентов важности бинарных (т.е. принимающих два значения) факторов осуществлен И.Ф. Шахновым [133]. При этом исходной информацией служат экспертные оценки, имеющие качественный характер.

Очевидна связь теории рейтингов с современной весьма математизированной теорией полезности [128], поскольку рейтинговая оценка — частный случай функций полезности, используемой для упорядочения объектов экспертизы.

Контрольные вопросы

1. Пусть рейтинговая оценка имеет четыре возможных значения. Как ее выразить через бинарные рейтинги?
2. Как соотносятся задачи группирования и задачи кластер-анализа?

3. Почему долю правильной диагностики нецелесообразно использовать как показатель качества алгоритма диагностики (прогностической «силы»)?
4. Как проверить возможность использования линейного рейтинга?

Темы докладов и рефератов

1. Алгоритм, с помощью которого любую рейтинговую оценку, принимающую конечное число значений, можно выразить через бинарные рейтинги.
2. Непараметрические методы диагностики на основе ядерных оценок плотности в пространствах произвольной природы.
3. Непараметрические ядерные оценки плотности в дискретных пространствах.
4. Доказательство теорем 13.1 и 13.2.
5. Алгоритмы оценивания коэффициентов линейного рейтинга в случае глобальных экспертных оценок по методу парных сравнений.
6. Современная теория рейтингов (с использованием [24]).
7. Подходы к выбору коэффициентов важности (на основе [99]).

ГЛАВА 14
ПРИМЕРЫ РАЗРАБОТКИ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ НА ОСНОВЕ ЭКСПЕРТНЫХ ОЦЕНОК

14.1. ЭКСПЕРТНЫЕ ОЦЕНКИ В МАРКЕТИНГОВОМ ИССЛЕДОВАНИИ

Экспертные оценки — мощный интеллектуальный инструмент организационно-экономических исследований. Рассмотрим ряд применений экспертных оценок в менеджменте и экономике при проведении конкретных исследований.

На различных этапах маркетинговых исследований активно применяются различные виды экспертных оценок [50; 121]. Например, в ходе разработки проекта развития инновационных технологий космического приборостроения на примере системы ГЛОНАСС Н.А. Чеботарева (МГТУ им. Н.Э. Баумана, 2005) провела опрос экспертов. Опишем это конкретное маркетинговое исследование.

Цель исследования — выявление направлений развития навигационных приборов в области автомобильного транспорта с целью улучшения их технических характеристик.

Объект исследования — прогнозирование предпочтений потребителей в области технико-функциональных характеристик навигационного прибора.

Содержание анкеты. В предложенной экспертам анкете были представлены 15 важнейших характеристик, определяющих технико-функциональные характеристики навигационного прибора:

1) точность определения навигационных параметров (координат, времени, скорости);
2) сохранение точностных характеристик: при механическом ударе, движении со скоростью 180 км/ч, воздействии тумана, динамической пыли и пр.;
3) габариты прибора;
4) масса прибора;
5) наличие жидкокристаллического цветового экрана;
6) наличие картографической базы данных;
7) объем памяти (число сохраняемых маршрутов);
8) наличие системы мониторинга о пробках на дорогах (выбор оптимальных путей объезда);

9) помехозащищенность (влияние прибора на работу других приборов автомобиля);
10) надежность и прочность в эксплуатации;
11) потребление энергии;
12) время непрерывной работы от аккумуляторной батареи;
13) соотношение цены и качества;
14) современный дизайн;
15) простота в обращении (быстрый доступ к функциям прибора).

Эксперты оценивали перечисленные характеристики по пятибалльной системе:

5 — «абсолютно необходимо»;
4 — «очень важно»;
3 — «может быть важно»;
2 — «не очень важно»;
1 — «абсолютно не важно».

Обработка результатов опроса экспертов. На основе полученных результатов осуществлена статистическая обработка данных.

В опросе приняло участие 12 экспертов, информация была собрана путем самостоятельного заполнения ими анкет.

Поскольку ответы экспертов измерены в порядковой шкале (используется балльная система), согласно теории измерений (см. главу 9) обоснованным является использование медиан в качестве средних баллов и рассмотрение их в качестве интегральных оценок (рейтингов). Полученные результаты можно обработать также путем вычисления средних арифметических из-за их привычности и распространенности, но следует учесть, что данная рекомендация противоречит теории измерений. Однако исходя из концепции устойчивости, согласно которой следует использовать различные методы для обработки одних и тех же данных с целью выделить выводы, получаемые одновременно при всех методах, целесообразно рассмотреть одновременно оба метода — и метод средних арифметических баллов, и метод медианных баллов (см. главу 10).

Метод средних арифметических баллов предполагает подсчет суммы баллов, присвоенных каждой из технико-функциональных характеристик навигационного прибора. Затем эта сумма должна быть разделена на число экспертов (12) для получения среднего арифметического балла (именно эта операция дала название методу). По средним арифметическим баллам проводится итоговая ранжировка (в другой терминологии — упорядочение), исходя из принципа — чем больше средний ранг, чем важнее характеристика (табл. 14.1).

Таблица 14.1

Сводная таблица результатов маркетингового исследования

Эксперт	Характеристика прибора														
	Точность определения навигационных параметров	Сохранение точностных характеристик в различных условиях	Габариты прибора	Масса прибора	Наличие жидкокристаллического цветного экрана	Наличие картографической базы данных	Объем памяти	Наличие системы мониторинга пробок	Помехозащищенность	Надежность	Потребление энергии	Время работы от батареи	Соотношение цены и качества	Дизайн	Простота в обращении
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	5	5	3	3	1	2	5	3	1	5	4	5	2	2	5
2	5	2	3	3	4	4	3	3	4	4	4	3	4	2	5
3	5	4	3	3	3	4	3	3	5	5	3	4	3	3	4
4	5	5	4	2	4	3	5	3	2	5	3	4	3	2	5
5	5	5	3	3	4	2	4	3	4	5	4	4	5	2	5
6	5	4	3	3	3	3	3	2	3	5	4	4	2	2	4
7	5	4	3	2	4	3	4	3	2	5	2	3	4	3	5
8	5	4	2	3	1	3	4	3	2	5	3	4	4	3	5
9	5	5	3	3	1	4	5	2	2	5	4	4	3	2	5
10	4	3	3	3	2	3	4	2	1	5	4	5	4	2	4
11	5	4	4	3	2	2	5	3	1	5	3	4	4	3	4
12	5	4	2	2	2	2	3	3	2	5	4	4	4	2	5

Наибольший средний ранг, равный 4,92, соответствует двум характеристикам прибора — «точность определения навигационных параметров» (№ 1) и «надежность и прочность в эксплуатации» (№ 10), следовательно, в итоговой ранжировке они должны стоять на 1-м и 2-м местах. Этим характеристикам присваивается средний ранг $(1 + 2)/2 = 1{,}5$ (табл. 14.1). Следующая по величине сумма, равная 4,67, соответствует эргономической характеристике — «простота в обращении» (№ 15), ей присваивается ранг 3. Характеристике № 2 — «сохранение точностных характеристик в различных условиях» — в итоговой ранжировке присвоен ранг 4. Дальнейшие результаты приведены в табл. 14.2.

Итак, ранжировка по суммам баллов (по средним арифметическим баллам) имеет вид:

$$\{1, 10\} < 15 < 2 < \{7, 12\} < \{11, 13\} < 3 < 6 < \{4, 8\} < 5 < 9 < 14 \quad (14.1)$$

(движение слева направо соответствует уменьшению среднего арифметического балла). Поскольку характеристики № 1 и 10, а также № 7 и 12, № 11 и 13, № 4 и 8 получили одинаковую сумму баллов, то по рассматриваемому методу они эквивалентны, а потому объединены в группы — классы эквивалентности (выделены фигурными скобками, как и в главе 10).

Метод медиан баллов. Поскольку ответы экспертов измерены в порядковой шкале, для них, согласно теории измерений (см. главу 9), неправомерно проводить усреднение методом средних арифметических баллов. Надо использовать метод медиан баллов.

Необходимо взять ответы экспертов, соответствующие одной из технико-функциональных характеристик, например характеристике № 11 (потребление энергии). Это баллы. Располагаем их в порядке неубывания (проще было бы сказать — «в порядке возрастания», но поскольку некоторые ответы совпадают, то приходится использовать термин «неубывание»). Получим последовательность: 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4. На центральных местах — 6-м и 7-м — стоят 4 и 4, следовательно, медиана равна $(4 + 4)/2 = 4$. Медианы совокупностей 12 баллов, соответствующих определенным технико-функциональным характеристикам прибора, приведены в табл. 14.1. Итоговый ранг по медианам приведен в последней строке таблицы 14.1 и имеет следующий вид:

$$\{1, 10, 15\} < \{2, 7, 11, 12, 13\} < \{3, 4, 6, 8\} < 5 < \{9, 14\}. \quad (14.2)$$

Выделено пять групп характеристик (кластеров), соответственно значениям медианы баллов 5; 4; 3; 2,5; 2.

Таблица 14.2

Результаты расчетов по методу средних арифметических баллов и методу медиан

Показатель	Характеристика														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Сумма баллов	59	49	36	33	31	35	48	33	29	59	42	48	42	28	56
Среднее арифметическое баллов	4,92	4,08	3,0	2,75	2,58	2,92	4,0	2,75	2,42	4,92	3,5	4,0	3,5	2,33	4,67
Итоговый ранг по среднему арифметическому баллов	1,5	4	9	11,5	13	10	5,5	11,5	14	1,5	7,5	5,5	7,5	15	3
Медианы баллов	5	4	3	3	2,5	3	4	3	2	5	4	4	4	2	5
Итоговый ранг по медианам	2	5	10,5	10,5	13	10,5	5	10,5	14,5	2	5	5	5	14,5	2

Сравнение ранжировок по методу средних арифметических и методу медиан баллов. Нет ни одной противоречивой пары характеристик (в смысле определения противоречивости, данного в подразделе 10.3). Сравнение ранжировок (14.1) и (14.2) показывает, что каждый кластер во второй из них представляет собой объединение нескольких кластеров из первой ранжировки. Формальное применение процедуры согласования ранжировок (см. подраздел 10.3) дает в качестве итоговой ранжировку (14.1). Отметим, что более крупные кластеры ранжировки (14.2) легче поддаются интерпретации.

По итоговой ранжировке можно сделать, в частности, следующие ***выводы***:

1) к наиболее важным характеристикам навигационного прибора следует отнести такие характеристики, как «точность определения навигационных параметров», «надежность и прочность в эксплуатации», «простота в обращении», а также «сохранение точностных характеристик в различных условиях». Следовательно, при разработке навигационного прибора в первую очередь необходимо обеспечить выполнение данных требований. Для этого научно-исследовательские и опытно-конструкторские работы должны быть направлены на создание универсальных микросхем, печатных плат, обеспечивающих получение качественных сигналов приема-передачи со спутников, на разработку программного обеспечения. Задача дизайнеров заключается в разработке эргономичного, удобного в использовании прибора. Необходимо обеспечить конструктивное исполнение корпуса прибора, позволяющее сохранять точностные характеристики в различных условиях. Для получения требуемых объема памяти, длительности работы аккумуляторной батареи целесообразно использовать покупные комплектующие изделия, исходя из целевого использования прибора, решаемых с его помощью задач;
2) такие характеристики, как габариты, масса, современный дизайн, были отнесены экспертами к менее значимым и не несут основной технико-функциональной нагрузки, однако при разработке прибора их следует учесть, чтобы обеспечить эстетичность и гармоничное сочетание с салоном автомобиля и другими устройствами.

14.2. ЭКСПЕРТНЫЕ ТЕХНОЛОГИИ В СИСТЕМЕ «ШЕСТЬ СИГМ»

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 новая система — «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 [97], «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 [127]. Концепция «Шесть сигм» ставит на 1-е 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 млрд дол. [97].

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 процентах). Так, 6 «сигма» соответствуют 3,4 дефектам на 1 000 000 возможностей, или выходу годной продукции 99,99966%. А пока такой высокий уровень не достигнут, можно оценивать ситуацию в 1 «сигма». И промежуточная задача может формулироваться так: с уровня 2,5 «сигма» подняться до уровня 4 «сигма».

Интеллектуальные инструменты. 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. Обратите внимание, что реализация практически всех этих методов осуществляется с активным использованием тех или иных экспертных технологий в сочетании с методами анализа объективных данных.

Инструментарий системы «Шесть сигм» весьма широк. Эти инструменты помогают принимать правильные решения, решать проблемы и управлять переменами. Среди них, как следует из приведенного ранее перечисления, основное место занимают различные экспертные и статистические инструменты. Однако нельзя считать, что система «Шесть сигм» и инструменты «Шести сигм» — это одно и то же.

Как справедливо подчеркнуто в цитированной книге о системе «Шесть сигм», возможно, вы говорите себе: «Мы уже делаем кое-что из этого». И уж, безусловно, читали почти обо всех из названных ранее инструментах, в том числе на страницах данного 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 «Шесть сигм». В этом и состоит

суть нового шага в науке и практике управления предприятием. И этот новый шаг реализуется с помощью интенсивного использования процедур экспертного оценивания.

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;
- стремление к совершенству без боязни поражений.

Конечно, каждый из этих элементов сам по себе хорошо известен. Дело в системе «Шесть сигм», в которую они объединены. В ней 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 место, которое занимают экспертные и статистические методы среди ее инструментов. Система «Шесть сигм» трудоемка, на внедрение нужны годы. Но и эффект велик [127].

Отметим, что «Шесть сигм» можно рассматривать и как новую систему внедрения математических методов исследования [66].

Проблемы внедрения математических методов исследования. Полезно проанализировать изменение представлений о проблемах внедрения современных научных достижений в отечественную практику. В качестве примера для обсуждения рассмотрим теорию и методы планирования эксперимента, об истории которых в нашей стране рассказано в [48]. Как известно, локомотивом работ по планированию эксперимента в нашей стране являлся «незримый коллектив» под руководством В.В. Налимова, чьи основные научные идеи и результаты их практического внедрения рассматривались на страницах журнала «Заводская лаборатория».

Очевидно, совершенно необходим первый этап — разработка самой научной теории до той стадии, когда предлагаемые рекомендации уже можно использовать на практике. Основной результат этого этапа — методические разработки и образцы внедрения. Первый этап в основном завершился к началу 1970-х гг.

Термин «завершился» требует уточнения. Научные исследования, разумеется, продолжались и после 1970 г. Они продолжаются сейчас и будут продолжаться в дальнейшем, поскольку любая научная область может развиваться — при наличии энтузиастов — до бесконечности. Речь о другом — к началу 1970-х гг. была создана методическая база для массового внедрения.

Следующий этап — пропаганда возможностей методов планирования эксперимента, преподавание и подготовка кадров. В статье [48] рассказано о многочисленных акциях 1960—1970-х гг. в этом направлении. Казалось, что дальше все пойдет самотеком. Но не получилось. Широкого потока внедренческих работ не последовало. Блестящие работы не стали образцами для подражания.

И не только для планирования эксперимента. Примерно так же развивалась ситуация с внедрением экономико-математических методов. Хотя были и некоторые незначительные отличия. Удалось организовать Центральный экономико-математический институт РАН, а вот академического института по планированию эксперимента нет до сих пор. И Межфакультетская лаборатория статистических методов МГУ им. М.В. Ломоносова, которая занималась развитием теории и внедрением методов планирования эксперимента, расформирована в середине 1970-х гг. Были и другие примеры того, что организационные успехи по тем или иным причинам не удавалось закрепить [48].

Стало ясно, что создания методов и их пропаганды недостаточно. Внедрение новшеств должно опираться на развитую структуру (систему) высококвалифицированных экспертов. Выявилась необходимость перехода к третьему этапу — этапу разработки организационных форм, обеспечивающих широкое внедрение. Наиболее ярким проявлением этого этапа было учреждение в 1990 г. Всесоюзной статистической ассоциации (ВСА), объединяющей — прежде всего в секции статистических методов — специалистов по математическим методам исследования [89]. В статье [74] тех лет, посвященной проблемам внедрения прикладной статистики и других статистических методов, была развернута программа создания сети научно-исследовательских и внедренческих институтов по этой тематике, аналогичной сети метрологических организаций. К сожалению, все эти глобальные планы организации внедрения рассматриваемых методов в государственном масштабе

остались нереализованными из-за развала СССР и развертывания экономических «реформ» 1990-х гг., приведших к сокращению (в разы!) объемов научных исследований и численности работников в сфере науки и научного обслуживания.

Сейчас мы находимся на четвертом этапе. Надо разрабатывать и широко использовать новые организационные формы внедрения математических методов исследования на отдельных предприятиях. С похожими проблемами сталкиваются разработчики крупных информационных систем управления предприятиями (типа SAP R/3, Oracle, JD Edwards, Baan), занимающиеся их внедрением в конкретных организациях [87]. В частности, необходимо создание соответствующей службы под непосредственным началом одного из высших руководителей организации. Недаром внедрение контроллинга — современных методов управления предприятиями — обычно начинается именно с создания службы контроллинга и прорабатывания ее взаимодействия со всеми остальными структурами предприятия [32].

Система «Шесть сигм» ценна прежде всего своей организационной составляющей — той, которой не уделяли внимания на ранних этапах истории внедрения современных математических методов исследования. Система «Шесть сигм», опирающаяся на постоянное и интенсивное использование знаний и интуиции высококвалифицированных экспертов, дает алгоритмы практической деятельности по организации внедрения, чем она и интересна для отечественных специалистов.

Подведем итоги. В России активно разрабатываются теоретические, программные и практические вопросы статистических методов сертификации и управления качеством продукции. Некоторые из них мы кратко рассмотрели. 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 в целях 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.

14.3. ИЕРАРХИЧЕСКАЯ СИСТЕМА ПОКАЗАТЕЛЕЙ ТЕХНИЧЕСКОГО УРОВНЯ И КАЧЕСТВА ПРОДУКЦИИ

Среди иерархических систем показателей одной из наиболее практически важных является система показателей качества продукции. Она представляет интерес для нас в связи с построением рейтингов (см. главу 13). Рассмотрим эту систему показателей подробнее.

Под качеством продукции обычно понимают совокупность свойств продукции, обусловливающих ее пригодность к удовлетворению определенных потребностей в соответствии с ее назначением. Для организации контроля качества, изучения динамики качества продукции, выпускаемой различными производителями, планирования производства, управления инвестициями и инновациями и решения других управленческих задач недостаточно словесного определения понятия качества, необходимо выразить уровень качества в числовой форме. Другими словами, необходимо использовать обобщенные показатели качества, ориентированные на решаемую управленческую задачу и описывающие качество числом. Как их сконструировать?

Выделяют следующие ***группы показателей*** качества:

1) назначения;
2) надежности и долговечности;
3) технологичности;
4) стандартизации и унификации;
5) эргономические;
6) эстетические;
7) патентно-правовые;
8) экономические;
9) экологические;
10) безопасности и т.п. (например, показатели транспортабельности [23, с. 217]).

Ясно, что набор используемых групп зависит от решаемой задачи. Потребителя не интересуют показатели технологичности, стандартизации и унификации, государственные органы сейчас контролируют экологические показатели, безопасность и соблюдение прав на интеллектуальную собственность и т.д. Часто обсуждают соотношение «цена — качество», ясно, что при этом группу экономических показателей не включают в понятие «качество».

Под ***техническим уровнем*** понимают меру использования достижений технического прогресса для удовлетворения конкретных потребностей, степень технического совершенства продукции, новизны и прогрессивности конструкторско-технологических решений. Технический уровень продукции — относительный показатель, который определяется на основе сравнения параметров и характеристик предлагаемого в проекте продукта с показателями базового образца соответствующего уровня, имеющегося в стране и за рубежом. При этом используют как групповые, так и единичные показатели. Как следует из сказанного, при определении технического уровня экономические показатели не учитываются. Конкретные правила расчета технического уровня зафиксированы в соответствующих нормативных документах.

Для практических расчетов каждая группа показателей подробно раскрывается. Например, надежность изделия — сложное свойство, состоящее из трех частных свойств — безотказности, ремонтопригодности, сохраняемости [129]. Эти три свойства стоят на третьем уровне иерархии показателей качества (на первом — верхнем — обобщенный показатель, на втором — перечисленные ранее групповые показатели, в том числе групповой показатель надежности).

Безотказность — свойство технического изделия сохранять работоспособность в течение некоторого времени эксплуатации без вынужденных перерывов. Показателями безотказности являются: вероятность безотказной работы, средняя наработка до первого отказа, наработка на отказ, интенсивность отказов, параметр потока отказов, гарантийная наработка. Все эти показатели стоят на четвертом уровне иерархической системы показателей.

Они могут быть раскрыты с помощью показателей пятого уровня иерархии. Так, вероятность безотказности работы является функцией (в математическом смысле) времени наблюдения. Функцию с достаточной степенью точности можно задать с помощью ее значений в конечном числе точек. Часто с помощью экспертов выбирают «естественные» единицы измерения — час, сутки, год. Средняя наработка до первого отказа — это математическое ожидание, или медиана, или иное теоретическое среднее. А реально все эти средние оцениваются по результатам предварительных испытаний, следовательно, характеристики имеют вид доверительных интервалов и т.д.

Казалось бы, показатели надежности должны оцениваться по результатам объективных измерений. Реально же велика роль экспертов. Именно они задают конкретные процедуры оценивания, выбирая вид характеристик нижнего уровня иерархии, модели и вытекающие из них алгоритмы нахождения числовых значений характеристик. Отметим, что эксперты могут и непосредственно оценивать надежность в качественных терминах, например как высокую, среднюю или низкую.

Заметно больше роль экспертов для других групп показателей. Так, ***эстетические показатели*** характеризуют такие свойства продукции, как выразительность, гармоничность, целостность, соответствие среде и стилю, колористическое (цветовое) оформление и др. В работе В.В. Солодова (МГТУ им. Н.Э. Баумана, 2006) указано, что эстетические показатели определяют информационную выразительность, рациональность формы, целостность композиции, совершенство исполнения продукции, а также стабильность товарного вида. Он строит двухуровневую систему эстетических показателей:

1. Показатели красоты формы (информационной выразительности).

   1.1. Показатель значимости формы — насколько социально значима форма изделия.

   1.2. Показатель соответствия моде формы изделия.

   1.3. Показатель стилевого соответствия, т.е. насколько форма изделия соответствует стилю.

   1.4. Оригинальность формы изделия.

2. Показатели рациональности формы — насколько форма соответствует функциям изделия, конструкции, процессу, в котором участвует изделие.

   2.1. Показатель функционально-конструктивной обусловленности — насколько форма изделия соответствует конструкции, функциям (показатель технологичности).

   2.2. Показатель эргономической обусловленности (эргономический показатель) — насколько форма изделия приспособлена к человеку, насколько она удобна.

3. Показатели целостности и гармоничности формы характеризуют гармоничность и единство элементов изделия.

   3.1. Показатель целостности объемно-пространственной структуры — насколько гармонично выглядит изделие.

   3.2. Показатель тектоничности — насколько в изделии имеет место единство материала и его формы.

   3.3. Показатель пластичности — насколько красивы переходы между поверхностями, составляющими формы изделия.

   3.4. Показатель упорядоченности, гармоничности знаков, табличек, эмблем, гармоничности изображений элементов.

   3.5. Показатель колорита и декоративности (отражает взаимосвязанность цветовых сочетаний и декоративных свойств материалов изделия).

Перечисленные показатели используют при технико-экономическом анализе при маркетинге — на первом этапе жизненного цикла продукции — и учитывают на следующих этапах.

Важны еще две группы показателей, которые определяются технологией производства изделия:

4. Показатели совершенства изготовления элементов формы и поверхностей (показатель четкости исполнения знаков, графических элементов, чистоты поверхностей, тщательности покрытия, напылений).

5. Показатели стабильности товарного вида.

Рассмотренные показатели оцениваются экспертным путем в баллах как дифференциальные эстетические показатели, и при анализе используется интегральный эстетический показатель, который определяется сравнением выбранных эстетических показателей нового изделия по отношению к образцовому (со своими весовыми коэффициентами).

Отметим, что среди чисто эстетических показателей оказался один эргономический (показатель 2.2). Он необходим для целостной оценки эстетической составляющей качества, формально говоря, для построения группового показателя качества. Вместе с тем, как уже отмечалось, ***эргономические показатели*** качества составляют самостоятельную группу показателей. Обычно выделяют четыре подгруппы — гигиенические, антропометрические, физиологические и психофизиологические, психологические показатели.

Гигиенические показатели оценивают соответствие изделия санитарно-гигиеническим нормам и рекомендациям. Речь идет об уровнях шума, температуры, давления, запыленности, освещенности и т.д.

Антропометрические показатели оценивают соответствие изделия размерам и форме человеческого тела в целом и его отдельных частей.

Физиологические и психофизиологические показатели характеризуют соответствие конструкции изделия и его элементов физиологическим свойствам человека, прежде всего особенностям и возможностям его органов чувств.

Психологические показатели оценивают соответствие изделия психологическим возможностям и особенностям человека, совершенство обеспечения информационного обмена в системе «человек—изделие—

среда», влияющего на легкость и быстроту формирования рабочих навыков человека при использовании изделия, на объем, полноту и скорость восприятия информации, поступающей от изделия.

Таким образом, группы эстетических и эргономических показателей имеют многоуровневую иерархическую структуру (не менее трех уровней).

Методы агрегирования показателей качества. После оценивания единичных показателей качества вычисляют значения групповых показателей, а затем обобщенного показателя, дающего итоговую единую оценку объекту оценки (изделию). При этом двигаются снизу вверх, от нижнего уровня иерархии показателей к верхним. Сначала единичные показатели нижнего уровня обобщают в групповые предпоследнего уровня. Затем на каждом шаге по групповым оценкам определенного уровня получают групповые оценки более высокого уровня. Заканчивается процесс расчетом обобщенного показателя (рейтинга). Примеры были приведены в главе 13.

На каждом шаге рассчитывают взвешенные агрегированные показатели. Опишем эту процедуру.

Пусть $X_1, X_2, ..., X_K$ — частные (единичные или групповые) числовые показатели. Пусть каждому из них приписан вес — $A_1, A_2, ..., A_K$, соответственно отражающий их относительную важность (оцененную экспертами или иным способом). Весовые коэффициенты неотрицательны и в сумме составляют 1.

Взвешенные агрегированные показатели можно определить следующим единообразным способом.

Введем (чисто формально) распределение вероятностей, приписывающее каждому значению X_M, $M = 1, 2, ..., K$, вероятность A_M. Для этого распределения обычным образом определим такие характеристики, как математическое ожидание, медиана, начальные моменты, мода и т.д., которые и будем использовать в качестве взвешенных агрегированных показателей или при их расчете.

При этом математическое ожидание дает взвешенное среднее арифметическое, медиана — взвешенную медиану (в частном случае, когда одна из ступенек функции распределения приходится на высоту 0,5, целесообразно ввести понятия левой и правой медиан, т.е. левого и правого концов указанной ступеньки соответственно).

Начальный момент p-го порядка после извлечения корня p-ой степени дает взвешенное степенное. Аналогичным образом получаем обобщенное среднее по Колмогорову общего вида.

Мода указывает на значение наиболее важного показателя.

В соответствии с методологией устойчивости (см. [88] и главу 10) при анализе конкретной ситуации целесообразно одновременно использовать несколько обобщенных показателей, например взвешенную медиану и взвешенное среднее арифметическое. Такая процедура предусмотрена в методике [89]. Хотя согласно теории измерений (см. главу 9), использование среднего взвешенного арифметического некорректно, но приходится учитывать традиции (проблема учета традиций подробно обсуждалась в главе 10).

Наиболее часто используют среднее взвешенное арифметическое

$$\overline{X} = \sum_{M=1}^{K} a_M X_M$$

и среднее взвешенное геометрическое

$$X_{geo} = \prod_{M=1}^{K} X_M^{a_M} = \exp\left\{\sum_{M=1}^{K} a_M \ln X_M\right\}.$$

Одним из широко обсуждаемых в литературе преимуществ среднего взвешенного геометрического является то, что оно равно 0, когда хотя бы одно из усредняемых значений равно 0, в то время как при использовании среднего взвешенного арифметического недопустимо низкое значение одних показателей может быть компенсировано высокими значениями других показателей.

Однако согласно теории измерений и среднее взвешенное арифметическое, и среднее взвешенное геометрическое нельзя использовать для усреднения показателей, измеренных в порядковой шкале. Именно с обсуждения методов агрегирования показателей качества началась разработка проблем поиска и изучения инвариантных алгоритмов в нашей стране [88]. В частности, для усреднения показателей, измеренных в порядковой шкале, рекомендуем использовать медиану или взвешенную медиану.

Иногда используют и иные виды средних, например среднее взвешенное гармоническое, как в [23, с. 235]: «Сводный показатель уровня конкурентности производства продукции предприятия может быть представлен индексом ее качества:

$$Y_0 = \left(\sum_{i=1}^{n} \frac{\alpha_i}{Y_i}\right)^{-1},$$

где n — число видов продуктов, производственных процессов или услуг; α_i — доля затрат на i-й продукт в общей сумме затрат на производство n видов продукции; Y_i — уровень качества i-го продукта (услуги) или эффективности i-го производственного процесса».

Применяют и иные виды показателей. Так, для оценки уровня качества и конкурентности промышленной продукции в качестве критерия используется ***интегральный показатель***, которым является численная характеристика превосходства или конкурентности, являющаяся отношением группового показателя по техническим параметрам к групповому показателю по стоимостным показателям. Подробнее: интегральный показатель качества представляет собой отношение натурального эффекта (в натуральных единицах — штуках, тоннах и др.) от эксплуатации или потребления продукции к суммарным затратам на ее создание и эксплуатацию или потребление, т.е. натуральный удельный эффект от использования продукции (приходящийся на единицу денежных затрат) [23, с. 235—236].

Иерархическая система показателей используется отнюдь не только в задачах управления качеством. Аналогичные системы строят, например, с целью анализа, оценки и управления рисками. Другим примером является иерархическая система показателей в методике сравнительного анализа родственных эконометрических моделей [88]. Эта методика имеет цели:

- оценка по единой схеме качества эконометрических моделей;
- проведение сравнения однотипных эконометрических моделей;
- выбор эконометрических моделей среди однотипных моделей в целях практического использования или углубленной доработки.

Рассматриваемая методика основана на выделении теоретических и эмпирических единичных показателей качества эконометрической модели, построении на их основе групповых и обобщенных показателей качества, их согласования и использования для решения сформулированных ранее задач.

Обобщенный показатель качества, построенный на основе иерархической системы показателей, частный, но весьма важный вид рейтинговой оценки (см. главу 13).

14.4. ПРИМЕНЕНИЕ ЭКСПЕРТНЫХ ОЦЕНОК ПРИ УПОРЯДОЧЕНИИ СИСТЕМЫ ГОСУДАРСТВЕННЫХ СТАНДАРТОВ

Крупное экспертное исследование было проведено с целью упорядочения системы государственных стандартов по статистическим методам управления качеством продукции. Расскажем о нем, обращая внимание на общие выводы и рекомендации, вытекающие из накопленного опыта.

Около 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 [14]. С тех пор в России в статистическом контроле качества было сделано многое, особенно в области теории [88]. Актуальными и в настоящее время остаются работы выдающегося русского математика А.Н. Колмогорова (1903—1987) и его учеников [37].

С начала 1970-х гг. началась разработка государственных стандартов по статистическим методам. В связи с обнаружением в них грубых ошибок 24 из 31 государственного стандарта по статистическим методам были отменены в 1986—1987 гг. (перечень стандартов и описание ошибок приведены в работе [81]). Это решение было основано на результатах крупного экспертного исследования, о котором пойдет речь далее.

Сначала рассмотрим процедуру подготовки такого ответственного нормативного документа, как государственный стандарт.

Типовая процедура подготовки деловых и нормативных документов. Внутри одной организации происходит координация действий специалистов и менеджеров при подготовке различных деловых документов — планов, приказов, предложений, направляемых в другие организации, ответов на распоряжения и запросы властей и др. Обычно один из сотрудников — назовем его Исполнителем — готовит первоначальный вариант документа, который размножают и рассылают на отзыв заинтересованным в нем менеджерам и специалистам, а иногда и в другие организации. Исполнитель составляет сводку отзывов, с одними из замечаний соглашается, против других высказывает возражения. Затем собирают так называемое «согласительное совещание», на которое приглашают всех тех, с чьим мнением Исполнитель не согласен. В результате дискуссии по ряду позиций достигается компромисс и возражения снимаются. Окончательное решение по проекту документа с учетом оставшихся возражений принимает генеральный директор или совет директоров, т.е. высшая инстанция в данной организации.

Во многих случаях подготовка отзывов заменяется на *визирование*, при котором свое согласие менеджеры выражают, накладывая на документ *визу*, т.е. расписываясь (иногда добавляя несколько слов по затрагиваемой проблеме). Например, подготовленное для отправки в другую организацию письмо визируют руководители нескольких отделов, и генеральный директор его подписывает от имени фирмы, не вникая в суть (поскольку каждый день он подписывает десятки писем, то вникать некогда). Адресату уходит письмо, на обратной стороне которого указаны фамилия и телефон Исполнителя (поскольку адресат тоже хорошо знаком с процедурой подготовки документов, он понимает, что по конкретным вопросам надо обращаться к Исполнителю, а не к генеральному директору). В архиве фирмы остается письмо с визами, так что в случае необходимости легко выяснить, кто составил и одобрил документ.

Подготовка, согласование и утверждение федеральных и региональных законов, постановлений органов исполнительной власти, государственных и международных стандартов и иных нормативных документов проходит по более сложной, но похожей схеме. Например, в роли Исполнителя выступает не отдельный специалист, а организация. Подготовленный ею (по описанной схеме) документ рассылается на отзыв. Список рассылки может включать сотни организаций, состав этого списка во многом определяется соответствующими нормативными документами. Отзывы, подготовленные организациями (по описанной схеме подготовки делового документа внутри организации), направляются Исполнителю.

Разрабатываемый документ, как правило, разбит на отдельные пункты, и в отзывах обычно даются замечания и предложения по отдельным пунктам, а не только общая оценка документа. Эти правила подготовки отзывов облегчают сотрудникам организации-разработчика составление сводки, в которой для каждого конкретного пункта приводятся замечания организаций-рецензентов и их предложения по доработке формулировок пунктов. Затем по каждому пункту, включенному в сводку, организация-исполнитель готовит свое заключение, в котором замечания и предложения рецензентов обсуждаются и в итоге обоснованно отклоняются или полностью или частично принимаются. На основе этих заключений Исполнитель готовит вторую редакцию документа.

Организации-рецензенты получают от Исполнителя сводку отзывов (включая заключения по документу в целом и по отдельным пунктам) и вторую редакцию документа. И процесс повторяется.

Оставшиеся несогласованными положения обсуждаются на специальном совещании, организованном Исполнителем, на котором эксперты из различных организаций встречаются лично. Как правило, итогом согласительного совещания является полученный в результате компромисса итоговый проект документа, который поддерживают все участники описанного ранее процесса разработки.

После этого проект государственного стандарта рассматривается на заседании научно-технической комиссии Госстандарта, участники которого уже, как правило, не являются узкими специалистами по тематике обсуждаемого документа. В случае положительного решения совещательного органа — научно-технической комиссии — проект стандарта поступает к лицу, принимающему решение, — председателю Госстандарта (в ранге министра). После утверждения ЛПР проект становится нормативным документом, государственным стандартом, несоблюдение которого преследуется по закону.

Приведенное краткое описание необходимо для понимания проблем работы экспертов в рассматриваемой области деятельности. Оно не заменяет системы нормативных документов, посвященных разработке государственных стандартов или иных нормативных документов. Не является нашей задачей также и обоснование или критика описанной системы подготовки нормативных документов.

Обратим внимание на то, что здесь речь идет об интенсивном использовании многоуровневых экспертных технологий. Велика роль Исполнителя, выступающего в роли коллективного председателя экспертной комиссии. Во многом от него зависит подбор экспертов — организаций-рецензентов, участников согласительного совещания. Причем вопрос о привлечении ведущих ученых и организаторов производства, специалистов по рассматриваемой тематике, остается в компетенции Исполнителя.

В качестве примера рассмотрим подготовку международных стандартов в рамках международной организации по стандартизации ИСО (International Organization for Standardization — ISO). В качестве «организаций» в приведенной ранее схеме выступают национальные органы по стандартизации. В СССР это был Госстандарт. В области статистических методов управления качеством продукции конкретную работу выполнял выделенный Госстандартом подчиненный ему институт — ВНИИ стандартизации. Участие ведущих ученых из вузов или Академии наук не предусматривалось. Аналогична ситуация и в других странах. Результаты очевидны — научно-технический уровень ряда международных стандартов ИСО, по оценке ряда экспертов, отстает от переднего края научных исследований на десятилетия.

Экспертный анализ стандартов по статистическим методам. В нашей стране в послевоенное время расширялся фронт прикладных исследований с использованием статистических методов. В качестве очередного шага с начала 1970-х гг. стали разрабатывать государственные стандарты по статистическим методам управления качеством. Однако вскоре в них были обнаружены грубые ошибки.

Руководство ВНИИ стандартизации в 1985 г. организовало «Рабочую группу по упорядочению системы стандартов по прикладной статистике и другим статистическим методам». Так была названа экспертная комиссия по рассматриваемой тематике. В ее работе приняли участие 66 специалистов, в том числе 15 докторов и 36 кандидатов наук, представляющих предприятия и организации различных отраслей промышленности, академическую и вузовскую науку.

Рабочая группа выделила из своего состава четыре комиссии, которые анализировали стандарты по своей тематике, а именно по прикладной статистике, статистическому контролю качества, статистическим методам регулирования технологических процессов и общим проблемам внедрения статистических методов. Каждый стандарт подробно анализировал специально выделенный рецензент, его письменное заключение размножалось и предоставлялось каждому члену соответствующей комиссии. После тщательного обсуждения на заседании комиссии с участием разработчиков стандартов принималось решение.

Выводы рабочей группы кратко отражены в статьях [74; 81]. 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 сих пор продолжают использоваться инженерами 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 многочисленных работах пропагандируются ошибочные стандарты, регламентирующие применение контрольных карт при статистическом регулировании технологических процессов. В ряде статей, опубликованных в научно-техническом журнале «Надежность и контроль качества» (1988. № 9. С. 48—52; 1991. № 4. С. 31—42 и др.) использовался уже отмененный к тому времени грубо ошибочный ГОСТ 11.006—74 (СТ СЭВ 1190—78). Перечисленные факты делают целесообразным публикацию и популяризацию результатов и выводов рабочей группы и в настоящее время, через 20 лет после окончания анализа стандартов по статистическим методам.

Примеры заключения экспертной комиссии. В качестве примера ошибочного стандарта обсудим ГОСТ 11.006—74 (СТ СЭВ 1190—78)

«Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим». Как следует из его обозначения, этот стандарт был утвержден в 1974 г., а в 1978 г. стал основой международного стандарта в рамках СЭВ — Совета экономической взаимопомощи, объединяющего европейские социалистические страны, Югославию и Кубу.

Этот стандарт является принципиально ошибочным. Основная ошибка известна специалистам с 1950-х гг. и подробно разбирается, например, в статье [79] и учебнике [77]. О нравах разработчиков ГОСТ 11.006—74 (СТ СЭВ 1190—78) свидетельствует судьба отзыва проф. И.Н. Володина (Казанский государственный университет), в котором на 28 машинописных страницах разоблачались ошибки первого проекта стандарта и предлагались способы их устранения. Разработчики стандарта во ВНИИ стандартизации отзыв получили и подшили в архивное «Дело ГОСТа». Все ошибки были бережно сохранены в окончательном тексте стандарта. Не реагировали разработчики стандарта и на дальнейшую критику. В соответствии с текстом стандарта попытки обрабатывать данные в соответствии с научными рекомендациями, описанными в [77; 79], должны были «преследоваться по закону»(!).

Применение ГОСТ 11.006—74 (СТ СЭВ 1190—78) в отраслях народного хозяйства приносило вред. Примеры были приведены, в частности, специалистами НИИ резиновой промышленности, входившими в состав рабочей группы. Из 30 выборок показателей качества при проверке по ГОСТ 11.006—74 (СТ СЭВ 1190—78) лишь 2 не были признаны нормально распределенными, а при применении корректных методов негауссовскими оказались не 2, а подавляющее большинство, а именно 25. Соответственно их дальнейшая обработка пошла по непараметрическому пути, т.е. совсем не так, как вытекало из применения ГОСТ 11.006—74 (СТ СЭВ 1190—78).

В качестве второго примера рассмотрим терминологический стандарт ГОСТ 15895—77 (СТ СЭВ 547—77, СТ СЭВ 3404—81) «Статистические методы управления качеством продукции. Термины и определения» (он не был отменен, несмотря на отрицательное заключение рабочей группы).

Этот стандарт содержит огромное количество грубейших ошибок. Достойно удивления и сожаления, что подготовленные на столь низком научно-техническом уровне документы, как два рассматриваемых стандарта, оказались утвержденными не только в СССР (и затем в России), но и в рамках международной стандартизации (в рамках

СЭВ). Не подействовали ни заключения ведущих специалистов, ни научные публикации. Только общие решения по переводу подобных стандартов на уровень рекомендательных документов избавил академиков и профессоров — авторов учебников по теории вероятностей и математической статистике, по статистическим методам и эконометрике — от «преследования по закону» (!) за использование определений и обозначений, отличающихся от включенных в рассматриваемый стандарт. Наличие подобных «стандартов» — одна из причин появления терминологического «Приложения 1» к учебнику [89]. В свое время этот текст (опубликован в [176, приложение 1]) был разработан взамен негодного стандарта СТ СЭВ 3404—81, однако ввести его в качестве стандарта не удалось.

Как избежать ошибок в нормативно-технической 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 и не должен этим заниматься. С другой стороны, математику (специалисту по статистическим методам) 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 появлении ошибок лица сделали все, чтобы помешать признанию и исправлению допущенных ими ошибок в государственных стандартах. До сих пор продолжаются попытки навязать промышленности (в качестве нормативных и методических документов, в том числе международных стандартов!) 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 и вообще достижений современной статистической науки, а также организацию борьбы с ошибками.

От экспертизы — к научным разработкам и их внедрению. 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 1600 с.) и на настоящий момент является наиболее полным руководством по рассматриваемой тематике. Информация о нем приложена к изданному массовым тиражом переводу книги японских авторов по аналогичной тематике [118].

К сожалению, Госстандарт не пожелал финансировать реализацию заказанного им «Аванпроекта». Тогда решено было действовать самостоятельно. В 1989 г. нами был организован Центр статистических методов и информатики (ЦСМИ; в настоящее время — Институт высоких статистических технологий и эконометрики МГТУ им. Н.Э. Баумана). 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 [68]). В работе участвовали 128 специалистов. В дальнейшем к ЦСМИ присоединялись новые группы научно-технических работников, уже к концу 1991 г. нас было более 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, использовались более чем в 100 организациях и на предприятиях. Среди них — производственные объединения: «Уралмаш», «АвтоВАЗ», «Пластик»; научно-исследовательские институты: 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; вузы: МИИТ, Казахский политехнический институт, Ульяновский политехнический институт, Донецкий государственный университет и др.

Параллельно ЦСМИ вел работу по объединению статистиков и эконометриков. 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 [89, глава 13]). В 1992 г. после развала СССР и фактического прекращения работы ВСА на основе секции статистических методов ВСА была организована Российская ассоциация по статистическим методам (РАСМ), а затем и Российская академия статистических методов, действующие и в настоящее время. В мероприятиях секции статистических методов ВСА и РАСМ активно участвовали несколько сот специалистов по статистическим методам и эконометрике. А одной из основных тематик этих специалистов являются, как следует из сказанного, статистические методы в сертификации (управлении качеством). В ЦСМИ и РАСМ, объединивших большинство ведущих российских специалистов, коллективными усилиями разработан единый подход к проблемам применения статистических методов в сертификации и управлении качеством, отраженный, в частности, в учебнике [89].

Ранее разработанные нормативно-техническая и методическая документация, диалоговые компьютерные системы по статистическим методам продолжают использоваться, несмотря на политические преобразования. 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, чтобы никто не мог навязать стране свои ошибки в качестве стандартов, проигнорировав протесты ведущих специалистов. При этом условии внедрение современных эконометрических методов сертификации и управления качеством продукции может дать нашей стране экономический эффект, измеряемый миллиардами долларов США в год [89].

14.5. ЭКСПЕРТНЫЕ ОЦЕНКИ В ОЦЕНОЧНОЙ ДЕЯТЕЛЬНОСТИ И ИНВЕСТИЦИОННОМ МЕНЕДЖМЕНТЕ

Необходимость оценки стоимости бизнеса, инвестиций, недвижимости и других активов актуальны как для организаций, так и для физических лиц [120]. Органы государственной власти заинтересованы, в частности, в подготовке нормативных и методических документов по проведению массовой оценки стоимости недвижимого имущества для целей налогообложения. Для простоты речи будем говорить об оценке стоимости предприятия (бизнеса). Для физических лиц важна оценка стоимости недвижимости, например квартиры.

Обычно выделяют три подхода к оценке стоимости предприятия — затратный, доходный и сравнительный. Проводят и комплексную оценку, в частности на основе соответствующего обобщенного показателя (рейтинга) [120].

Затратный подход состоит в оценке стоимости активов предприятия на основе бухгалтерской отчетности и анализа его финансово-хозяйственной деятельности. Используют метод накопления активов (т.е. расчета балансовой стоимости активов), метод расчета восстановительной стоимости активов, стоимости замещения, ликвидационной стоимости, показатель «чистые активы». Затратный подход обращен в прошлое, основан на анализе истории экономической деятельности предприятия. Он активно использует индивидуальные экспертные оценки, в частности при сопоставлении результатов применения конкретных методов оценки, учете инфляции, оценке стоимости производственных запасов [137].

Доходный подход предполагает оценку будущих доходов, которые может принести данное предприятие. Можно сказать, что речь идет об оценке экономической эффективности инвестиций в покупку предприятия. Оценка стоимости предприятия при доходном подходе — это тот максимальный объем инвестиций, при котором они еще остаются экономически выгодными. С помощью экспертов составляют сценарий развития предприятия (бизнес-план), рассчитывают соответствующий финансовый поток и его характеристики, в частности чистую текущую стоимость *NPV* (для чего экспертно оценивают приемлемый для инвестора уровень доходности, т.е. коэффициент дисконтирования), анализируют и оценивают риски, разрабатывают планы управления рисками. Доходный подход основан на прогнозе будущего развития предприятия и его рыночного окружения, а такой прогноз предполагает интенсивное использование современных технологий экспертного прогнозирования.

Сравнительный подход называют также рыночным. Выделяют совокупность аналогичных предприятий, проданных в последнее время (решают, какие предприятия считать сравнимыми), и анализируют рыночные цены, зафиксированные в актах их купли-продажи. Сравнительный подход предполагает движение не по оси времени (для затратного подхода — в прошлое, для доходного — в будущее), а в пространстве — анализ стоимостей аналогичных предприятий в тот момент времени, когда проводится оценка. Необходимо предсказать цену, за которую будет продано оцениваемое предприятие. Поскольку предприятия отличаются друг от друга, то с помощью экспертов проводят соответствующие исследования:

1) выделяют совокупность аналогичных объектов оценки, как теоретическую генеральную совокупность, так и обучающую выборку (здесь использованы термины прикладной статистики [77]);
2) выделяют совокупность характеристик, описывающих предприятия из рассматриваемой генеральной совокупности, и формируют информационную базу для решения задачи оценки стоимости предприятия;
3) по эмпирическим данным восстанавливают функциональную зависимость рыночной стоимости (цены) от характеристик предприятий;
4) на основе характеристик оцениваемого предприятия рассчитывают прогнозируемое значение его рыночной стоимости.

Например, при разработке алгоритмов оценивания рыночной стоимости жилых квартир в качестве генеральной совокупности естественно взять все квартиры определенного региона (например, Москвы). В качестве обучающей выборки не менее естественно взять доступную информацию о продажах за последнюю единицу времени (месяц или год). Проблема сбора такой информации нетривиальна. Может оказаться доступной лишь цена продавца, как правило превышающая цену сделки. Перечень характеристик объектов оценки (квартир) достаточно очевиден:

1) общая площадь (в квадратных метрах);
2) жилая площадь;
3) число комнат;
4) тип дома (панельный; кирпичный и т.п.);
5) этаж (первый; последний; иные);
6) оценка привлекательности (престижности) района;
7) оценка экологической безопасности;

8) оценка текущего состояния и степени износа квартиры и дома;
9) оценка инфраструктуры и т.д.

С точки зрения прикладной статистики [77] речь идет о регрессионном анализе разнотипных данных, входящем в статистику нечисловых данных. Прагматичный, но не вполне корректный подход состоит в оцифровке всех характеристик (т.е. введении баллов для характеристик, измеренных в качественных шкалах) с целью перехода к количественным признакам, а затем в применении стандартного метода наименьших квадратов. Методы восстановления зависимостей от разнотипных переменных продолжают развиваться [117], находя все большее применение в решении организационно-экономических задач [17].

При рассмотрении конкретных постановок задач оценки стоимости предприятий (бизнеса) или иных экономических единиц расчетные значения, полученные с помощью трех рассмотренных подходов (или двух, если какой-либо из подходов применить не удается), как правило, различаются, иногда на десятки процентов. Для получения единой оценки необходимо их усреднить с помощью того или иного метода агрегирования (другими словами, построения рейтинговой оценки, обобщающего показателя). Как уже не раз обсуждалось, весовые коэффициенты в таких случаях обычно определяют экспертным путем.

Роль экспертных оценок в инвестиционном менеджменте. Как уже отмечалось, доходный подход к оценке стоимости предприятия по своей организационно-экономической сути является частью инвестиционного менеджмента. Обсудим место и роль экспертных оценок в инвестиционном менеджменте. Напомним, что термин «инвестиции» в переводе на русский язык означает «капиталовложения». С точки зрения теории принятия решений, в том числе на основе интенсивного использования экспертных технологий, проблемы инвестиционного менеджмента рассмотрены в [50; 87].

Инвестирование представляет собой один из наиболее важных аспектов деятельности любой развивающейся организации. Причины, обусловливающие необходимость инвестиций, могут быть различными, однако в целом их можно подразделить на три вида:

- обновление имеющейся материально-технической базы;
- наращивание объемов производственной деятельности;
- освоение новых видов деятельности.

Любой инвестиционный проект может быть охарактеризован с различных сторон: финансовой, технологической, организационной, временно́й, экологической, социальной и др. Каждая из них по-своему важна, однако финансовые аспекты инвестиционной деятельности во многих случаях имеют решающее значение. Например, речь идет о реконструкции действующего предприятия или строительстве нового завода. Помимо финансовой выгодности, нельзя забывать, скажем, и о социальном окружении: с одной стороны, появятся новые рабочие места (легко ли будет их заполнить?), с другой — население может выступить против проекта, сочтя его экологически вредным. В соответствии с Законом РФ «Об экологической экспертизе» любая намечаемая хозяйственная или иная деятельность рассматривается как имеющая потенциальную экологическую опасность, а потому любая такая деятельность подлежит государственной экологической экспертизе (за счет заказчика). Только при ее положительном заключении разрешается финансирование и кредитование проекта.

Инвестиционные проекты разрабатывают не только частные структуры, но и государственные организации. Так, изменение налоговой системы — тоже инвестиционный проект.

С экономической точки зрения инвестиционные проекты описываются потоками платежей, т.е. функциями от времени, значениями которых являются сальдо поступлений и затрат за очередной интервал времени. Как правило, вначале необходимо вкладывать деньги (производить затраты), а затем за счет поступлений возмещать затраты и получать прибыль. Однако возможны и ситуации, когда завершение проекта (например, закрытие атомной электростанции и утилизация отработанного ядерного топлива) требует существенных вложений.

В конкретный промежуток времени обычно происходят как поступления, так и платежи. Как элемент финансового потока рассматривается итоговый результат — сальдо, т.е. поступления минус платежи. Этот результат может быть как положительным, так и отрицательным.

Для различных вариантов управляющих воздействий на процессы налогообложения (например, различных вариантов изменения ставок налогов) при сравнении их с действующей системой ситуация аналогична. Если в результате управляющих воздействий налоговые сборы в некоторый момент меньше тех, что при действующей системе, то элемент финансового потока является отрицательным (приращение поступлений отрицательно), в противном случае — положительным (приращение налоговых поступлений положительно). Для любого управляющего воздействия часть поступлений оказывается отрицательной, часть — положительной, и проблема состоит в их соизмерении, поскольку они относятся к различным моментам времени.

Ясно, что финансовый поток инвестиционного проекта определяется разработанным в результате экспертного исследования сценарием развития проекта.

В финансовом плане, когда речь идет о целесообразности принятия того или иного инвестиционного проекта, необходимо получить ответы на три вопроса:

а) каков необходимый объем финансовых ресурсов;
б) где найти источники финансирования (кредитования) в требуемом объеме и какова цена их услуг;
в) окупятся ли сделанные вложения, т.е. достаточен ли объем прогнозируемых поступлений по сравнению со сделанными инвестициями?

Ответ на первый вопрос определяется инженерной сутью проекта и выражается в виде финансового потока, обоснованного в бизнес-плане, разработанном на основе сценария, полученного в результате экспертного исследования. Ответ на второй вопрос зависит от конкретной ситуации на финансовом рынке, для оценки которой необходимы эксперты. Для ответа на третий вопрос необходимо от финансового потока как функции времени перейти к той или иной его обобщенной характеристике. Такой переход целесообразен также при сравнении различных проектов.

Отметим, что как при изменении налоговой системы путем варьирования значений управляющих параметров, так и при реализации иных крупных инвестиционных проектов меняются также и значения социальных, технологических, экономических, экологических и политических факторов (сокращенно, СТЭЭП-факторов). Например, строятся или приходят в упадок дороги, создаются или сокращаются рабочие места и т.д. Другими словами, оценку управляющих воздействий на процессы налогообложения, как и крупных инвестиционных проектов, нельзя проводить только с экономической точки зрения, должен учитываться весь комплекс СТЭЭП-факторов. При этом, очевидно, необходимо применять разнообразные процедуры экспертных оценок для комплексного учета СТЭЭП-факторов, нельзя опираться лишь на чисто экономические расчеты.

Обсудим подходы к сравнению инвестиционных проектов (и оценке управляющих воздействий на процессы налогообложения). Прежде всего отметим, что сравнение инвестиционных проектов — это сравнение функций от времени. Кроме того, имеется внешняя среда, которая проявляется в виде дисконт-функции как результата воздействия СТЭЭП-факторов и представлений законодателя или инвестора. Эти априорные представления проявляются в основном в виде ограничений на потоки платежей (в частности, могут быть заданы ограничения на объем кредитов или налогов) и на горизонт планирования, рассматриваемый лицом или лицами, принимающими решения (законодателями, работниками государственных органов, занимающихся проблемами налогов и сборов, или инвестором).

Одна из основных проблем при сравнении инвестиционных проектов такова: что лучше — меньше, но сейчас, или больше, но потом? Например, существенно увеличив сбор налогов сейчас, мы уменьшим рост производства и, следовательно, в дальнейшем из-за уменьшившейся налоговой базы будем собирать налогов меньше, чем в ситуации, когда мы вначале сократим ставки налогов, что даст быстрый рост производства и налоговой базы, и сборы в бюджет возрастут, но потом, а не сейчас. Похожая ситуация описана в пословице: лучше синица в руках, чем журавль в небе.

Та же проблема возникает при сравнении инвестиционных проектов, рассматриваемых частным инвестором. Как правило, чем больше вкладываем сейчас, тем больше получаем в более или менее отдаленном будущем. Вопрос в том, достаточны ли будущие поступления, чтобы покрыть нынешние платежи и обеспечить приемлемую прибыль?

Выбирая для реализации тот или иной инвестиционный проект, как и выбирая тот или иной вариант налоговой политики, те или иные управляющие воздействия на процесс налогообложения, мы сравниваем потоки платежей. С одной стороны, ситуация с частными инвестиционными проектами проще, поскольку мы можем существенно более точно предсказать моменты и размеры будущих поступлений и платежей для конкретного проекта, чем в случае системы налогообложения, охватывающей всех юридических и физических лиц. С другой стороны, будущие налоговые сборы должны учитываться при оценке эффективности инвестиционных проектов.

Среди экономических характеристик (критериев) инвестиционных проектов один из основных — ***чистая текущая стоимость***. Этот критерий основан на сопоставлении величины исходных инвестиций (IC) с общей суммой дисконтированных чистых денежных поступлений, генерируемых проектом в течение прогнозируемого срока. Поскольку приток денежных средств распределен во времени, он дисконтируется с помощью коэффициента q. Выбор значения этого коэффициента может осуществляться из различных соображений. Например, он может быть установлен аналитиком (выступающим от имени инвестора), исходя из ежегодного процента возврата, который инвестор хочет или может иметь на инвестируемый им капитал.

Допустим, делается прогноз, что исходные инвестиции (IC) будут генерировать в течение n лет годовые доходы в размере $P_1, P_2, \ldots, P_n$. Общая накопленная *величина дисконтированных доходов* (*Present Value, PV*) и *чистая текущая стоимость* (*Net Present Value, NPV*) соответственно рассчитываются по формулам:

$$PV = \sum_{k=1}^{n} \frac{P_k}{(1+q)^k}; \quad NPV = PV - IC.$$

Очевидно, что если $NPV > 0$, то проект экономически выгоден; если $NPV < 0$, то проект экономически невыгоден и его целесообразно отвергнуть; при $NPV = 0$ проект не является ни прибыльным, ни убыточным.

Разработано много методов определения коэффициента дисконтирования q. Выбор среди них — дело экспертов. Ряд методов предусматривает использование экспертов. Кроме того, очевидно, что коэффициент дисконтирования должен меняться со временем, поскольку меняются определяющие его ставка рефинансирования, индекс инфляции, риски хозяйственных операций. Необходимо оценивать погрешности характеристик инвестиционного проекта. В [3; 77] разработан метод изучения отклонений NPV в зависимости от допустимых отклонений погодовых коэффициентов дисконтирования. При этом максимально допустимое отклонение коэффициентов дисконтирования определяется в результате экспертного исследования. И риски инвестиционного проекта оцениваются прежде всего экспертами.

Необходимость применения экспертных оценок при сравнении инвестиционных проектов. Из сказанного ранее и анализа ситуации, проведенного в [87], вытекает, что разнообразные формальные методы оценки характеристик и рисков инвестиционных проектов во многих случаях (реально — во всех нетривиальных ситуациях) не могут дать однозначных рекомендаций.

Поэтому процедуры экспертного оценивания нужно применять не только на заключительном этапе — при принятии решения о целесообразности реализации, но и на всех остальных этапах анализа инвестиционного проекта. При этом необходимо использовать весь арсенал теории и практики экспертных оценок, весьма развитой области научной и практической деятельности. В конце процесса принятия решения — всегда человек.

Мы не призываем отказаться от формально-экономических методов. Вычисление чистой текущей стоимости и других характеристик финансовых потоков, использование соответствующих программных продуктов полезно для принятия обоснованных решений. Однако нельзя абсолютизировать формально-экономические методы. Например, на основной вопрос: что лучше — быстро, но мало, или долго, но много — ответить могут только эксперты.

Поэтому система поддержки принятия решений в области управления инвестициями, а также, например, в совершенствовании налогообложения (в том числе оценки управляющих воздействий на процессы налогообложения) должна сочетать формально-экономические и экспертные процедуры.

14.6. ПРОГНОЗИРОВАНИЕ И МЕТОД СЦЕНАРИЕВ

В главе 7 были рассмотрены основные идеи одной из самых популярных экспертных технологий — технологии сценарных экспертных прогнозов. В предыдущем подразделе показана роль этих технологий в оценочной деятельности и при управлении инвестициями.

Академия прогнозирования. В нашей стране исследования по прогнозированию ведут многие организации и отдельные специалисты. Среди них выделяется общественная Академия прогнозирования (исследований будущего). Она является одной из дочерних организаций Всемирной федерации исследований будущего (статус II общественных организаций ООН и статус «В» консультативных организаций ЮНЕСКО). Академия прогнозирования создана по инициативе пяти ассоциаций и центров и объединяет специалистов, которые изучают перспективы развития различных процессов и явлений.

Академия прогнозирования опирается в своей деятельности на концепцию технологического прогнозирования (проблемно-целевой подход в исследованиях будущего), сформулированную в 1924—1927 гг. В.А. Базаровым-Рудневым и затем, независимо от него, в конце 1950 — и начале 1960-х гг. О. Гелмером, Т. Гордоном, Д. Беллом, Б. Де Жувенелем и другими футурологами. Академия развивает прогностические традиции, заложенные в трудах Н.Д. Кондратьева, В.И. Вернадского, К.Э. Циолковского, А.П. Сорокина, Л.А. Чижевского, А.Н. Колмогорова, Н. Винера и др.

В России были предприняты три попытки создания общественной (внегосударственной) научной организации теоретиков и практиков прогнозирования. Первая попытка состоялась в 1966—1970 гг., когда работала Советская ассоциация научного прогнозирования, которая ежемесячно собирала семинары с сотнями участников, выпускала ряд изданий и т.д. Она была распущена в 1971 г. Вторая попытка началась с неофициальных семинаров в Московском авиационном институте, что привело к созданию в 1976 г. Комиссии по научно-техническому прогнозированию в одном из комитетов Всесоюзного Совета научно-технических обществ (ВСНТО). В 1979 г. комиссия была реорганизована в Комитет по научно-техническому прогнозированию и разработке комплексных программ научно-технического прогресса ВСНТО (затем Совета научных и инженерных обществ СССР), в состав которого входило более десятка комиссий: по теории и методологии прогнозирования, по социальным, экономическим, экологическим, глобальным проблемам прогнозирования. Формально Комитет существует до сих пор, а фактически парализован крушением СССР.

Третья попытка началась в 1988—1990 гг. с семинаров Ассоциации содействия Всемирной федерации исследований будущего (президент И.В. Бестужев-Лада, академик РАО). Академия прогнозирования (исследований будущего) создана в апреле 1997 г. как межрегиональная и междисциплинарная общественная организация, объединяющая специалистов, которые изучают перспективы развития различных процессов и явлений. Она имеет целью содействовать обмену информацией, представляющей интерес для ее членов, подготовке и переподготовке прогнозистов, координировать деятельность в сфере прогнозирования. В сентябре 1999 г. Академия прогнозирования (исследований будущего) стала институциональным членом Международной академии исследований будущего (International Future Research Academy). Сегодня Академия прогнозирования работает по 40 основным направлениям научного прогнозирования и представляет собой кооперацию специалистов в сфере прогнозирования.

Основные методы прогнозирования — статистические и экспертные.

Статистические методы прогнозирования. Основными задачами, решаемыми с помощью этих методов, являются:

- разработка, изучение и применение современного математико-статистического прогнозирования на основе объективных данных (в том числе непараметрических методов наименьших квадратов с оцениванием точности прогноза, адаптивных методов, методов авторегрессии и др.);

- развитие теории и практики вероятностно-статистического моделирования экспертных методов прогнозирования (в том числе методов анализа субъективных экспертных оценок на основе статистики нечисловых данных; методов прогнозирования в условиях риска и комбинированных методов прогнозирования с использованием совместно экономико-математических и эконометрических (как математико-статистических, так и экспертных) моделей).

Научной базой статистических методов прогнозирования является прикладная статистика [77] и теория принятия решений [87].

Простейшие методы восстановления используемых для прогнозирования зависимостей исходят из заданного временно́го ряда, т.е. функции, определенной в конечном числе точек на оси времени. Временной ряд при этом часто рассматривается в рамках той или иной вероятностной модели, вводятся другие факторы (независимые переменные), помимо времени, например, объем денежной массы. Временной ряд может быть многомерным. Основные решаемые задачи — интерполяция и экстраполяция. Метод наименьших квадратов в простейшем случае (линейная функция от одного фактора) был разработан К. Гауссом в 1794—1795 гг. Могут оказаться полезными предварительные преобразования переменных, например логарифмирование. Наиболее часто используется метод наименьших квадратов при нескольких факторах. Метод наименьших модулей, сплайны и другие методы экстраполяции применяются реже, хотя их статистические свойства зачастую лучше.

Опыт прогнозирования индекса инфляции и стоимости потребительской корзины накоплен в Институте высоких статистических технологий и эконометрики [89]. Оказалось полезным преобразование (логарифмирование) переменной — текущего индекса инфляции. При стабильности условий точность прогнозирования оказывалась достаточно удовлетворительной — 10—15%. Однако спрогнозированное на осень 1996 г. значительное повышение уровня цен не осуществилось. Причина — руководство страны перешло к стратегии сдерживания роста потребительских цен путем массовой невыплаты долгов юридическим и физическим лицам. Условия изменились — и статистический прогноз оказался непригодным. Другой пример: влияние решений руководства Москвы проявилось в том, что в ноябре 1995 г. (перед парламентскими выборами) цены в Москве упали в среднем на 9,5%, хотя обычно для ноября характерен более быстрый рост цен, чем в другие месяцы года, кроме декабря и января.

Оценивание точности прогноза (в частности, с помощью доверительных интервалов) — необходимая часть процедуры прогнозирования. Обычно используют вероятностно-статистические модели восстановления зависимости, например, строят наилучший прогноз по методу максимального правдоподобия. Разработаны параметрические (обычно на основе модели нормальных ошибок) и непараметрические оценки точности прогноза и доверительные границы для него (на основе Центральной предельной теоремы теории вероятностей). Так, предложены и изучены непараметрические методы доверительного оценивания точки наложения (встречи) двух временны́х рядов и их применения для оценки динамики технического уровня собственной продукции и продукции конкурентов, представленной на мировом рынке. Применяются также эвристические приемы, не основанные на вероятностно-статистической теории: метод скользящих средних, метод экспоненциального сглаживания.

Многомерная регрессия, в том числе с использованием непараметрических оценок плотности распределения, — основной на настоящий момент статистический аппарат прогнозирования. Подчеркнем, что нереалистическое предположение о нормальности погрешностей измерений и отклонений от линии (поверхности) регрессии использовать не обязательно. Однако для отказа от предположения нормальности необходимо опереться на иной математический аппарат, основанный на многомерной Центральной предельной теореме теории вероятностей, технологии линеаризации и наследования сходимости [88]. Он позволяет проводить точечное и интервальное оценивание параметров, проверять значимость их отличия от 0 в непараметрической постановке, строить доверительные границы для прогноза.

Весьма важна проблема проверки адекватности модели, а также проблема отбора факторов. Априорный список факторов, оказывающих влияние на отклик, обычно весьма обширен (желательно его сократить) и крупное направление современных исследований посвящено методам отбора «информативного множества признаков». Однако эта проблема пока еще окончательно не решена. Проявляются необычные эффекты. Так, установлено, что обычно используемые оценки степени полинома имеют в асимптотике геометрическое распределение [77, 89]. Перспективны непараметрические методы оценивания плотности вероятности и их применения для восстановления регрессионной зависимости произвольного вида. Наиболее общие результаты в этой области получены с помощью подходов статистики нечисловых данных.

К современным статистическим методам прогнозирования относятся также модели авторегрессии, модель Бокса — Дженкинса, системы эконометрических уравнений, основанные как на параметрических, так и на непараметрических подходах.

Для установления возможности применения асимптотических результатов при конечных (так называемых «малых») объемах выборок полезны компьютерные статистические технологии. Они позволяют также строить различные имитационные модели. Отметим полезность методов размножения данных (бутстреп-методов). Системы прогнозирования с интенсивным использованием компьютеров объединяют различные методы прогнозирования в рамках единого автоматизированного рабочего места прогнозиста.

Прогнозирование на основе данных, имеющих нечисловую природу, в частности прогнозирование качественных признаков, основано на результатах статистики нечисловых данных. Весьма перспективными для прогнозирования представляются регрессионный анализ на основе интервальных данных, включающий, в частности, определение и расчет нотны и рационального объема выборки, а также регрессионный анализ нечетких данных, разработанный в [70]. Общая постановка [77] регрессионного анализа в рамках статистики нечисловых данных и ее частные случаи — дисперсионный анализ и дискриминантный анализ (распознавание образов с учителем), давая единый подход к формально различным методам, полезны при программной реализации современных статистических методов прогнозирования.

Экспертные методы прогнозирования. Основными процедурами обработки прогностических экспертных оценок являются проверка согласованности, кластер-анализ и нахождение группового мнения. Проверка согласованности мнений экспертов, выраженных ранжировками, проводится с помощью коэффициентов ранговой корреляции Кендалла и Спирмена, коэффициента ранговой конкордации Кендалла и Бэбингтона Смита. Используются параметрические модели парных сравнений — Терстоуна, Бредли — Терри — Льюса — и непараметрические модели теории люсианов [77; 89]. Полезна процедура согласования ранжировок и классификаций путем построения согласующих бинарных отношений. При отсутствии согласованности разбиение на группы мнений экспертов, сходных между собой, проводят методом ближайшего соседа или другими методами кластерного анализа (автоматического построения классификаций, распознавания образов без учителя). Классификация люсианов осуществляется на основе вероятностно-статистической модели.

Используют различные методы построения итогового мнения комиссии экспертов. Своей простотой выделяются методы средних арифметических и медиан рангов. Компьютерное моделирование [89] позволило установить ряд свойств медианы Кемени, часто рекомендуемой для использования в качестве итогового (обобщенного, среднего) мнения комиссии экспертов. Интерпретация закона больших чисел для нечисловых данных в терминах теории экспертного опроса такова: итоговое мнение устойчиво, т.е. мало меняется при изменении состава экспертной комиссии, и при росте числа экспертов приближается к «истине». При этом в соответствии с принятым в [88] подходом предполагается, что ответы экспертов можно рассматривать как результаты измерений с ошибками, все они — независимые одинаково распределенные случайные элементы, вероятность принятия определенного значения убывает по мере удаления от некоторого центра — «истины», а общее число экспертов достаточно велико.

Многочисленны примеры ситуаций, связанных с социальными, технологическими, экономическими, политическими, экологическими и другими рисками. Именно в таких ситуациях обычно и необходимо прогнозирование. Известны различные виды критериев, используемых в теории принятия решений [87] в условиях неопределенности (риска). Из-за противоречивости решений, получаемых по различным критериям, очевидна необходимость применения оценок экспертов.

В конкретных задачах прогнозирования необходимо провести классификацию рисков, поставить задачу оценивания конкретного риска, провести структуризацию риска, в частности построить деревья причин (в другой терминологии, деревья отказов) и деревья последствий (деревья событий). Центральной задачей является построение групповых и обобщенных показателей, например показателей конкурентоспособности и качества. Риски необходимо учитывать при прогнозировании экономических последствий принимаемых решений, поведения потребителей и конкурентного окружения, внешнеэкономических условий и макроэкономического развития России, экологического состояния окружающей среды, безопасности технологий, экологической опасности промышленных и иных объектов.

Современные компьютерные технологии прогнозирования основаны на интерактивных статистических методах прогнозирования с использованием баз эконометрических данных, имитационных (в том числе на основе применения метода статистических испытаний) и экономико-математических динамических моделей, сочетающих экспертные, математико-статистические и моделирующие блоки.

Литература по статистическим методам прогнозирования весьма обширна (содержит не менее 100 тыс. статей и книг). В данной главе процитированы источники, в которых информация приведена в единую систему в соответствии с рекомендациями Российской академии статистических методов и Международной академии исследований будущего.

Технологии экспертного прогнозирования, прежде всего методом сценариев, систематически рассмотрены в [115]. Разработке конкретных прогнозов методом сценариев посвящены работы [69; 85; 86], методологические аспекты которых продолжают оставаться актуальными.

В сценарном подходе мы формулируем исходные предположения и затем отвечаем на вопрос: «Что будет, если... ?» В качестве примера рассмотрим сюжет, посвященный последствиям перехода к схеме «открытой торговли».

Нобелевский лауреат по экономике Пол Самуэльсон в своем учебнике [110] пишет: «Беспрепятственно осуществляемая торговля способствует взаимовыгодному международному разделению труда, в большой степени увеличивает потенциально реальный национальный продукт всех стран и создает возможность повышения уровня жизни на всем земном шаре». Конечно, СССР и Россия всегда были на мировом рынке, всегда торговали с зарубежными странами, но сейчас речь об ином — о снятии таможенных барьеров.

Посмотрим, что будет с Россией, если мы войдем в мировой рынок (в смысле П. Самуэльсона), т.е. примем безграничную свободу торговли, уберем все таможенные барьеры.

Начнем с прогноза развития сельского хозяйства. У нас климат — не лучший для земледелия, длительные зимы и длинные дороги, маленькая плотность населения, сравнительно (с Францией и США) низкая урожайность. Невыгодно сажать пшеницу и разводить коров, дешевле купить за рубежом. И с этим нам придется смириться — такова география России.

Поэтому на нашем рынке западные продовольственные товары будут дешевле отечественных. Неконкурентоспособные фирмы должны исчезнуть — таков закон «свободного рынка». Значит, сельское хозяйство России обречено.

Сначала разорятся хозяйства, поставляющие продукты на рынок, крестьяне станут безработными. Затем наиболее активные из них уедут из деревни, а оставшиеся перейдут к натуральному хозяйству на своих приусадебных участках. Отомрет инфраструктура — магазины, связь,

школы, больницы. Шансы выжить есть лишь у сельскохозяйственного товаропроизводителя юга России — на Кубани, в Ставрополье. А в Центральной России все больше будет заброшенных полей.

В чем же мы конкурентоспособны? Еще совсем недавно мы производили половину военной техники, выпускаемой в мире. Но «холодная война» кончилась, ракеты уничтожают. Разве будут США, Германия и Япония поддерживать развитие нашей оборонной промышленности? Наш идеал — мир на Земле, а потому оборонная промышленность обречена.

Может быть, выход в конверсии? Но совсем не так легко перейти от выпуска танков к производству тракторов, как казалось 15—20 лет назад. Нужны инвестиции, и немалые. Оправдаются ли они для мировой экономики? И тут опять придется вспомнить о российской зиме, о необходимости строительства заводских корпусов и их отоплении, о длинных русских дорогах. Дешевле построить и содержать завод в Алжире, чем в Архангельске, и никуда от этого не уйти. Наш климат требует дополнительных вложений в 2000 дол. в год на одно рабочее место по сравнению с Алжиром и Китаем.

Зачем американцам и европейцам переносить современные технологии в Россию, да и вообще давать работу русским, когда у себя — безработица? Да они, русские, и английского языка не знают, и работать под надзором западных менеджеров не умеют. Возможно, и не захотят.

Короче, будущее российской промышленности в условиях свободного рынка почти столь же мрачно, как и будущее нашей деревни. Большая ее часть обречена на поражение в конкурентной борьбе ближайших десятилетий, а потому — на уничтожение.

В чем же Россия имеет преимущества по сравнению с другими странами? В производстве сырья — нефти, газа, леса, стали, алюминия... И сырьевые отрасли войдут (уже вошли!) в мировое хозяйство без границ. Правда, до тех пор, пока добыча нефти остается экономически выгодной, т.е., по разным оценкам, до 2010—2015 г.

Большие пространства России хороши для создания мировых хранилищ вредных отходов, в том числе радиоактивных, да и попросту мусора. В этом виде «деятельности» у нас мало конкурентов.

Итак, сельское хозяйство и почти вся промышленность обречены. Сколько-то рабочих мест останется — охранники банков и продавцы продуктов, могильщики и мэры всегда будут нужны. Но бóльшая часть населения лишится работы. Как ни странно может показаться, для мирового сообщества дешевле кормить и одевать (в поношенную на Западе одежду) безработных россиян, чем дать им возможность трудиться на своих неконкурентоспособных предприятиях.

И тогда проявится еще одно преимущество России перед другими странами — большое число дешевых рабочих рук. Они найдут применение во всем мире. Возможно, к 2030 г. начальные школы Сомали и Чада будут укомплектованы русскими учителями. Такие рабочие места уже предлагаются.

Что будет дальше? Население, лишаясь нормальных условий для жизни, будет уходить не только из деревень, но и с северных территорий. Этот процесс уже сейчас активно идет. При добыче сырья придется перейти на вахтовый метод. При отсутствии отечественной промышленности сырьевые отрасли будут насыщаться зарубежной техникой, а потому и вахты будут состоять в основном из иностранцев.

Примерно к 2050 г. Россия превратится в территорию, покрытую запретными зонами, окружающими свалки опасных отходов, на которой в разрушающихся от старости городах еще живут несколько десятков миллионов читателей настоящего учебника, наших детей и внуков, которых из гуманных соображений кормят и одевают миссии ООН (на кредиты Международного валютного фонда и Всемирного банка). Конечно, детей учат в ООНовских школах. Однако на каком языке? Поскольку деньги на учебу дает мировой капитал, то и занятия идут на английском языке. Родители не возражают — ведь работу их дети смогут найти лишь за пределами России. Но на каком языке учат, на каком разговаривают на работе и после работы, на том начинают и думать. Наши внуки и правнуки будут думать на английском языке.

Итак, к 2050 г. Россия превратится в огромную резервацию типа тех, что правительство США содержит для индейских племен. Существенная часть пока еще нашей земли будет занята мировыми свалками или отторгнута сырьевыми компаниями с иностранным персоналом. На оставшейся территории мы сможем жить «привычной» жизнью, питаясь подаяниями мирового сообщества.

Вот что будет, если Россия войдет в мировой рынок, вооружась лозунгом свободной торговли. Конкретные политические действия и экономические решения могут несколько замедлить или убыстрить общий ход событий, но не могут изменить финальное состояние, если только не будет поставлено пределов свободе торговли и мировому рынку.

То, что вы прочитали, — не фантастическая антиутопия, это — научный прогноз на основе западной экономической теории. Применен метод сценариев из теории экспертных оценок.

Обсудим результаты экспертного прогнозирования. Конечно, такого ужаса не будет. Любое мало-мальски разумное правительство России будет защищать отечественного товаропроизводителя. Но под влиянием постоянной пропаганды средств массовой дезинформации в 1990-е г. лозунги «свободы торговли» и «ликвидации неконкурентоспособных предприятий» стали догмами массового сознания, многие россияне приняли их без критики. Поэтому необходимо продемонстрировать, к чему они, эти лозунги, ведут. Впервые это было сделано в рамках исследований так называемого «Римского клуба», объединяющего западных ученых и промышленников еще в начале 1980-х гг.

Итак, что будет, если Россия откажется от защиты отечественных товаропроизводителей, снимет все таможенные барьеры и войдет в мировой рынок? Погибнут как неконкурентоспособные сельское хозяйство и бóльшая часть промышленности. Выживут сырьевые отрасли и свалки для всего мира, а также организации по подготовке рабочей силы для работы за рубежом. Обучение будет проходить на английском языке. Россия как самостоятельная страна погибнет. Вывод из этой антиутопии — необходимость защиты отечественных товаропроизводителей, в том числе путем установления таможенных барьеров.

Технологии экспертного прогнозирования, как и другие экспертные методы, необходимы широким кругам современных экономистов, управленцев (менеджеров), инженеров, специалистов практически всех отраслей народного хозяйства, самых разных направлений деятельности. Недаром они постоянно используются при рассмотрении вопросов организации и экономики производства [93; 137], как это подробно показано в работе [93].

Контрольные вопросы

1. Как с помощью экспертных оценок выявляют предпочтения потребителей в области технико-функциональных характеристик изделия (навигационного прибора)?
2. Какие экспертные технологии используются в системе «Шесть сигм»?
3. Что представляют собой иерархические системы эстетических и эргономических показателей качества?
4. В чем причина появления грубых ошибок в государственных стандартах по статистическим методам управления качеством?
5. Какова роль экспертных оценок в оценочной деятельности?
6. Почему необходимо опираться на мнения экспертов в инвестиционном менеджменте?
7. Как соотносятся статистические и экспертные методы прогнозирования?

Темы докладов и рефератов

1. Экспертные технологии оценки эффективности рекламы.
2. Метод фокус-групп.
3. Использование экспертных технологий при разработке управленческих решений.
4. Перспективы применения системы «Шесть сигм» в России.
5. Система «Шесть сигм» как образец для организации внедрения контроллинга и современных математических методов исследования (на основе работ [66; 127]).
6. Использование экспертных технологий для управления качеством.
7. Иерархическая система показателей надежности (на основе монографии [129]).
8. Технический уровень и конкурентоспособность продукции.
9. Методы построения обобщенных показателей технического уровня и качества (проблема агрегирования).
10. Научно-организационное развитие в области статистических методов (1985—1990) — от «Рабочей группы по упорядочению системы стандартов по прикладной статистике и другим статистическим методам» до Всесоюзной статистической ассоциации.
11. Оценочная деятельность в России (с использованием портала «Вестник оценщика», URL: http://www.appraiser.ru).
12. Соотношение экспертных и расчетных методов в инвестиционном менеджменте.
13. Система сценариев социально-экономического развития России (на основе работ [85; 86]).
14. Разработки Международной академии исследований будущего (URL: http://www.maib.ru/).

Часть IV

МОДЕЛИРОВАНИЕ В ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

ГЛАВА 15
ОСНОВЫ МОДЕЛИРОВАНИЯ

15.1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ МОДЕЛИРОВАНИЯ

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 [62, с. 44]. Для теории принятия решений наиболее полезны модели, которые выражаются словами или формулами, алгоритмами и иными математическими средствами.

Пример словесной модели [2]. Обсудим необходимость учета эффекта лояльности при управлении организацией в современных условиях. Под *лояльностью* понимается честное, добросовестное отношение к чему-либо или к кому-либо. Базу менеджмента, основанного на лояльности, заложил в 1908 г. профессор Гарварда Джошуа Ройс — автор книги «Философия лояльности», где впервые научно определено понятие «лояльность».

В рамках предлагаемой словесной модели бизнес-лояльность рассматривается с точки зрения трех самостоятельных базисных аспектов: лояльность потребителей, лояльность сотрудников и лояльность инвесторов. Каждый раз за словом «лояльность» понимается что-то свое:

- приверженность (с точки зрения покупателей),
- добросовестность (с точки зрения сотрудников),
- взаимное доверие, уважение и поддержка (с точки зрения инвесторов).

Но несмотря на ярко выраженные компоненты, эта система должна рассматриваться только как единое целое, поскольку невозможно создать лояльных покупателей, не обращая внимания на лояльность сотрудников, или воспитать лояльность сотрудников без должного внимания к лояльности инвесторов. Ни одна из частей не может существовать отдельно от двух других, но все три вместе позволяют организации достигать невиданных высот в развитии.

Менеджмент, основанный на лояльности, обращен на людей. В нем рассматриваются люди и их роль в бизнесе. Это скорее модель мотивации и поведения, чем маркетингового, финансового или производственного развития. Лишь во вторую очередь менеджмент, основанный на лояльности, обобщает людей в более абстрактные категории и управляет техническими процессами.

Как показывает практика, люди всегда оказываются более готовыми работать на организацию, которая имеет цель служения, чем на организацию, которая существует только ради того, чтобы «делать деньги». Поэтому люди охотно работают в церкви или в общественных организациях.

Менеджеры, желающие успешно использовать модель управления, основанную на эффекте лояльности, должны рассматривать прибыль не как первоочередную цель, а как необходимый элемент благосостояния и выживания трех составляющих каждой бизнес-системы: покупателей, сотрудников и инвесторов. Еще в начале XX в. Генри Форд говорил, что «...организация не может работать без прибыли ... иначе она умрет. Но и создавать организацию только ради прибыли... значит привести ее к верной гибели, так как у нее не будет стимула к существованию».

Основа рассматриваемой модели лояльности — не прибыль, а привлечение дополнительного количества покупателей, процесс, осознанно или неосознанно лежащий в основе большинства преуспевающих организаций. Создание целевого количества покупателей пронизывает все сферы бизнеса компании. Силы, управляющие взаимосвязями между покупателями, сотрудниками и инвесторами, называют *силами лояльности*. Критерий успешности — возвращаются ли покупатели, чтобы купить больше, или они идут куда-то еще, т.е. проявляют ли они лояльность.

Как причина лояльность инициирует несколько экономических эффектов, влияющих на всю бизнес-систему примерно следующим образом.

1. Прибыли и рыночная доля растут, когда наиболее перспективные покупатели охватывают весь спектр деятельности компании, создавая о ней хорошее общественное мнение и повторно приходя за покупками. За счет большого и качественного предложения компания может себе позволить быть более привередливой при выборе новых покупателей и концентрироваться на более прибыльных и потенциально лояльных проектах их привлечения, дальше стимулируя свой долгосрочный рост.

2. Долгосрочный рост позволяет фирме привлекать и сохранять лучших сотрудников. Постоянное поддержание целевого количества покупателей увеличивает лояльность сотрудников, давая им чувство

гордости и удовлетворения своей работой. Далее, в процессе взаимодействия постоянные сотрудники узнают больше о своих постоянных покупателях, в частности как улучшить обслуживание, чтобы их объем покупок рос. Этот увеличивающийся объем продаж подстегивает и лояльность покупателей, и лояльность сотрудников.

3. Лояльные сотрудники в долгосрочном периоде учатся снижать издержки и повышать качество работы (эффект научения). Организация может использовать эту дополнительную продуктивность для расширения системы вознаграждения, для покупки лучшего оборудования и обучения. Все это, в свою очередь, подстегнет продуктивность сотрудников, рост вознаграждений и, следовательно, лояльность.

4. Рассмотренная спираль продуктивности дает такое преимущество в издержках, которое очень сложно скопировать для чисто конкурентных организаций. Долгосрочные преимущества в издержках, соединенные с устойчивым ростом числа лояльных покупателей, приносят прибыль, очень привлекательную для инвесторов. Это, в свою очередь, расширяет возможности компании по привлечению и сохранению «правильных» инвесторов.

5. Лояльные инвесторы ведут себя как партнеры: стабилизируют систему, снижают издержки по поиску капитала и дают гарантии, что полученные отвлеченные денежные потоки будут вложены обратно в бизнес как инвестиции. Это укрепляет организацию и увеличивает ее производственный потенциал.

Обсудим основные идеи модели лояльности. Покупатели — активы любой организации, и для достижения успеха ей необходимо управлять ими так же эффективно, как и другими активами. Но для этого нужно быть в состоянии сегментировать покупателей, предсказывать их поведение, а также жизненный цикл их денежных потоков.

В основе большинства провалов лежит общепринятый бизнес-язык организации — бухгалтерский учет, ограничивающий возможности формирования лояльности. Бухгалтеры не в состоянии провести черту между выручкой, полученной от вновь пришедших покупателей, и выручкой, полученной от постоянных, лояльных покупателей. Это происходит потому, что они не знают, точнее, их не заботит тот факт, что обслуживание нового покупателя оказывается более дорогим, нежели обслуживание постоянного покупателя. Хуже того, в большинстве организаций бухгалтеры считают вложения в привлечение покупателей краткосрочными. И это вместо того, чтобы относить их на специальный счет покупателя и амортизировать в течение всего времени отношений с ним.

Как сформировать портфель лояльных покупателей? Существует два варианта действий:

- увеличение списка покупателей. Организация постоянно добавляет новых покупателей к началу списка, но ее старые покупатели также постоянно «вымываются» снизу из этого списка. Получается эффект дырявой корзины. Чем больше в ней дыра, тем тяжелее ее наполнить и сохранять наполненной;
- эффект прибыли от каждого покупателя. В большинстве организаций прибыль, которую приносит каждый покупатель, растет, пока он остается ее клиентом. Другими словами, для организации невыгодно терять постоянных покупателей, даже заменяя их новыми. Получается ситуация, когда «за одного битого двух небитых дают».

При подборе покупателей необходимо помнить, что существует три основных типа лояльных покупателей (это помогает определить, сможет ли организация сделать покупателя лояльным):

1) некоторые покупатели изначально предсказуемы и лояльны, вне зависимости от того, как организация с ними работает. Они просто лояльны по своей природе и предпочитают более стабильные и длительные отношения;
2) некоторые покупатели более прибыльны, чем другие. Они тратят деньги в большем количестве, чем другие, оплачивают покупки безотлагательно и требуют меньше внимания обслуживающего персонала;
3) некоторые покупатели находят продукты или услуги организации (в силу их особенностей) более привлекательными, чем у конкурентов. Нет такой организации, товары которой нравились бы всем без исключения. Сильные стороны ее товаров или услуг будут просто лучше подходить для определенных покупателей, более полно удовлетворяя их желаниям и возможностям.

Без сомнения, каждая организация уникальна, но все же в той или иной мере показатели ее прибылей будут укладываться в общую модель экономических эффектов, получаемых от постоянства или лояльности покупателей. Среди них стоит особо отметить следующие:

- издержки привлечения (реклама, направленная новым покупателям, комиссионные по продажам новым покупателям, накладные расходы продаж и т.д.);
- базовая прибыль (цена, которую платят вновь появившиеся покупатели, превышает затраты организации на создание товара);

- рост выручки (как правило, если покупатель доволен параметрами товара, он склонен увеличивать объемы покупок с течением времени);
- издержки сбережений (близкое знакомство с товарами организации уменьшает зависимость покупателей от ее сотрудников в вопросах информации и советов);
- отзывы (удовлетворенные уровнем обслуживания покупатели рекомендуют организацию своим друзьям и знакомым);
- дополнительная цена (постоянные покупатели, сотрудничающие с организацией достаточно долго, чтобы изучить все ее товары и услуги, получают несоизмеримо больше от продолжения отношений и не нуждаются в дополнительных скидках или рекламных акциях).

Чтобы оценить истинный долгосрочный потенциал лояльности покупателя или группы покупателей, необходимо знать их предрасположенность к проявлению постоянства. Так некоторые покупатели перебегут к конкуренту и за 2%-ную скидку, а другие останутся и при 20%-ной разнице в цене. Те усилия, которые требуются для переманивания различных типов покупателей, определяют *коэффициент лояльности*. В некоторых организациях для оценки коэффициентов лояльности используются история развития или поведение покупателей на отдельных сегментах. В других, особенно в тех, чье будущее слабо связано с прошлым, пытаются методами анализа данных нащупать, насколько велика должна быть скидка, чтобы покупатели перешли к их организации. Но несмотря на все трудности в измерении, использование коэффициента лояльности позволяет организациям идентифицировать сохранение покупателей и внедрять оправданную практику, проверенную на одном департаменте, во всю организацию.

Развитие систем измерения, анализа и управления денежными потоками, полученными от лояльности, может привести организацию к инвестициям, которые в дальнейшем обеспечат рост числа покупателей и организации в целом.

Итак, модель лояльности подробно обоснована на словесном уровне [87]. В этом обосновании упоминалось математическое и компьютерное обеспечение. Однако для принятия первоначальных решений их использование не требуется.

Математические модели при принятии решений. При более тщательном анализе ситуации словесных моделей, как правило, недостаточно. 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.

В процессе подготовки и принятия решений 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 [55, с. 213].

Основные термины математического моделирования. Прежде чем начать рассматривать конкретные математические модели процессов управления, необходимо вспомнить определения следующих основных терминов:

- *компоненты системы* — части системы, которые могут быть вычленены из нее и рассмотрены отдельно;
- *независимые переменные* — величины, которые могут изменяться, но являются внешними, не зависящими от проходящих в системе процессов;
- *зависимые переменные* — величины, значения которых есть результат (функция) воздействия на систему независимых внешних переменных;
- *управляемые (управляющие) переменные* — величины, значения которых могут изменяться исследователем (инженером, менеджером);
- *эндогенные переменные* — величины, значения которых определяются в ходе деятельности компонентов системы (т.е. «внутри» системы);
- *экзогенные переменные* — величины, определяемые либо исследователем, либо извне, т.е. в любом случае действующие на систему извне.

При построении любой модели процесса управления 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 следующие виды:

1) функциональные модели, выражающие прямые зависимости между эндогенными и экзогенными переменными;
2) модели, выраженные с помощью систем уравнений относительно эндогенных величин и, в свою очередь, выражающие балансовые соотношения между различными экономическими показателями (например, модель межотраслевого баланса);
3) модели оптимизационного типа, основная часть которых — система уравнений относительно эндогенных переменных. Цель данных моделей — нахождение оптимального решения для некоторого экономического показателя (например, таких величин ставок налогов, чтобы обеспечить максимальный приток средств в бюджет за заданный промежуток времени);
4) имитационные модели — достаточно точное отображение экономического явления. Математические уравнения при этом могут содержать сложные, нелинейные, стохастические зависимости.

Но модели также можно делить на управляемые и прогнозные. Управляемые модели отвечают на вопросы: «Что будет, если...?» или «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 вопрос: «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, так как информация поступает дискретно: отчеты, балансы и иные документы составляются периодически. Но с формальной точки зрения непрерывная модель может оказаться более простой для изучения. В физической науке продолжается дискуссия: 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 и правильно учесть такие факторы, как трудоотдача, структура потребления, мотивация и т.п.

При построении моделей, использующих дискретное время, часто применяют методы эконометрики [89]. Среди них популярны регрессионные уравнения и их системы. Различные системы регрессионных уравнений, построенные для решения практически важных задач, рассмотрены в [60]. Часто используют лаги (запаздывания в реакции). Для систем, нелинейных по параметрам, применение метода наименьших квадратов встречает трудности [89].

Большое количество конкретных эконометрических и математических моделей принятия решений рассмотрено в [60; 88]. Обратим внимание, что популярные в настоящее время подходы к процессам бизнес-реинжиниринга основаны на активном использовании математических и информационных моделей.

15.2. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ УПРАВЛЕНИЯ

Математическое моделирование экономических явлений и процессов с целью оптимизации процессов управления — область научно-практической деятельности, получившая мощный стимул к развитию во время и сразу после Второй мировой войны. Эта тематика развивалась 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 [14], и по нашему мнению, основанному на опыте СССР и России, в частности, анализе организационно-экономических результатов работы служб технического контроля на промышленных предприятиях) наибольший экономический эффект среди всех экономико-математических методов управления. Только дополнительный доход от их применения в промышленности США оценивается как 0,8% валового национального продукта США.

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 [88].

Особое место занимают имитационные системы, позволяющие отвечать на вопросы типа: «Что будет, если...?» (как подчеркнуто в [55, 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 [55, 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 [60, с. 9], при этом имитационный эксперимент состоит из следующих шести этапов:

1) формулировки задачи;
2) построения математической модели;
3) составления программы для ЭВМ;
4) оценки пригодности модели;
5) планирования эксперимента;
6) обработки результатов эксперимента.

Имитационное моделирование (*simulation modelling*) широко применяется в различных областях, в том числе в экономике [60].

Экономико-математические методы управления 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.

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 [127], вышедшей вскоре после Второй мировой войны. В учебниках по экономике обычно разбираются «дилемма заключенного» и «точка равновесия по Нэшу» (ему присуждена Нобелевская премия по экономике за 1994 г.).

15.3. О МЕТОДОЛОГИИ МОДЕЛИРОВАНИЯ

Моделирование процессов управления предполагает последовательное осуществление трех этапов исследования:

1) от исходной практической проблемы до теоретической чисто математической задачи;
2) внутриматематическое изучение и решение этой задачи;
3) переход от математических выводов обратно к практической проблеме.

В области моделирования процессов управления, как, впрочем, и в иных областях применения математики, целесообразно выделять четверки составляющих:

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? 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 переформулируется в рамках этой модели как 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 двум совокупностям [89].

Задача может быть порождена также обобщением потребностей ряда прикладных областей. Приведенный ранее 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.

Выделение перечня задач находится вне математики. Выражаясь инженерным языком, этот перечень является сутью технического задания, которое специалисты различных областей деятельности дают специалистам по математическому моделированию.

Метод, используемый в рамках определенной математической модели, — это уже во многом, если не в основном, дело математиков. В эконометрических моделях речь идет, например, о методе оценивания, о методе проверки гипотезы, о методе доказательства той или иной теоремы и т.д. В первых двух случаях алгоритмы разрабатываются и исследуются математиками, но используются экономистами, в то время как метод доказательства касается лишь самих математиков.

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 ранее проблеме однородности для проверки одной и той же гипотезы совпадения функций распределения могут быть применены самые разные методы — Смирнова, Лемана — Розенблатта, Вилкоксона и др. [89].

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 экономист оценить это достижение не сможет. Для него, как и во времена Ньютона и Лейбница, непрерывные функции мало отличаются от кусочно дифференцируемых. Точнее, они одинаково хорошо (или одинаково плохо) могут быть использованы для описания реальной действительности.

Методологический анализ — первый этап моделирования процессов управления, да и вообще любого исследования. Он определяет исходные постановки для теоретической проработки, а потому во многом и успех всего исследования. Анализ динамики развития методов моделирования позволяет выделить наиболее перспективные методы. В частности, при вероятностно-статистическом моделировании наиболее перспективными оказались методы нечисловой статистики [89].

Рассмотрим два конкретных примера — детерминированную модель управления обучением и вероятностно-статистическую модель помех, создаваемых электровозами.

15.4. МОДЕЛЬ УПРАВЛЕНИЯ ОБУЧЕНИЕМ

В качестве примера конкретной модели процесса управления рассмотрим модель распределения времени между овладением знаниями и развитием умений [71].

Любое знание состоит частично из «информации» («чистое знание») и частично из «умения» («знаю как»). Умение — это мастерство, это способность использовать имеющиеся у вас сведения для достижения своих целей; умение можно еще охарактеризовать как совокупность определенных навыков, в конечном счете умение — это способность методически работать [101, с. 308].

Пусть $x(t)$ — объем сведений, накопленных учащимся к моменту времени t («чистое знание»), $y(t)$ — объем накопленных умений: рассуждать, решать задачи, разбираться в излагаемом преподавателем материале; $u(t)$ — доля времени, отведенного на накопление знаний в промежутке времени $(t; t + dt)$.

Примем в модели в качестве исходного положения, что увеличение $x(t + dt) - x(t)$ объема знаний учащегося пропорционально потраченному на это времени $u(t)dt$ и накопленным умениям $y(t)$. Следовательно,

$$\frac{dx(t)}{dt} = k_1 u(t) y(t), \tag{15.1}$$

где коэффициент $k_1 > 0$ зависит от индивидуальных особенностей учащегося.

Примем также, что увеличение знаний за то же время пропорционально потраченному на это времени $[1 - u(t)]dt$, имеющимся умениям $y(t)$ и знаниям $x(t)$. Следовательно,

$$\frac{dy(t)}{dt} = k_2 [1 - u(t)] x(t) y(t). \tag{15.2}$$

Коэффициент $k_2 > 0$ также зависит от индивидуальности. Учащийся тем быстрее приобретает умения, чем больше он уже знает и умеет. Тем быстрее усваивает знания, чем больше умеет. Но нельзя считать, что чем больше он запомнил, тем быстрее запоминает. На правую часть уравнения (15.1) влияют только приобретенные в прошлом активные знания, примененные при решении задач и перешедшие в умения. Отметим, что модель (15.1)—(15.2) имеет смысл применять на таких интервалах времени, чтобы, например, пять минут можно было считать бесконечно малой величиной.

Можно управлять процессом обучения, выбирая при каждом t значение функции $u(t)$ из отрезка $[0; 1]$. Рассмотрим две задачи.

1. Как возможно быстрее достигнуть заданного уровня знаний x_1 и умений y_1? Другими словами, как за кратчайшее время перейти из точки фазовой плоскости $(x_0; y_0)$ в точку $(x_1; y_1)$?

2. Как быстрее достичь заданного объема знаний, т.е. выйти на прямую $x = x_1$?

Двойственная задача: за заданное время достигнуть как можно большего объема знаний. Оптимальные траектории движения для второй задачи и двойственной к ней совпадают (двойственность понимается в обычном для математического программирования смысле).

С помощью замены переменных $z = k_2 x$, $w = k_1 k_2 y$ перейдем от системы (15.1)—(15.2) к более простой системе дифференциальных уравнений, не содержащей неизвестных коэффициентов:

$$\frac{dz}{dt} = uw, \quad \frac{dw}{dt} = (1 - u)zw. \tag{15.3}$$

Описанная линейная замена переменных эквивалентна переходу к другим единицам измерения знаний и умений, своим для каждого учащегося.

Решения задач 1 и 2, т.е. наилучший вид управления $u(t)$, находятся с помощью математических методов оптимального управления, а именно с помощью принципа максимума Л.С. Понтрягина. В задаче 1 для системы (15.3) из этого принципа следует, что быстрейшее движение может происходить либо по горизонтальным ($u = 1$) и вертикальным ($u = 0$) прямым, либо по особому решению — параболе $w = z^2$ ($u = {}^1/_3$). При $z_0^2 > w_0$ движение начинается по вертикальной прямой, при $z_0^2 < w_0$ — по горизонтальной, при $z_0^2 = w_0$ — по параболе. По каждой из областей $\{z^2 > w\}$ и $\{z^2 < w\}$ проходит не более одного вертикального и одного горизонтального отрезка оптимальной траектории.

Используя теорему о регулярном синтезе, можно показать, что оптимальная траектория выглядит следующим образом. Сначала надо выйти на «магистраль» — добраться до параболы $w = z^2$ по вертикальной ($u = 0$) или горизонтальной ($u = 1$) прямой. Затем пройти основную часть пути по магистрали ($u = {}^1/_3$). Если конечная точка лежит под параболой, добраться до нее по горизонтали, сойдя с магистрали. Если она лежит над параболой, заключительный участок траектории является вертикальным отрезком. В частности, в случае $w_0 < z_0^2 < w_1 < z_1^2$ оптимальная траектория такова. Сначала надо выйти на магистраль — добраться по вертикальной ($u = 0$) прямой до параболы. Затем двигаться по магистрали ($u = {}^1/_3$) от точки $(z_0; z_0^2)$ до точки $(\sqrt{w_1}; w_1)$. Наконец, по горизонтали ($u = 1$) выйти в конечную точку.

В задаче 2 из семейства оптимальных траекторий, ведущих из начальной точки $(z_0; w_0)$ в точки луча $(z_1; w_1)$, $w_0 \le w_1 < +\infty$, выбирается траектория, требующая минимального времени. При $z_1 \le 2z_0$ оптимально $w_1 = z_0(z_1 - z_0)$, траектория состоит из вертикального и горизонтального отрезков. При $z_1 > 2z_0$ оптимально $w_1 = z_1^2/4$, траектория проходит по магистрали $w = z^2$ от точки $(z_0; z_0^2)$ до точки $(z_1/2; z_1^2/4)$. Чем большим объемом знаний z_1 надо овладеть, тем большую долю времени надо двигаться по магистрали, отдавая при этом ${}^2/_3$ времени увеличению умений и ${}^1/_3$ времени — накоплению знаний.

Полученное для основного участка траектории оптимального обучения значение $u = {}^1/_3$ можно интерпретировать приблизительно так: на одну лекцию должно приходиться два семинара, на 15 мин объяснения — 30 мин решения задач. Результаты, полученные в математической модели, вполне соответствуют эмпирическим представлениям об оптимальной организации учебного процесса. Кроме того, модель определяет численные значения доли времени (${}^1/_3$), идущего на повышение знаний, и доли материала (${}^1/_2$), излагаемого на заключительных лекциях (без проработки на семинарах).

При движении по магистрали, т.е. в течение основного периода учебного процесса, оптимальное распределение времени между объяснениями и решением задач одно и то же для всех учащихся независимо от индивидуальных коэффициентов k_1 и k_2. Этот факт устойчивости оптимального решения показывает возможность организации обучения, оптимального одновременно для всех учащихся. При этом время движения до выхода на магистраль зависит, естественно, от начального положения $(x_0; y_0)$ и индивидуальных коэффициентов k_1 и k_2.

Таким образом, модель процесса управления обучением (15.1)—(15.2) позволила получить ряд практически полезных рекомендаций, в том числе выраженных в числовой форме. При этом не понадобилось уточнять способы измерения объемов знаний и умений, имеющихся у учащегося. Достаточно было согласиться с тем, что эти величины удовлетворяют качественным соотношениям, приводящим к уравнениям (15.1) и (15.2).

Многочисленные модели процессов управления описаны в литературе [60; 87; 88; 89; 127]. Их практическим использованием обычно занимаются информационно-аналитические подразделения, службы контроллинга, качества и надежности, маркетинга и др.

15.5. ВЕРОЯТНОСТНО-СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПОМЕХ, СОЗДАВАЕМЫХ ЭЛЕКТРОВОЗАМИ

В качестве примера моделирования рассмотрим задачу, относящуюся к железнодорожному транспорту.

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 [88], m-ю гармонику ($m \geq 7$) создаваемой электровозом помехи можно описать двумерным случайным вектором

$$(\xi_1, \xi_2) = (r\cos\varphi, r\sin\varphi), \tag{15.4}$$

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,

$$\begin{gathered} M(\xi_1) = M(\xi_2) = M(\xi_1\xi_2) = M(\xi_1^3) = M(\xi_2^3) = 0, \\ D(\xi_1) = D(\xi_2) = \frac{1}{2}M(r^2). \end{gathered} \tag{15.5}$$

В случае n электровозов рассматриваемая гармоника суммарной помехи описывается вектором (α_n, β_n), причем из физических соображений

$$(\alpha_n, \beta_n) = \sum_{i=1}^{n} r_i(\cos\varphi_i, \sin\varphi_i). \tag{15.6}$$

Каждый вектор в правой части (15.6) соответствует электровозу с тем же номером. Будем считать, что случайные величины $\{r_i, \varphi_i, i = 1, 2, ..., n\}$ независимы в совокупности, а все фазы φ_i распределены равномерно на $[0; 2\pi]$. Из (15.5) и (15.6) вытекает, что

$$M(\alpha_n^2 + \beta_n^2) = \sum_{i=1}^{n} M(r_i^2), \tag{15.7}$$

если все слагаемые в правой части (15.7) существуют. Последнее будем считать выполненным. Правую часть (15.7) обозначим R_n^2.

Из многомерной центральной предельной теоремы (см., например, [77, подразд. 4.2])вытекает следующее утверждение.

Лемма 15.1. Пусть выполнены условия теоремы Линдеберга — Феллера, в данном случае

$$\lim_{n\to\infty} R_n = +\infty, \ \lim_{n\to\infty}\left(\frac{1}{R_n^2}\max_{1\leq i\leq n} M(r_i^2)\right) = 0. \tag{15.8}$$

Тогда для любых чисел x, y

$$\lim_{n\to\infty} P\left\{\frac{\sqrt{2}}{R_n}\alpha_n < x; \ \frac{\sqrt{2}}{R_n}\beta_n < y\right\} = \Phi(x)\Phi(y), \tag{15.9}$$

где $\Phi(z)$ — функция нормального распределения с математическим ожиданием 0 и дисперсией 1.

Нам понадобится следующее утверждение.

Лемма 15.2. Пусть ξ и η — независимые случайные величины, имеющие стандартное нормальное распределение $\Phi(z)$. Пусть выполнены условия леммы 15.1. Пусть функция $f : R^2 \to R^1$ интегрируема по Риману по любому квадрату. Тогда

$$\lim_{n\to\infty} P\left\{f\left(\frac{\sqrt{2}}{R_n}\alpha_n, \ \frac{\sqrt{2}}{R_n}\beta_n\right) < u\right\} = P\{f(\xi, \eta) < u\} \tag{15.10}$$

при всех u, при которых правая часть (15.10) непрерывна по u.

Доказательство вытекает из леммы 15.1 и результатов о наследовании сходимости [77, подраздел 4.3].

В методиках расчетов, связанных с защитой проводных линий связи, используется математическое ожидание амплитуды суммарной помехи (α_n, β_n). Следовательно, надо рассмотреть функцию

$$f(z, w) = \sqrt{x^2 + w^2}.$$

Из леммы 15.2 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, [35, с. 262]), имеет плотность

$$h(x) = x\exp\left\{-\frac{x^2}{2}\right\},\ x \geq 0. \tag{15.11}$$

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 15.1. Пусть предельные соотношения (15.8), (15.9) имеют место и, кроме того,

$$\overline{\lim_{n\to\infty}} \frac{1}{R_n^4}\sum_{i=1}^{n} M(r_i^4) < +\infty. \tag{15.12}$$

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 (15.8) и (15.12) выполнены, например, в случае, когда существуют числа $0 < r < R$ такие, что $r < r_{in} < R$ при всех n, $i = 1, 2, ..., n$, т.е. амплитуды помех, создаваемых электровозами, ограничены сверху и снизу. Это условие является весьма естественным с прикладной точки зрения.

Доказательство теоремы 15.1. При любом $a > 0$ справедливо равенство

$$M(\gamma_n) - M(\delta) = \left\{\int_{|x|<a} xdP(\gamma_n < x) - \int_{|x|<a} xh(x)dx\right\} +$$

$$+\left\{\int_{|x|\geq a} xdP(\gamma_n < x) - \int_{|x|\geq a} xh(x)dx\right\}. \tag{15.13}$$

Каждую из скобок в представлении (15.13) будем обрабатывать своим способом. Поскольку

$$\int_{-a}^{a} xdF(x) = xF(x)\Big|_{-a}^{a} - \int_{-a}^{a} F(x)dx, \tag{15.14}$$

то, подставляя в (15.14) вместо $F(x)$ соответственно $P(\gamma_n < x)$ и $P(\delta < x)$, заключаем с помощью леммы 15.2, что при фиксированном значении a первая скобка в (15.13) стремится к 0 при росте n.

Вторую скобку оценим следующим образом:

$$\left|\int_{|x|\geq a} xdP(\gamma_n < x) + \int_{|x|\geq a} xh(x)dx\right| \leq \int_{|x|\geq a} |x|dP(\gamma_n < x) + \int_{|x|\geq a} |x|h(x)dx. \tag{15.15}$$

Из явного вида функции $h(x)$ — см. формулу (15.11) — следует, что второе слагаемое в (15.15) стремится к 0 при росте n. Для оценки первого слагаемого в (15.15) воспользуемся соотношением

$$\int_a^{+\infty} xdF(x) = \int_a^{+\infty} xd[F(x)-1] = \int_a^{+\infty} xd[F(x)-1]\Big|_a^{+\infty} - \int_a^{+\infty}[F(x)-1]dx,$$

в котором положим $F(x) = P(\gamma_n < x)$. Оценим сначала порядок убывания по x величины

$$1 - P(\gamma_n < x) = 1 - P\left[\frac{2}{R_n^2}\left(\alpha_n^2 + \beta_n^2\right) < x^2\right].$$

С помощью несложных, но достаточно продолжительных вычислений можно найти $M[(\alpha_n^2 + \beta_n^2)^2]$ и на основе неравенства (15.12) доказать, что существует константа C такая, что при всех n

$$M\left[\left(\frac{2}{R_n^2}\left(\alpha_n^2 + \beta_n^2\right)\right)^2\right] < C.$$

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

$$P\left[\frac{2}{R_n^2}\left(\alpha_n^2+\beta_n^2\right) \geq x^2\right] = 1-P(\gamma_n < x) \leq \frac{C}{x^4}.$$

Следовательно,

$$0 \leq \int_a^{+\infty} x dP(\gamma_n < x) < \int_a^{+\infty}\left[1-P(\gamma_n < x)\right]dx \leq \int_a^{+\infty}\frac{C}{x^4}dx = \frac{C}{3a^3}. \quad (15.16)$$

Из соотношений (15.15), (15.11) и (15.16) следует, что первое слагаемое в правой части равенства (15.13) 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 15.1 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 15.1 доказана.

Рассмотрим важный частный случай. Пусть амплитуды помех от всех электровозов детерминированы, одинаковы и равны 1, т.е. $r_i = 1$, $i = 1, 2, ..., n$. Тогда по следствию из теоремы 15.1 математическое ожидание амплитуды суммарной помехи приблизительно равно $0{,}886\sqrt{n}$, в то время как специалисты-электротехники использовали значение $\sqrt{n}$ [88]. Таким образом, согласно развитой ранее 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» является приближенным и выполняется с точностью порядка 1% [88]. 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, принятая электротехниками формула 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 данной теории позволяет сэкономить достаточно большие материальные и финансовые ресурсы.

Однако приведенные 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 15.1—15.2 и особенно в теореме 15.1.

Пусть сначала $n = 2$. Нетрудно показать, что в случае детерминированных амплитуд r_1 и r_2

$$M\left[|\alpha_n \beta_n|\right] = R_2 M\left(\sqrt{1+\gamma\cos\varphi}\right), \quad (15.17)$$

где $R_2 = \sqrt{r_1^2 + r_2^2}$; фаза φ — равномерно распределенная на $[0; 2\pi]$ случайная величина, а

$$\gamma = \frac{2r_1 r_2}{r_1^2 + r_2^2}.$$

Легко видеть, что при $\gamma = 1$ (т.е. при $r_1 = r_2$) математическое ожидание в правой части (15.17) равно 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:

$$M\left(\sqrt{1+\gamma\cos\varphi}\right) = \sum_{k\geq 0}\binom{0{,}5}{2k}2^{-2k}\binom{2k}{k}\gamma^{2k} = 1-\frac{1}{16}\gamma^2-\frac{15}{1024}\gamma^4-\frac{105}{16384}\gamma^6-\ldots$$

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 формуле (15.5). 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 15.1 может быть сделан переход к интересующей нас разности

$$M(\gamma_n) - M(\delta) = O\left(\frac{1}{n}\right). \quad (15.18)$$

Обосновать описанную цепочку утверждений, приводящую к (15.18), и получить численные оценки скорости сходимости можно, оценивая остаточные члены во всех описанных ранее 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 (15.18), 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 [77, подраздел 4.7], константы, которые можно таким путем получить для уточнения формулы (15.18), 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 [124] использовался алгоритм перемешивания Макларена — Марсальи (М-алгоритм).

В таблице 15.1 приведены значения оценок величины

$$\rho = \frac{M\left[\left|(\alpha_n, \beta_n)\right|\right]}{\sqrt{\sum\limits_{1 \le i \le n} r_i^2}} - 0{,}886 \tag{15.19}$$

для пяти наборов амплитуд.

Таблица 15.1

Оценка погрешности ρ методом Монте-Карло

Номер эксперимента	Число электровозов n									
	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
	—	0,009	0,018	0,010	0,013	0,013	0,001	0,001	0,010	0,010
2	0,40	0,41	0,72	0,64	1,18	0,66	1,60	1,70	0,87	2,60
	—	0,016	0,029	0,021	0,022	0,022	0,025	0,018	0,008	0,007
3	0,40	0,40	0,41	0,86	0,80	0,86	0,85	0,90	2,97	0,72
	—	0,010	0,022	0,030	0,016	0,012	0,009	0,007	0,040	0,035
4	0,40	0,53	2,52	0,72	0,85	3,00	0,65	1,46	0,63	0,84
	—	0,031	0,095	0,077	0,067	0,025	0,029	0,017	0,023	0,022
5	0,40	0,41	0,85	0,56	0,82	0,58	1,63	1,64	0,87	1,00
	—	0,014	0,041	0,017	0,017	0,007	0,025	0,011	0,009	0,012

Примечание. Числитель — значения амплитуд, знаменатель — значения ρ из (15.19), 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 (15.19) оценивались как средние арифметические 10 000 наблюдений случайной величины $|(\alpha_n, \beta_n)|$.

Дисперсия результата одного испытания согласно лемме 15.2, аналогу теоремы 15.1 для вторых моментов и соотношению (15.11) оценивается как $D(\delta) / 2 \approx 0{,}22$, а потому среднеквадратичное отклонение табличных значений оценивается как 0,0015 (в предположении идеальности датчика псевдослучайных чисел).

Данные табл. 15.1 показывают, что наиболее быстрая сходимость ρ в теореме 15.1 имеет место в случае равенства амплитуд. Напротив, если значение одной из амплитуд безгранично возрастает, а значения остальных фиксированы, то $\rho \to 0{,}114$. В целях изучения сходимости в случае различных амплитуд были взяты четыре независимые выборки значений амплитуд [88]. Объемы выборок — 10 наблюдений. Выборки приведены в табл. 15.1. 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 таблицы 15.1 следует, что существенные различия между амплитудами приводят к возрастанию ρ, однако при росте числа электровозов n это возрастание становится все менее существенным. Так, для $n \ge 6$ максимальное табличное значение равно 0,040, т.е. при $n \ge 6$ имеем основания полагать, что

$$M\left[\left|(\alpha_n, \beta_n)\right|\right] \le 0{,}926\sqrt{\sum_{1 \le i \le n} r_i^2} \tag{15.20}$$

для подавляющего большинства реальных ситуаций. Замена 1 на 0,926 в методиках защиты проводных линий связи сулит значительный экономический эффект. Кроме формулы (15.20), данные табл. 15.1 позволяют дать ряд других полезных практических рекомендаций.

Замечание. Предположение о равномерности распределения фазы φ использовалось ранее лишь при выводе соотношений (15.5) и (15.17). Поэтому все результаты настоящего раздела, кроме соотношения (15.17), справедливы для любых распределений фаз, удовлетворяющих соотношениям (15.5). В частности, для различных электровозов распределения могут быть различными.

Контрольные вопросы

1. В чем сходство и различие словесных и математических моделей?
2. Согласны ли вы с моделью лояльности, описанной в подразделе 15.1?
3. Каковы основные виды переменных в математических моделях процессов управления?
4. Какие виды математических моделей принятия решений обычно выделяют?

5. В чем суть методологии математического моделирования?
6. Как в соответствии с моделью подраздела 15.4 должны соотноситься затраты времени на накопление знаний и развитие умений во время основного периода обучения?
7. Что представляет собой вероятностно-статистическая модель, предназначенная для изучения помех, создаваемых электровозами?

Темы докладов и рефератов

1. Классификация математических моделей процессов управления.
2. Сравнение словесных и математических моделей.
3. Модели процессов управления предприятием.
4. Модели процессов управления качеством.
5. Макроэкономические модели управления.
6. Соотношение задач, моделей, методов и условий применимости.
7. Место принципа максимума Понтрягина среди математических методов оптимального управления.
8. Статистическое моделирование с использованием датчиков псевдослучайных чисел (метод Монте-Карло).
9. Методы проверки качества датчиков псевдослучайных чисел.

ГЛАВА 16
ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МОДЕЛИ И ПРИНЯТИЕ РЕШЕНИЙ

16.1. ПРИМЕРЫ ТИПОВЫХ МАКРОЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

Принятие решений проводится на основе прогнозирования развития ситуации с учетом динамических связей между переменными. Эти связи описываются экономико-математическими моделями, краткий анализ которых необходим для рассмотрения задач принятия решений.

Модель межотраслевого баланса (модель В. Леонтьева). Каждая из n отраслей производит свой (обобщенный) продукт. Выпуск распределяется в заданной пропорции между конечным потреблением, другими отраслями и внутренними потребностями отрасли. Кроме того, описывается прирост производственных мощностей. Модель описывается уравнением

$$v_i = \sum_{j=1}^{n}\left[a_{ij}v_j(t) + b_{ij}\frac{dV_j(t+\tau_j)}{dt}\right] + P_i(t),\ i = 1, 2, \dots, n,$$

где $v_i(t)$ — поток выпуска продукта i в момент времени t (единица измерения = единица продукта / единица времени); a_{ij} — коэффициенты прямых сырьевых затрат (количество продукта i, необходимое для производства продукта j); $v_j(t)$ — поток выпуска продукта j в момент времени t; b_{ij} — количество фондообразующего продукта i, идущее на единичный прирост мощности в отрасли j; $V_j(t+\tau_j)$ — мощность j-го производства в момент времени $t+\tau_j$; τ_j — продолжительность строительства мощности в отрасли j; $P_i(t)$ — поток конечного (непроизводственного) потребления.

Таким образом, выпуск $v_i(t)$ расходуется на покрытие сырьевых и фондообразующих затрат и конечное потребление.

Эконометрические модели (типа Брукингской и Уортоновской). В основе этих моделей лежат: 1) балансовые соотношения; 2) функциональные зависимости — производственная функция и функция потребительского спроса.

Производственная функция F задает зависимость национального дохода Y от стоимости основных фондов (капитала) K и от используемых трудовых ресурсов L:

$$Y(t) = F[K(t), L(t)].$$

Функция спроса $P = S(c, q)P=S(c,q)$ задает зависимость вектора P конечного потребления, т.е. набора потребляемых товаров, от вектора c цен на эти товары и дохода q.

Паутинообразные модели. Применяются данные модели при изучении динамики спроса и предложения. Пусть D — спрос, S — предложение, P — цена, P^* — равновесная цена, X — объем производства, X^* — равновесный объем производства. Равновесные P^* и X^* находят из условия совпадения спроса и предложения $D(P) = S(P)$.

Однако более реалистичной является гипотеза запаздывания предложения. Например, пусть при цене в прошлый период P_{t-1} объем предложения в данный период есть $S_t = S(P_{t-1})$. Считаем, что цена P_t устанавливается на рынке так, чтобы был куплен весь объем выпущенной продукции X_t. Следовательно,

$$X_t = D(P_t) = S(P_{t-1}).$$

Пусть спрос и предложение достаточно точно описываются линейными функциями от цены:

$$D = \alpha + aP;$$

$$S = \beta + bp.$$

Такое предположение вполне естественно, если в модели рассматривается окрестность точки равновесия, а функции спроса и предложения гладкие. Тогда

$$X_t = \alpha + aP_t = \beta + bP_{t-1}. \tag{16.1}$$

Равновесие наступает, когда

$$X^* = \alpha + aP^* = \beta + bP^*. \tag{16.2}$$

Вычитая (16.1) из (16.2), получаем, что

$$X^* - X_t = a(P^* - P_t) = b(P^* - P_{t-1}) \tag{16.3}$$

Обозначим $x_t = X^* - X_t$; $p_t = P^* - P_t$ — отклонения от равновесия. Из (16.3) получим $x_t = ap_t = bp_{t-1}$, откуда $p_t = \dfrac{b}{a} p_{t-1}$. Решение этого уравнения имеет вид $p_t = p_0 \left(\dfrac{b}{a}\right)^2$.

В зависимости от того, чему равно $\dfrac{b}{a}$, получим либо затухающие колебания $\left(\left|\dfrac{b}{a}\right| < 1\right)$, сходящиеся к $P = P^*$ и $X = X^*$, либо колебания с возрастающей амплитудой $\left(\left|\dfrac{b}{a}\right| > 1\right)$. В промежуточном случае $a = b$ амплитуда колебаний постоянна.

Тот же результат справедлив и в модели с непрерывным временем. Будем считать, что спрос меняется не только в зависимости от цены, но и в зависимости от ее динамики, т.е.

$$D = D\left(P, \frac{dP}{dt}\right); \; S = S(P).$$

Тогда аналогом формулы (16.1) является уравнение $X = \alpha + aP + a_1 \dfrac{dP}{dt} = \beta + bP$, решением которого является $p = p_0 e^{ct}$.

В рассматриваемых моделях считалось, что производители ожидают, что цена останется, как в предшествующий период (и устанавливают объем изготавливаемого товара исходя из этих ожиданий). Модель может быть усовершенствована. Для установления объема изготавливаемого товара производителям более реалистично считать, что в момент времени t цена на товар будет равна $P_{t-1} - \rho(P_{t-1} - P_{t-2})$, где $0 < \rho < 1$, т.е. цена изменится в направлении, обратном тому, в котором она изменялась в прошлый период. Тогда $X_t = \alpha + aP_t = \beta + b(1 - \rho)P_{t-1} + b\rho P_{t-2}$, следовательно, $x_t = ap_t = b(1 - \rho)p_{t-1} + b\rho p_{t-2}$.

Дальнейшее развитие модели состоит во введении в нее запасов. Ожидая повышения цен, продавцы создают запасы товара.

Запасы в момент времени t обозначим Q_t. Тогда изменение запасов за период времени от $t - 1$ до t есть $Q_t - Q_{t-1} = S_t - D_t$. В модели цену можно устанавливать различными способами, например $P_t = P_{t-1} - \lambda(Q_{t-1} - Q_{t-2})$ или $P_t = P_{t-1} - \lambda(Q_{t-1} - Q^*)$, где Q^* — запасы в точке равновесия. В первом случае получим $P_t = P^* + (P_0 - P^*)c^t$, где $c = 1 - \lambda(b - a)$, а во втором — $P_t = [2 - \lambda(b - a)]P_{t-1} - P_{t-2}$.

Модель экономического цикла. Сначала рассмотрим простую модель без учета запаздывания, а также без учета экспорта-импорта, налогов и государственных расходов:

$$C = (1 - s)Y + A; \tag{16.4}$$

$$DK = \gamma(vY - K); \tag{16.5}$$

$$DY = \lambda(C + DK - Y), \tag{16.6}$$

где C — реальное потребление; A, $s < 1$, v, γ, λ — положительные константы; Y — реальный чистый доход; $D = \dfrac{d}{dt}$ — символ операции дифференцирования; K — объем основного капитала.

Более точно, Y — сумма всех видов конечных доходов, полученных в экономике, деленная на индекс инфляции (т.е. реальный валовой национальный продукт за вычетом затрат на возмещение основного капитала); C — общие затраты на потребительские товары конечных покупателей в экономике, деленные на индекс инфляции; K — объем основного капитала всей экономики страны (в сопоставимых ценах).

Уравнение (16.4) вытекает из теории Кейнса, а именно из соотношения: потребление = национальный доход – сбережения + автономное потребление. Значит, sY — часть дохода, идущая на сбережения; s — предельная склонность к сбережениям; A — автономное потребление (та доля потребления, которая не зависит от дохода, своеобразный прожиточный минимум).

Уравнение (16.5) допускает несколько интерпретаций. Рассмотрим две из них.

1. В первой интерпретации DK — это норма капитальных вложений в основной капитал. Допустим, существует оптимальный объем основного капитала и он равен некоторой доле от национального дохода — vY, где v — оптимальное соотношение «капитал—выпуск». Тогда уравнение (16.5) означает, что норма капитальных вложений в основной капитал пропорциональна превышению оптимального объема основного капитала над действительным.

2. Основное соотношение, описывающее капитальные вложения, имеет вид:

$$\frac{DK}{K} = \gamma\left[\frac{P}{(1+c)rK} - 1\right], \tag{16.7}$$

где P — реальная прибыль; c — премия за риск; r — норма процента.

Из соотношения (16.7) легко получить (16.5).

В уравнении (16.6) $DY = \dfrac{dY}{dt}$ — рост производства (поскольку все производство = всему доходу = Y). Рост производства зависит от избытка спроса. Потребление C + накопление (оно превращается в капитальные вложения DK) — чистый национальный доход Y — это и есть избыток спроса (то, что потребляется и накопляется, может быть не равно чистому доходу).

Для равновесной системы все производные по времени равны 0. Равновесные значения Y, C и K таковы:

$$Y^* = C^* = \frac{A}{s}; \tag{16.8}$$

$$K^* = \frac{vA}{s}. \tag{16.9}$$

Этот результат не предназначен для непосредственного практического использования, так как в модели не учитываются ограничения на выпуск, накладываемые рабочей силой и объемом основного

капитала. Однако он нужен, чтобы найти отклонения от равновесия $Y = Y^* + B_1 e^{x_1 t} - B_2 e^{x_2 t}$ — решение системы (16.4)—(16.5)—(16.6)—(16.8)—(16.9), где B_1 и B_2 зависят от λ, γ, v, s. В зависимости от B_1 и B_2 получим, согласно теории линейных дифференциальных уравнений, следующие четыре варианта траекторий Y: 1) незатухающие колебания (экономические циклы); 2) затухающие колебания; 3) взрывоподобные колебания; 4) взрывоподобная, но не колебательная траектория.

Довольно часто в экономике реально осуществляется приближение к первому варианту — экономические циклы.

Усложним модель, введем запаздывание. В модели (16.4)—(16.6) предполагается мгновенная реакция потребления на изменение дохода. На самом деле это неверно. Вместо уравнения (16.4) напишем

$$DC = \alpha[(1 - s)Y + A - C], \tag{16.10}$$

где α — параметр, определяющий быстродействие системы.

Теперь добавим запасы. Вместо уравнения (16.6) получим:

$$DY = \lambda(C + DK - Y) + \mu(S^0 - S); \tag{16.11}$$

$$S^0 = b(C + DK) + c; \tag{16.12}$$

$$DS = Y - C - DK, \tag{16.13}$$

где S^0 — оптимальный уровень запасов, равный некоторой постоянной величине + часть потребления и капитальных вложений; S — фактический уровень запасов.

Уравнение (16.11) отражает тот факт, что рост производства зависит от избытка спроса и от превышения оптимальных запасов над фактическими. Уравнения (16.10) и (16.11) аналогичны соответствующим соотношениям для паутинообразных моделей.

Добавим в систему экспорт-импорт налоги и государственные расходы. Теперь с учетом (16.11)—(16.13) получим модель в виде системы уравнений:

$$DC = \alpha[(1 - s)(Y - T) + A - C], \tag{16.14}$$

$$DK = \gamma(vY - k), \tag{16.15}$$

$$DY = \lambda(C + DK + G + E - I - Y) + \mu(S^0 - S), \tag{16.16}$$

$$S^0 = b(C + DK + G + E) + c, \tag{16.17}$$

$$DS = Y + I - E - G - C - DK, \tag{16.18}$$

$$I = m(C + DK + G + E), \tag{16.19}$$

$$T = \tau Y - B, \tag{16.20}$$

где T — реальный объем налогов за вычетом государственных трансфертных платежей; G — реальные государственные расходы на товары и услуги; E — реальный экспорт; I — реальный импорт.

В уравнении (16.14) национальный доход, идущий на потребление и накопление, уменьшился на сумму налогов, т.е. по сравнению с (16.10) произошла замена $Y \to Y - T$.

Далее заметим, что теперь C — общее потребление товаров как отечественного, так и импортного производства, а DK теперь есть рост основного капитала частного сектора. Накопление основного капитала частного сектора входит в G.

Уравнение (16.16) отличается от (16.11) на величину $G + E - I$, так как DY — рост производства, зависящий от избытка спроса, который теперь равен тому, что общество расходует, т.е. потребляет (C) + вкладывает (DK) + экспорт (E) + государственные расходы (G), за вычетом того, что общество получает: национальный доход (Y) + импорт (I).

Уравнение (16.17) предполагает, что желаемый уровень запасов есть линейная функция валового сбыта, а валовой сбыт — это: 1) сбыт потребительских товаров отдельным потребителям C; 2) сбыт капитальных благ фирмам (капитальные вложения) DK; 3) сбыт товаров в государственном секторе G; 4) сбыт иностранным производителям E.

Уравнение (16.18) означает, что изменение запасов равно всем товарам ($Y + I$) минус весь сбыт ($C + DK + G + E$).

Уравнение (16.19) предполагает, что импорт — это доля всего сбыта.

Уравнение (16.20) предполагает, что налоги — линейная функция доходов, тогда τ — аналог процентной ставки. То, что в уравнении имеется отрицательная константа B, говорит о том, что $\frac{T}{Y}$ — возрастающая функция, т.е. чем больше доход, тем больше налог.

При решении системы (16.14)—(16.20) выяснилось, в частности, что введение налогов и импорта оказывает на экономику стабилизирующее воздействие.

Модель экономического роста. В этой модели, в отличие от модели экономического цикла, считается, что предложение денег пропорционально $\exp(mt)$ и предложение труда пропорционально $\exp(lt)$, т.е явно учитываются процессы инфляции и изменение численности необходимой рабочей силы, причем и в том и в другом случае предполагается экспоненциальный рост.

Без учета бюджетной политики модель выглядит так:

$$C = (1 - s)Y, \tag{16.21}$$

$$\frac{DK}{K} = \gamma \log\left[\frac{pY - wL}{(1+c)rKp}\right], \tag{16.22}$$

$$DY = \gamma(C + DK - Y), \tag{16.23}$$

$$L = Be^{-\rho t}Y^{b}K^{1-b}, \tag{16.24}$$

$$\frac{Dw}{w} = \log\left(\frac{L}{L_s}\right)\beta + a, \tag{16.25}$$

$$p = (1+\pi)w\frac{dL}{dY} = \frac{b(1+\pi)wL}{Y}, \tag{16.26}$$

$$\frac{M_d}{p} = AY^{u}r^{-v}, \tag{16.27}$$

$$M_d = M_s, \tag{16.28}$$

$$L_s = L_0e^{lt}, \tag{16.29}$$

$$M_s = M_0e^{mt}, \tag{16.30}$$

где s — склонность к сбережениям; p — уровень цен; w — ставка заработной платы; L — численность используемой рабочей силы; r — норма процента; L_s — предложение труда; M_d — спрос на деньги; M_s — предложение денег; m — темп роста предложения денег.

Остальные переменные определены ранее при рассмотрении модели экономического цикла.

Уравнение (16.21) означает, что «доход = сбережение + потребление». Уравнение (16.22) — формула для нормы прироста основного капитала, аналогичная (16.7). Уравнение (16.23) означает, что рост производства равен избытку спроса. Уравнение (16.24) отражает тот факт, что количество рабочей силы, требуемой для выпуска одного и того же количества продукции, все время убывает благодаря научно-техническому прогрессу. Таким образом, это уравнение учитывает технический прогресс. Уравнение (16.25) описывает изменение цен на рынке труда. Уравнение (16.26) утверждает, что уровень цен равен предельным издержкам (издержки на рабочую силу wL, предельные издержки $\frac{d(wL)}{dY}$) плюс некоторая добавка. В уравнении (16.27) M_d — те активы, которые население желает держать в денежной форме. Реальный спрос на деньги $\frac{M_d}{p}$ тем выше, чем выше доход Y и ниже норма процента r. Уравнение (16.28) означает, что спрос на деньги равен предложению денег. Это возможно, если считать, что норма процента все время подстраивается так, чтобы выполнялось это равенство. В уравнении (16.29) зафиксировано, что предложение труда растет со временем. В уравнении (16.30) предполагается, что предложение денег растет со временем.

При решении этой системы выяснилось, что, как и раньше, чем больше s, тем стабильнее K и Y, но в отличие от модели экономического цикла, равновесные K и Y теперь растут при увеличении m — темпа роста и предложения денег.

Теперь отразим в модели экономическое регулирование. Существование денежной политики можно выразить заменой уравнения (16.30) на

$$\log M_s = \log \widehat{M} + \theta \log\left(\frac{\hat{L}e^{lt}}{L}\right), \quad (16.31)$$

где $\hat{L}, \widehat{M}, \theta$ — положительные константы; $\hat{L}e^{lt}$ — оптимальная траектория занятости; $\widehat{M}$ — оптимальное предложение денег при оптимальном уровне занятости.

Чтобы отразить существование государственных расходов и налогов, изменим в системе уравнений (16.21)—(16.29), (16.31) значения некоторых переменных: C — личное потребление и государственные расходы; K — сумма государственного и частного основного капиталов; sY — сумма частных и государственных сбережений.

Государственные сбережения — это налоги минус государственные расходы, поэтому, чтобы отразить налоги, сделаем s переменной величиной:

$$s = \varepsilon_1 + \varepsilon_2 \log\left(\frac{L}{\hat{L}e^{lt}}\right), \quad (16.32)$$

где $\varepsilon_1, \varepsilon_2$ — параметры бюджетной политики.

В параметре s учитываются: 1) отношение личного потребления к личному доходу; 2) отношение поступлений от налогов к доходу; 3) отношение текущих государственных расходов к поступлениям от налогов. Все это можно учесть с помощью параметров $\varepsilon_1, \varepsilon_2$, которые являются управляющими.

Модель межотраслевых взаимодействий. Рассмотрим типичную макроэкономическую модель открытого типа (незамкнутую) — модель межотраслевых взаимодействий. Ее формируют две группы математических зависимостей:

1) система уравнений — баланс объема производства каждого вида продукции и его распределение между потребителями (другими производителями и конечными потребителями);
2) система неравенств, которые описывают зависимость между производственными возможностями каждой отрасли и ограничивающими наличными ресурсами (основные фонды и живой труд).

В эту модель нужно ввести извне вектор Y — конечный продукт и учесть его деление на потребление, накопление, экспорт, государственные резервы, налоги. Далее, следует задать вектор F — производственные фонды и вектор L — ресурсы живого труда. Это означает, что «вокруг» модели межотраслевых взаимодействий необходимо построить модель доходов и потребления населения — для определения Y, модель формирования национального дохода — для определения F, модель «демография — трудовые ресурсы» для определения L и т.п., т.е. создать так называемый «макромодельный комплекс».

Макроэкономические модели можно условно разделить на два вида. Одни из них описывают, так сказать, типовую страну без привязки к ее конкретным особенностям. Другие предназначены для использования в конкретных условиях, описывают вполне определенную экономическую реальность. Рассмотрим модели экономики отдельных стран и мирового хозяйства в целом.

Модель влияния государственной финансовой политики на экономику США. В эту модель входят всего 6 переменных, она подходит для аналитического анализа и иллюстрации влияния правительственного фонда заработной платы, правительственного заказа, налога на деловую активность, на личное потребление, заработную плату частного сектора, прибыли, инвестиции, основной капитал и национальных доход.

В рассматриваемой модели переменные управления таковы:

$W2_t$ — правительственный фонд заработной платы на t-м отрезке времени;

G_t — правительственные заказы на t-м отрезке времени;

XT_t — налог на деловую активность.

Используются эндогенные (заданные извне) переменные:

C_t — потребление на t-м отрезке времени;

$W1_t$ — фонд заработной платы в частном секторе на t-м отрезке времени;

PP_t — прибыли на t-м отрезке времени;

I_t — инвестиции на t-м отрезке времени;

K_t — основной капитал в конце t-го отрезка времени;

Y_t — национальный доход на t-м отрезке времени.

В модель входят уравнения функционирования и тождества. Уравнения функционирования касаются потребления:

$$C_t = a_1 + a_2(W1_t + W2_t) + a_3PP_t + a_4PP_{t-1} + \mu_{1t};$$

инвестиций:

$$I_t = b_1 + b_2PP_t + b_3PP_{t-1} + b_4K_{t-1} + \mu_{2t};$$

и спроса на рабочую силу:

$$W1_t = c_1 + c_2(Y_t + TX_t - W2_t) + c_3(Y_{t-1} + TX_{t-1} - W2_{t-1}) + c_4T + \mu_{3t},$$

где μ_{it}, $i = 1, 2, 3$ — случайные возмущения.

Тождества имеют смысл балансовых соотношений (законов сохранения):

$$Y_t = C_t + I_t + G_t - TX_t;$$

$$PP_t = Y_t - (W1_t + W2_t);$$

$$K_t = K_{t-1} + I_t.$$

Таким образом, в уравнении потребления зафиксировано, что потребление зависит от заработной платы в частном и государственном секторах, от прибыли в настоящий и предшествующий периоды времени. В уравнении инвестиций принято, что инвестиции зависят от прибылей в настоящий и предшествующий периоды времени и от основного капитала в предшествующий период времени. Спрос на рабочую силу фактически зависит от прибыльности в настоящий и предшествующий периоды времени.

Коэффициенты в уравнениях и тождествах определяются путем анализа конкретных экономико-статистических данных эконометрическими методами.

Модель экономики США. Существует множество моделей экономики США. Рассмотрим сначала так называемую Уортонскую модель (фактически макромодельный комплекс). Эта модель содержит 734 соотношения, из них 292 уравнения поведения и 442 тождества. Модель составляют 8 блоков: 1) конечный спрос; 2) межотраслевые потоки; 3) потребность в трудовых ресурсах; 4) заработная плата; 5) цены производства; 6) цены конечного потребления; 7) прочие доходы; 8) финансы.

Используемые в модели сценарии состоят в том или ином изменении: 1) федеральных закупок товаров; 2) закупок товаров и услуг органами штатов и местного управления; 3) трансфертных платежей; 4) экспорта; 5) налога на инвестиции. Управляющими параметрами были следующие: 1) статьи расходов государственного бюджета; 2) ставки налогов; 3) цены и заработная плата; 4) курс доллара; 5) импортные пошлины.

Цель модели — оценка эффективности деятельности федерального правительства. Упрощенная схема этой модели была приведена ранее.

Рассмотрим более простую, нежели Уортонская, модель, содержащую гораздо меньше уравнений, однако хорошо иллюстрирующую принципы построения моделей рассматриваемого типа.

Сначала выделяются блоки, из которых будет состоять модель, затем перечисляются переменные, которые входят в модель (их 35). Формируется таблица объясняемых переменных и объясняющих факторов. На основании этой таблицы строится система уравнений. Например, по таблице находим, что основной капитал X_1 зависит: 1) от основного капитала в предшествующий период времени $(X_1)_{-1}$; 2) инвестиций производственного назначения в предшествующий момент времени $(X_{16})_{-1}$; 3) краткосрочного процента в предшествующий момент времени $(X_{12})_{-1}$; 4) возмещения выбытия фондов $(X_{20})_{-1}$; 5) занятости в частном секторе $(X_2)_{-1}$. Теперь строим линейное регрессионное уравнение с авторегрессионным членом:

$$X_1 = a_1 + a_2(X_1)_{-1} + a_3(X_{16})_{-1} + a_4(X_{12})_{-1} + a_5(X_{20})_{-1} + a_6(X_2)_{-1}.$$

Сложный вопрос состоит в выборе тех переменных, от которых зависит X_1. Он решается с помощью того или иного алгоритма нахождения «информативного подмножества переменных» в регрессионном анализе. Используются парные и множественные коэффициенты линейной или непараметрической корреляции.

Модель мирового хозяйства. Рассмотрим проект ЛИНК, который разработан в 1970-х гг. Уортонской ассоциацией эконометрических прогнозов под руководством нобелевского лауреата по экономике Л. Клейна.

Макромодельный комплекс ЛИНК — совокупность разрабатываемых независимо друг от друга, различных по размерам и структуре эконометрических моделей национальной экономики ряда стран и регионов, которые увязываются в единую систему посредством субмодели мировой торговли.

В систему ЛИНК включены:

1) модель экономики США — 207 уравнений;
2) модель экономики Канады — 183 уравнения;
3) модель экономики Франции — 32 уравнения;
4) модель экономики ФРГ — 137 уравнений;
5)—13) — (модели национальных экономик Великобритании, Италии, Швеции, Финляндии, Бельгии, Нидерландов, Австрии, Японии, Австралии);
14) единая модель экономики развивающихся стран;
15) единая модель экономики социалистических стран;
16) единая модель экономики стран остального мира.

Модель для каждой страны (группы стран) разрабатывалась независимо. Сначала модели опробовались для каждой страны (группы стран) отдельно. Потом все эти модели объединялись в мировую модель посредством модели мировой торговли.

Модели развитых стран содержали блоки: 1) производства; 2) потребления; 3) инвестиций; 4) доходов и занятости; 5) цен; 6) денежного обращения; 7) внешней торговли.

Каждая страна описывалась с помощью моделей верхнего и нижнего уровня.

Верхний уровень состоит из перечисленных блоков. Далее каждый блок раскрывается. Например, в блок денежного обращения включены параметры: 6.1) количество денег в обращении; 6.2) дефицит бюджета; 6.3) сальдо платежного баланса; 6.4) индекс цен (дефлятор ВНП); 6.5) индекс розничных цен; 6.6) индекс оптовых цен; 6.7) учетная ставка по долгосрочным кредитам; 6.8) учетная ставка по краткосрочным кредитам. Модели верхнего уровня содержат взаимосвязи между этими параметрами.

Нижний уровень модельного комплекса содержит детализированные модели, описывающие регионально-страновые и проблемно-функциональные отношения.

С помощью системы ЛИНК были выявлены нетривиальные экономические связи. Например, оказалось, что снижение налогов в США приводит к улучшению платежного баланса Франции.

Модель мировой торговли. Рассмотрим моделирование товарных потоков между парами стран. Для этого используются, например, гравитационные методы, приводящие к соотношениям:

$$E_{ij}^t = \left(Z_{ij}^t; Z_j^t; R_{ij}^t\right),$$

где E_{ij}^t — экспорт из страны i в страну j в интервал времени t; Z_{ij}^t — факторы, определяющие потенциальное предложение экспорта страной i для страны j в интервал времени t; Z_j^t — факторы, определяющие потенциальный спрос страны j на импорт в интервал времени t; R_{ij}^t — факторы, относящиеся к продвижению товарного потока из страны i в страну j в интервал времени t.

С помощью этой и других моделей независимо разработанные модели отдельных стран можно увязать в единую мировую макромодель.

Вопросам построения, изучения и использования макроэкономических моделей при разработке и принятии управленческих решений посвящено огромное количество литературы.

В системы поддержки принятия решений входят не только общие макроэкономические модели, но и модели, касающиеся отдельных сторон функционирования народного хозяйства, в частности модели налогообложения. Обсудим некоторые из них [89].

Модель вычетов при налогообложении прибыли. Математические модели налогообложения, используемые в зарубежных странах,

весьма разнообразны. Начнем с канадской модели Т2. Она посвящена моделированию изменения нормы вычетов из налоговых обязательств затрат капитальных активов (при уплате налогов на прибыль). Предполагается, что каждая фирма самостоятельно проводит максимизацию скидок и минимизацию налоговых сборов (в рамках действующего налогового законодательства).

Анализируются изменения в первый год после управляющего воздействия и в «зрелой» системе через большой промежуток времени. Любопытно, что управляющим воздействием является не изменение ставки налога, а изменение правил расчета амортизационных начислений, причем это изменение касается лишь вновь приобретаемых единиц основных фондов (поэтому новые ставки амортизации лишь постепенно распространяются на налоговую базу).

В модели «зрелой» системы налогообложения используются параметры:

- средняя прошлая норма прироста капитала,
- показатель экспоненциальной амортизации,
- индекс цен капитала (учитывающий, например, рост во времени стоимости участков земли в неизменных ценах),
- средний коэффициент (индекс) инфляции (для народного хозяйства в целом),
- норма вычетов из налоговой базы затрат капитальных активов.

Моделируется также влияние на налоговые поступления изменения ставки зачета налога на инвестиции.

Модель построена на основе анализа данных, приведенных в выборке налоговых деклараций 15 тыс. из 760 тыс. фирм Канады.

Модели поступлений от налога на добавленную стоимость. В Румынии и Венгрии для оценки суммарных поступлений от налога на добавленную стоимость сначала оценивают налоговую базу на основе макроэкономических показателей. Считают, что она равна:

(валовой внутренний продукт) + (импорт) –
– (экспорт) – (фиксированные капиталовложения) – (изменение
запасов) – (добавленная стоимость по освобожденным
от налога секторам) – (оценка НДС для малого бизнеса
и строительства частного жилья — до 1992 г.).

Прогноз на следующий год осуществляется умножением налоговой базы предыдущего года на коэффициент, равный сумме прогнозов индекса инфляции и экономического роста за следующий год.

Действительная ставка налоговых поступлений от НДС рассчитывается путем деления объема чистых поступлений на налоговую базу. Для целей прогноза налоговая ставка считается постоянной (либо прогнозируется с помощью теории временны́х рядов).

Прогноз объема налоговых поступлений получают перемножением прогнозов объема налоговой базы и прогноза действительной ставки поступлений от НДС.

В Венгрии в Министерстве финансов разработана модель налоговой базы и поступлений от налога на добавленную стоимость на основе межотраслевой модели «вход—выход» (21 отрасль).

Модель подоходного налога в Великобритании построена на основе репрезентативной выборки, включающей 80 тыс. налогоплательщиков (физических лиц) из 25 млн плательщиков подоходного налога. Используются данные налоговых деклараций.

С использованием соображений демографии, социологии, медицинской статистики и макроэкономики прогнозируется изменение налоговой базы, при этом структура модели определяется экспертами из представителей перечисленных наук, а коэффициенты оцениваются по выборочным данным.

Знание налоговой базы позволяет прогнозировать первоначальное (в первый год) изменение налоговых сборов при применении управляющих воздействий. Для оценки дальнейшей динамики необходимо учитывать реакцию налогоплательщиков на изменение системы налогообложения (в первом приближении — линейный отклик с коэффициентами эластичности в качестве множителей перед приращениями переменных). Прогнозирование на далекую перспективу возможно лишь методом сценариев, поскольку необходимо спрогнозировать, в частности, динамику народонаселения.

Моделирование процессов налогообложения в России. Разработка имитационных моделей процессов налогообложения с целью оценки влияния управляющих воздействий на эти процессы, сбора и обобщения информации о процессах налогообложения на основе компьютерных систем представляет собой достаточно наукоемкую и трудоемкую задачу [89]. Основные задачи, которые необходимо решить при разработке подобной модели, таковы:

- анализ нормативной базы и практической реализации процессов налогообложения;
- постановка основных задач оценки управляющих воздействий на процессы налогообложения;
- разработка и изучение системы математических моделей, имитирующих процессы налогообложения в реально действующей налоговой системе;
- решение тех же задач для будущей, модифицированной, согласно решениям государственной власти, налоговой системы;

- разработка диалоговой компьютерной системы и соответствующих программных средств, позволяющих сотрудникам налоговых служб решать стоящие перед ними задачи оценки управляющих воздействий на процессы налогообложения.

Целесообразна разработка моделей для анализа предлагаемых различными организациями и лицами налоговых систем, а также для оценки влияния процессов налогообложения на статику и динамику микро- и макроэкономических характеристик.

Сформулируем основные требования к выполнению подобного исследования в российских условиях:

- работа должна основываться на анализе действующей системы сбора налогов и других обязательных платежей в бюджетную систему Российской Федерации (в федеральный бюджет и в бюджеты территорий);
- математические модели и соответствующие компьютерные разработки, предназначенные для оценки управляющих воздействий на процессы налогообложения, должны позволять рассчитывать объемы налоговых поступлений при тех или иных значениях управляющих воздействий — ставок налогов, льгот, штрафов;
- модели должны предоставлять возможности для анализа модификаций налоговой системы (в частности, путем изменения ставок, системы льгот и штрафов, правил относительно времени внесения платежей, а также введения новых видов налогов);
- конечный программный продукт должен предназначаться для эксплуатации специалистами государственной налоговой службы, не имеющими специальных знаний в программировании и математическом моделировании.

Проблемам моделирования процессов налогообложения в России посвящена монография [89].

16.2. ПРИНЯТИЕ РЕШЕНИЙ В МАЛОМ БИЗНЕСЕ

При принятии решений на уровне предприятия, в том числе малого, весьма полезны соответствующие экономико-математические и эконометрические модели. Рассмотрим несколько примеров.

Модель функционирования промышленного предприятия. Рассмотрим модель предприятия, являющегося частью более крупной экономической структуры (системы) — государственного сектора экономики, финансово-промышленной группы, транснациональной корпорации, холдинга и т.п. Предприятие действует в плановом сегменте экономики, план определяется вышестоящими органами управления.

Основное внимание в модели предприятия уделяется влиянию фондов экономического стимулирования — фонда развития производства (ФРП) и фонда материального поощрения (ФМП) — на темпы роста прибыли.

Наполняются ФРП и ФМП из прибыли, которую получает предприятие. Основная часть этой прибыли идет в бюджет вышестоящей структуры, оставшаяся — в фонды. Сверху вводятся нормативы распределения прибыли, т.е. доли от прибыли, которые идут в фонды.

Важно найти оптимальные величины этих нормативов, так как если нормативы будут малы, то фонды практически перестанут зависеть от темпа роста прибыли и рентабельности, как следствие, их воздействие на деятельность предприятия окажется минимальным. Фактически тут можно говорить о чрезмерно больших изъятиях средств вышестоящей структурой. С другой стороны, если в фонды идет слишком большая часть прибыли, это может привести к дефициту бюджета вышестоящей структуры.

Размер ФРП

$$\Phi = f(A, I, R, \alpha_1, \alpha_2),$$

где A — стоимость основных производственных фондов; I — отношение $\frac{P_ф}{P_п}$ (здесь $P_ф$, $P_п$ — фактический и плановый объемы реализованной продукции); R — рентабельность; α_1, α_2 — отраслевые нормативы отчисления части прибыли соответственно в ФРП и ФМП.

Предположим, что вышестоящей структурой выделяются капитальные вложения в количестве, прямо пропорциональном объему произведенной продукции P. Пусть:

$$\frac{dA}{dt} = \alpha\frac{dA}{dt} + \gamma P.$$

Это уравнение означает, что полный прирост капитальных вложений в основные производственные фонды равен той доле, которая выделяется на это из ФРП $\left(\alpha\frac{dA}{dt}\right)$ плюс вложения вышестоящей экономической структуры (γP). Коэффициент γ определяется вышестоящим управляющим органом; величина $\alpha\frac{dA}{dt}$ зависит от величины ФРП, т.е. в конечном итоге от принятых в системе нормативов.

Примем, что производственная функция пропорциональна стоимости основных фондов, т.е. $P = kA$. Тогда

$$\frac{dA}{dt} = \frac{\gamma kA}{1-\alpha}.$$

Решение этого уравнения $A(t) = A(t_0)\exp\left[\frac{\gamma k(t-t_0)}{1-\alpha}\right]$ описывает рост стоимости основных фондов. При этом пропорционально растет и объем произведенной продукции P.

Необходимо иметь современные научно-экономические инструменты анализа такого сравнительно нового для нашей страны вида хозяйственных объединений, как *холдинги*. Ясно, что без предварительного научного анализа невозможно выработать обоснованные рекомендации по повышению эффективности оперативного финансового управления компаниями холдингового типа.

Малые предприятия. Малое предпринимательство — важная составная часть современной российской экономики. Например, в Москве более 10% населения трудится на малых предприятиях. Поэтому весьма актуальным является изучение сферы малого бизнеса с позиций экономической теории, в частности методами экономико-математического моделирования [22].

Развитие малого предпринимательства необходимо для эффективного функционирования экономики России. Для понимания особенностей этого развития и управления им могут оказаться полезными разнообразные экономико-математические модели. Рассмотрим подходы к построению и изучению некоторых из них, дадим широкую панораму возможных подходов к построению моделей, которые могут оказаться полезными для описания динамики развития малых предприятий, а также и управления ими.

Проблемы маркетинга малого бизнеса. Во всех странах с развитой рыночной экономикой нестабильность малого бизнеса во многом связана с его сильной зависимостью от внешней среды — как от СТЭЭП-факторов (социальных, технологических, экологических, экономических, политических), так и от факторов конкурентного окружения (в том числе от поставщиков и потребителей). Для того чтобы выжить и занять свою рыночную нишу, малый бизнес должен хорошо ориентироваться и адаптироваться в условиях достаточно высокой степени неопределенности и риска. Это означает, что маркетинг малого бизнеса изначально носит рисковый характер.

Для снижения степени риска маркетинга малого бизнеса требуется высокий профессионализм менеджмента малой организации в области управления рыночной информацией и быстрота реакции в принятии решений при изменении условий внешней среды. То есть как лицо, принимающее решения (ЛПР), менеджер малой организации должен быть одновременно хорошим маркетологом.

Маркетинг малого бизнеса имеет особенности. Для того чтобы малая организация могла выжить и занять свою рыночную нишу, ее маркетинг с самого начала должен быть ориентирован не на абстрактные производство и сбыт, а на конкретного потребителя с его индивидуальными запросами. Иными словами, приоритетной формой маркетинга малого бизнеса является целевой специализированный маркетинг. Он позволяет сконцентрировать объективно небольшие ресурсы малой организации на наиболее важном направлении. Однако цена ошибки ЛПР, цена принятия неправильного решения в малом бизнесе многократно возрастает, так как у малой организации, как правило, нет финансовых возможностей диверсифицировать свою деятельность и свой риск.

Следовательно, для менеджера малой организации наиболее важные и сложные задачи таковы: проведение маркетинговых исследований по изучению рынка, сегментация рынка, выбор целевого сегмента, оценка его потенциальной мощности, оценка риска выбора рыночной ниши и силы потенциальных конкурентов. Успешное решение перечисленных задач требует от менеджера малой организации достаточно серьезной подготовки в области теории принятия решений, эконометрики и экономико-математического моделирования, поскольку оплата услуг консалтинговых фирм по этим вопросам стоит достаточно дорого.

Вместе с тем для того, чтобы быстро реагировать на изменения внешней среды, оказывающей сильное воздействие на малую организацию, ее менеджер должен проводить постоянный мониторинг рыночной ситуации по определенным наиболее значимым параметрам (спрос, предложение, цены, товары-конкуренты, альтернативные технологии и др.). Сбор и оперативное использование такой информации является решающим фактором успеха в маркетинге малого бизнеса при принятии решений и требует определенных знаний и навыков у менеджера по формированию банка данных и работе с маркетинговой информацией. Наиболее доступными для менеджеров малого бизнеса являются экономико-статистические (эконометрические) методы и методы математического моделирования, позволяющие (при определенной подготовке менеджеров и наличии программной поддержки) достаточно быстро обрабатывать и использовать оперативную информацию на практике.

Математические методы и модели для решения задач малого бизнеса. Достаточно известными примерами применения методов экономико-математического моделирования в маркетинге для структурирования и анализа рыночной информации являются модели жизненного цикла товара (фирмы), модели маркетингового комплекса $4p$ ($7p$), матрица «Бостон-консалтинг групп», SWOT-анализ конкурентов, матрица определения проблемы и др. Они могут быть простейшими инструментами управления маркетингом в малом бизнесе и позволяют достаточно оперативно оценить место и конкурентные преимущества организаций. Вместе с тем возможности экономико-математического моделирования позволяют менеджеру самостоятельно структурировать свою собственную ситуацию и создавать собственные модели (или варианты типовых моделей с собственными значениями параметров) оптимального поведения на рынке в условиях неопределенности и риска. Так, известная среди маркетологов и менеджеров матрица «Бостон-консалтинг групп» является, на наш взгляд, не двухмерной, а трехмерной моделью, в которой наряду с долей на рынке и темпом роста продаж обязательно должен рассматриваться такой параметр, как прибыль организации.

При разработке системы экономико-математической поддержки малого бизнеса математические модели развития малого предпринимательства должны изучаться специалистами теоретически на основе вероятностных и имитационных методов и сопоставляться со статистическими данными, характеризующими реальное положение в рассматриваемой области экономики. Методология математического моделирования позволяет ставить и решать различные задачи, возникающие в маркетинге малого бизнеса. В частности, отметим задачи анализа и прогнозирования рыночной ситуации, оценки различных видов рисков.

Для повышения эффективности исследовательской работы целесообразно разделять экономико-статистические (эконометрические) методы и экономико-математическое моделирование, хотя такое деление и условно. Примером первых (т.е. методов прикладной статистики применительно к конкретным экономическим данным) являются методы выборочного изучения потребителей. Так, в 1994 г. сотрудниками Института высоких статистических технологий и эконометрики были опрошены 500 потребителей и продавцов растворимого кофе; полученные результаты использованы фирмой-заказчиком при маркетинге, в частности при планировании рекламной кампании [89]. Технология проведения таких маркетинговых исследований близка к технологии социологических опросов, а также имеет много общего со статистическим управлением качества продукции, в частности с оценкой качества при сертификации [81].

При экономико-математическом моделировании применяют нацеленные на конкретные применения модели, в отличие от моделей прикладной статистики, которые можно использовать в любой сфере деятельности. Примерами являются экономико-математические модели управления запасами (см. далее), с помощью которых удается находить оптимальные размеры поставок и процедуру их поступления. Обычно применение таких моделей позволяет, по крайней мере, вдвое сократить суммарные издержки. Набор подобных компьютерных моделей должен быть рабочим инструментом менеджера малого предприятия.

При математическом моделировании маркетинговых проблем малого бизнеса используют эконометрические методы и методы экспертных оценок, а также методы имитационного моделирования. В настоящее время быстрых перемен в социальной, экономической и политической сферах отсутствуют достаточно длинные временны́е ряды экономических данных и интерес исследователей и практических работников переместился из статистики временны́х рядов в области теории и практики экспертных оценок.

В маркетинговых исследованиях для малого бизнеса большую роль играют факторы нечисловой природы — качественные признаки, интервальные и нечеткие оценки и др. Развиваются и применяются современные методы статистического анализа нечисловых данных [89]. Оригинальность и эффективность математического аппарата в области статистики нечисловых данных определяется тем, что он основан на использовании расстояний в выборочных пространствах, а не операций суммирования.

При изучении экономических рисков, в частности связанных с осуществлением инвестиционных проектов, необходимо моделировать различные неопределенности будущего и настоящего. Неопределенность описывают с помощью вероятностно-статистических, нечетких, в частности интервальных, моделей. Вероятностно-статистические модели нацелены прежде всего на анализ массовых явлений. Неопределенность единичных событий более целесообразно описывать с помощью нечетких множеств, в частности с помощью интервальных чисел, задающих нижние и верхние границы для неизвестных в точности параметров. Около 30 лет назад доказано [88], что теория нечетких множеств в определенном смысле сводится к теории случайных множеств. Однако при практическом применении математический аппарат теории нечеткости существенно отличается от вероятностно-статистического инструментария, а также и от аппарата статистики интервальных данных [89].

При применении математических моделей весьма важным является исследование устойчивости выводов по отношению к допустимым

отклонениям исходных данных и предпосылок модели [88]. Только та модель может быть рекомендована для практического использования, для которой полученные с ее помощью выводы мало меняются при подобных отклонениях. Накоплен определенный опыт применения методологии экономико-математического моделирования при решении практических задач маркетинга малого бизнеса [89], в частности в области товаров народного потребления и производственного назначения, образовательных услуг, а также при анализе и моделировании инфляционных процессов, в сфере налогообложения и др.

Перейдем к более подробному рассмотрению некоторых экономико-математических моделей, предназначенных для описания маркетинговой деятельности и жизненного цикла предприятий малого бизнеса.

Маркетинговые модели принятия решений. Для структурирования и анализа рыночной информации могут быть успешно применены такие известные инструменты принятия управленческих решений, как SWOT-анализ и матрица «Бостон консалтинг групп», а также некоторые их обобщения. Эти обобщения позволяют эффективно использовать современные методы экспертного оценивания, в том числе основанные на применении статистики нечисловых, в частности интервальных данных.

В обобщении SWOT-анализа предприятия оцениваются (в количественных или в качественных шкалах) по четырем группам показателей — сильные и слабые стороны, угрозы и возможности. Частные показатели сводятся в групповые, а групповые — в итоговый (обобщенный) показатель. Эта процедура дает возможность ранжировать и классифицировать конкурентов (например, на весьма опасных, опасных и неопасных), а также отслеживать и моделировать динамику показателей и итоговых оценок предприятий.

В обобщенной матрице «Бостон-консалтинг групп» используют трехмерную модель, в которой предприятие описывается долей на рынке, темпом роста продаж и прибылью. От качественных значений перечисленных переменных переходим к количественным, а также строим итоговый показатель и прогностические правила.

Рассматриваемые модели основаны на применении технологии построения единичных, групповых и обобщенных показателей (оценок отдельных сторон деятельности фирм-конкурентов и их экономического положения в целом), развитой ранее для решения задач экологического страхования. Компьютерная поддержка этой технологии может быть осуществлена с помощью автоматизированного рабочего места организатора экспертного опроса АРМ МАТЭК (МАТематические методы в ЭКспертных исследованиях).

Как уже говорилось, экспертные оценки как самостоятельное направление научно-практической деятельности развивается в нашей стране с 1970-х гг. В частности, с 1973 г. работает неформальный научный коллектив вокруг научного семинара «Математические методы экспертных оценок и нечисловая статистика», часто обращающийся к проблемам принятия решений в условиях малого бизнеса. Проведена масса исследований, опубликованы десятки монографий и сборников, сотни статей.

В настоящее время возникла масса аналитических центров, бизнес-инкубаторов и других учреждений, которым рассматриваемые разработки явно полезны. Однако важно установить контакты между теоретиками и менеджерами аналитических центров, наладить систему обучения. Накопленные теоретиками знания должны быть основой для компьютерных систем, например таких, как АРМ МАТЭК.

О теории ранжировок и рейтингов. В настоящее время в практике работы малых предприятий распространены маркетинговые, экспертные и социологические опросы. При их проведении опрашиваемых просят выставить баллы инвестиционным проектам, направлениям работ или исследований, товарам, идеям, проблемам, программам или политикам. Затем рассчитывают средние арифметические баллов и рассматривают их как интегральные оценки, выставленные фирмой или обществом в целом инвестиционным проектам, направлениям работ или исследований, товарам, идеям, проблемам, программам или политикам. Но уже около 20 лет известно, что согласно теории измерений такой способ расчета интегральных оценок некорректен (см. главу 9).

Хорошо известно, каким условиям должны удовлетворять методы обработки данных, измеренных в тех или иных шкалах. Например, для порядковых данных в качестве интегрального показателя использовать среднее арифметическое нельзя, а медиану — можно [89]. К сожалению, распространены некорректные методы расчетов. В качестве примера отметим, что методы расчета рейтингов «ведущих политиков» на основе усреднения ответов экспертов, публикуемые в «Независимой газете», являются математически некорректными. Впрочем, к этим публикациям есть много иных претензий связанных, в частности, с нерепрезентативным составом экспертов.

Как известно, максимальными инвариантами в порядковой шкале являются ранжировки (нестрогие порядки). Поэтому от использования результатов теории измерений менеджеру малой организации естественно перейти к применению методов статистики объектов нечисловой природы [89].

Моделирование потока проектов. Кратко рассмотрим несколько экономико-математических моделей, описывающих развитие малых предприятий в течение их жизненного цикла.

При построении математических моделей типа «поток проектов» будем считать, что малое предприятие ассоциируется с последовательностью выполняемых им проектов. Новые малые предприятия порождаются в соответствии с пуассоновским процессом переменной интенсивности (аналогично потоку заявок в теории массового обслуживания [15]). Каждое новое малое предприятие выполняет вначале один проект, величина (стоимость) и продолжительность которого — случайные величины с заданными (в модели) распределениями.

Точнее, с учетом известных в менеджменте представлений о жизненном цикле продукции экономический эффект (на единицу времени) от выполнения проекта описывается случайной функцией от времени (с отсчетом от момента начала осуществления проекта). Типовой вид этой функции таков: сначала отрицательные значения (в начале проекта необходимы капиталовложения), затем — рост до максимального значения, продолжительное «плато» на достигнутом уровне, переходящее в спад (окончание проекта). В модель порождения малых предприятий необходимо внести новую переменную — случайную величину начального капитала, которая, в частности, ограничивает круг проектов, возможных для данного малого предприятия. Возможно и разорение малого предприятия, если в силу случая стартовый капитал окажется недостаточным для осуществления проекта. Отметим, что потоки платежей необходимо оценивать путем приведения элементов этих потоков к сопоставимым ценам, а при этом не обойтись без учета инфляции, изучение и прогнозирование которой встречает известные трудности [89].

Однако для некоторых видов деятельности, например оказания научно-технических услуг, можно считать, что экономический эффект (в сопоставимых ценах) имеет простой частный вид — является ступенчатой функцией, равной положительной константе C на отрезке $[0, T]$ и 0 вне его (здесь C и T — случайные величины).

Поскольку каждый проект рано или поздно заканчивается, малое предприятие, как правило, должно переходить к осуществлению новых проектов еще до окончания жизненного цикла предшествующего проекта. В модели принимаем, что каждый проект порождает своих потомков — новые проекты с определенной интенсивностью. С этой точки зрения малое предприятие — это совокупность проектов, в которую входят: 1) исходный проект (если он еще продолжается); 2) его непосредственные потомки; 3) потомки его потомков и т.д. Развитие малого предприятия состоит в возникновении, выполнении и прекращении проектов, его образующих. Если все эти проекты прекращаются, то малое предприятие ликвидируется. Аналогом является развитие популяции фамилий, изучаемое с помощью теории ветвящихся процессов [112].

Рассматриваемые модели позволяют, в частности, изучать динамику распределения малых предприятий по размерам и длительности жизни, например оценивать долю предприятий, прекративших деятельность в течение определенного интервала времени после организации. Можно продемонстрировать положительную роль технопарков как «инкубаторов» малых предприятий, влияние экспертизы бизнес-плана и других планов — поддержка проектов на начальных стадиях при условии отсечения малоперспективных проектов существенно повышает вероятность «выживания» остальных.

Пример модели потока проектов. Приведенное ранее описание задает достаточно обширное семейство математических моделей. Рассмотрим одну из них.

Пусть процесс порождения новых предприятий в регионе описывается пуассоновским процессом с постоянной интенсивностью q. Это означает, что за единицу времени возникает случайное число X малых предприятий, причем X имеет пуассоновское распределение с параметром q. В среднем за единицу времени возникает q малых предприятий, поскольку математическое ожидание X равно q. Величина q зависит от числа жителей и уровня социально-экономического развития региона.

Следующий шаг — моделирование начального капитала и стоимости проекта. При этом в случае когда необходимые капиталовложения больше начального капитала, предприятие погибает, не приступив к деятельности. Хорошо известно, что в современной России большое число зарегистрированных малых предприятий (по крайней мере до 30%) не проявляют производственной активности. Все такие предприятия можно считать погибшими еще до начала выпуска продукции.

Рассмотрим предприятия с достаточным начальным капиталом. Пусть для простоты экономический эффект при выполнении проекта является ступенчатой функцией, равной положительной константе C на отрезке $[0, T]$ и 0 вне его, где C и T — случайные величины. Далее следует смоделировать процесс порождения «потомков» проекта. Естественно считать, что число потомков случайно, но при этом их в среднем больше у проекта большей стоимости и более длительного. Дальнейшее

опустим, поскольку основные идеи, лежащие в основе моделирования, уже сформулированы.

Модель занятия ниш. Предположим, что имеется конечный набор «ниш», которые могут занять вновь возникающие предприятия. В соответствии с некоторым распределением вероятностей порождаются новые предприятия (т.е. указываются для них ниши). Если ниша занята, то предприятие гибнет. Если нет — занимает нишу и функционирует некоторое случайное время, после чего прекращает деятельность и освобождает нишу. Действующее предприятие может захватывать свободные ниши на тех же основаниях, что и вновь возникающие предприятия. Нетрудно получить расчетные формулы для числа свободных ниш и вероятности того, что ниша занята, а также для иных характеристик, описывающих развитие популяции малых предприятий.

Модель выбора ниши. Для описания поведения малого предприятия предлагается использовать модель выбора ниши на основе теории принятия решений с использованием дерева целей. Рассматривая выбор на каждом этапе как случайную величину, получаем возможность расчета распределения малых предприятий по вариантам окончательных решений. А это порождает итерационный процесс пересмотра решений, поскольку знание итогового распределения влечет пересмотр некоторых из ранее принятых решений, например о числе возможных конкурентов. Модель целесообразно реализовать в виде имитационной компьютерной системы, пригодной также для индивидуального обучения и проведения деловых игр. Интересны варианты модели с использованием интервальных или нечетких ответов, что делает и итоговое решение интервальным или нечетким.

Проблемам малого предпринимательства посвящено большое число официальных и научных публикаций, что объясняется, очевидно, заметным вкладом малых предприятий в отечественное производство, а также — что представляется нам более важным — пионерской ролью малых предприятий в опробовании различных вариантов организации экономической жизни, взаимодействия государственных и негосударственных структур. Именно малые предприятия лучше всего демонстрируют роль конкуренции в экономике.

Итак, экономико-математическое моделирование имеет широкие перспективы практического применения при принятии решений в малом бизнесе. Еще более интересные возможности раскрываются в области теоретических исследований проблем малого бизнеса. Совместная работа экономистов, эконометриков, математиков и практикующих менеджеров малого бизнеса приносит пользу как теории, так и практике.

16.3. ПРИНЯТИЕ РЕШЕНИЙ В ЗАДАЧАХ ЛОГИСТИКИ

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, получившей свое

развитие, в основном начиная с 1950-х гг. Предложенная, видимо, еще в 1915 г. Ф. Харрисом классическая модель теории управления запасами, называемая также моделью Вильсона (в связи с тем, что получила известность после публикации работы Р.Г. Вильсона в 1934 г.), является одним из наиболее простых и наглядных примеров применения математического аппарата для принятия решений в экономической области. В то же время 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 [88]. 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 зубчатой ломаной линией (рис. 16.1), состоящей из наклонных и вертикальных звеньев, причем наклонные отрезки параллельны.

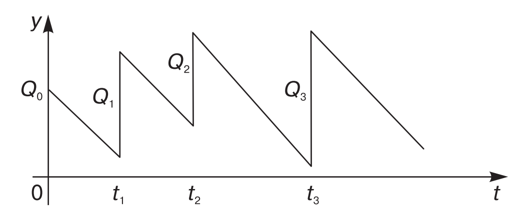


Рис. 16.1. График изменения величины запаса на складе

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 $n(T)$ — число 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

$$F(T;y) = F\left[y(t), 0 \le t < T\right] = gn(T) + s\int_0^T y(t)dt.$$

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 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:

$$f(T;y) = f\left[y(t), 0 \le t < T\right] = \frac{1}{T}F(T;y) = \frac{1}{T}\left\{gn(T) + s\int_0^T y(t)dt\right\}.$$

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 — фиксация функции $y = y(t)$, $0 \le t < T$, рассматриваемого вида (рис. 16.1) 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 $t_0 = 0, t_1, t_2, \ldots$ пополнения запаса на складе и размеры поставляемых партий товара $Q_0, Q_1, Q_2, \ldots$ 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).

Решение задачи оптимизации. 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 запас равен 0 (т.е. $y(t) = 0$) в моменты доставок очередных партий, назовем *напряженным*.

Утверждение 16.1. 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, 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 $y(t_1-) \ne 0$ (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 $(0; t_1)$, что видно на рис. 16.2.

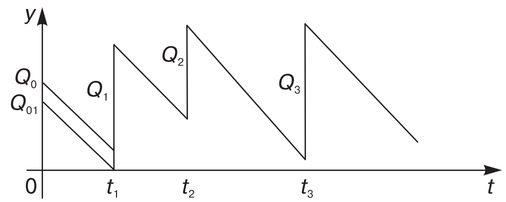


Рис. 16.2. Первый шаг перехода к напряженному плану

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 шаге 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 16.1 следует, что оптимальный план следует искать только среди напряженных планов. Другими словами, план, не являющийся напряженным, не может быть оптимальным.

Утверждение 16.2. 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:

$$Q_{i-1} = \mu(t_i - t_{i-1}),\ i = 1, 2, ..., n(T) - 1,\ Q_{n(T)-1} = \mu(T - t_{n(T)-1}).$$

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

$$s\int_0^T y(t)dt = s\sum_{i=1}^{n(T)} \frac{Q_{i-1}(t_i - t_{i-1})}{2} = s\sum_{i=1}^{n(T)} \frac{\mu(t_i - t_{i-1})^2}{2} = s\sum_{i=1}^{n(T)} \frac{\mu\Delta_i^2}{2} = \frac{\mu s}{2}\sum_{i=1}^{n(T)} \Delta_i^2,$$

где $\Delta_i = t_i - t_{i-1}$, $i = 1, 2, ..., n(T)$, $t_{n(T)} = T$.

Ясно, что Δ_i, $i = 1, 2, ..., n(T)$ — произвольные неотрицательные числа, в сумме составляющие T. Следовательно, для минимизации издержек среди напряженных планов с фиксированным числом поставок достаточно решить задачу оптимизации

$$\begin{cases} \Delta_1^2 + \Delta_2^2 + \ldots + \Delta_n^2 \to \min, \\ \Delta_1 + \Delta_2 + \ldots + \Delta_n = T, \\ \Delta_i \ge 0,\ i = 1, 2, \ldots, n, \end{cases}$$

где $n = n(T)$.

Полученная задача оптимизации формально никак не связана с логистикой, она является чисто математической. Для ее решения целесообразно ввести новые переменные $\alpha_i = \Delta_i - \dfrac{T}{n}, i = 1, 2, \ldots, n.$ Тогда

$$\sum_{i=1}^{n}\alpha_i = \sum_{i=1}^{n}\left(\Delta_i - \frac{T}{n}\right) = \left(\sum_{i=1}^{n}\Delta_i\right) - n\frac{T}{n} = T - T = 0\ .$$

Поскольку $\Delta_i = \dfrac{T}{n} + \alpha_i$, то $\Delta_i^2 = \dfrac{T^2}{n^2} + 2\dfrac{T}{n}\alpha_i + \alpha_i^2$, 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

$$\Delta_i = \frac{T}{n}, i = 1, 2 \ldots, n.$$

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:

$$f(T;y) = g\frac{n(T)}{T} + \mu s\frac{T}{2n(T)} = f_1(Q) = \frac{\mu g}{Q} + \frac{sQ}{2}. \qquad (16.33)$$

Задача состоит в минимизации $f_1(Q)$ по Q. При этом возможная величина поставки принимает дискретные значения, $Q \in \left\{\frac{\mu T}{n}, n = 1, 2, \ldots\right\}$.

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

$$\frac{df_1(Q)}{dQ} = -\frac{\mu g}{Q^2} + \frac{s}{2}.$$

Производная монотонно возрастает, поэтому рассматриваемая функция имеет единственный минимум в точке, в которой производная равна 0, т.е. при

$$Q_0 = 0\sqrt{\frac{2\mu g}{s}}. \qquad (16.34)$$

Получена знаменитая «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

$$Q \notin \left\{\frac{\mu T}{n}, n = 1, 2, \ldots\right\}.$$

Всегда можно указать неотрицательное целое число n такое, что

$$Q_1 = \frac{\mu T}{n+1} < Q_0 \le \frac{\mu T}{n} = Q_2. \qquad (16.35)$$

Утверждение 16.3. 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 $Q \in \left\{\frac{\mu T}{n}, n = 1, 2, \ldots\right\}$ 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 функции $f_1(Q)$ 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 $Q \in \left\{\frac{\mu T}{n}, n = 1, 2, \ldots\right\} \cap \left\{Q : Q \ge Q_0\right\}$ достигается при $Q = Q_2$, а минимум по $Q \in \left\{\frac{\mu T}{n}, n = 1, 2, \ldots\right\} \cap \left\{Q : Q < Q_0\right\}$ — при $Q = Q_1$. Последнее утверждение эквивалентно заключению утверждения 16.3.

Итак, алгоритм построения оптимального плана таков:

1. Найти Q_0 по формуле квадратного корня (16.34).
2. Найти n из условия (16.35).
3. Рассчитать $f_1(Q)$ по формуле (16.33) для $Q = Q_1$ и $Q = Q_2$, где Q_1 и Q_2 определены в (16.35).
4. Наименьшее из двух чисел $f_1(Q_1)$ и $f_1(Q_2)$ является искомым минимумом, а то из чисел Q_1 и Q_2, на котором достигается минимум, — решением задачи оптимизации. Обозначим его Q_{opt}.

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 существуют два оптимальных плана.

Пример 16.1. 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 $\mu = 5$ (т/день), $s = 50$ (руб./т·день), $g = 980$ (руб./партия), $T = 10$ (дней). По формуле (16.34) рассчитываем

$$Q_0 = \sqrt{\frac{2\mu g}{s}} = \sqrt{\frac{2\times5\times980}{50}} = \sqrt{196} = 14.$$

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

$$f_1(Q) = \frac{5\times980}{Q} + \frac{50Q}{2} = \frac{4900}{Q} + 25Q,$$

то

$$f_1(Q_1) = f_1(12{,}5) = \frac{4900}{12{,}5} + 25\times12{,}5 = 392 + 312{,}5 = 704{,}5$$

и

$$f_1(Q_2) = f_1(50/3) = \frac{4900\times3}{50} + 25\times\frac{50}{3} = 294 + 416{,}67 = 710{,}67.$$

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 придти 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 неполного. К моменту $T = 10$ пройдет $10 - 8{,}4 = 1{,}6$ дня с момента последней поставки, значит, со склада будет извлечено $5 \times 1{,}6 = 8$ 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 $\frac{14\times2{,}8}{2} = 19{,}6$, площадь 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 $\frac{(14+6)\times1{,}6}{2} = 16.$

Следовательно, площадь под графиком равна $58{,}8 + 16 = 74{,}8$, а плата за хранение составляет $50 \times 74{,}8 = 3740$ 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 $4 \times 980 = 3920$ руб. Общие издержки за 10 дней составляют $3740 + 3920 = 7660$ 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

$$f_1(Q_0) = \frac{4900}{Q_0} + 25Q_0 = \frac{7900}{14} + 25\times14 = 350 + 350 = 700,$$

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,

$$\frac{\mu g}{Q_0} = \frac{\mu g}{\sqrt{\frac{2\mu g}{s}}} = \sqrt{\frac{\mu g s}{2}}; \quad \frac{sQ_0}{2} = \frac{s\sqrt{2\mu g/s}}{2} = \sqrt{\frac{\mu g s}{2}}.$$

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

$$\frac{Q_0}{\mu} = \frac{\sqrt{\frac{2\mu g}{s}}}{\mu} = \sqrt{\frac{2g}{\mu s}}.$$

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. для

$$T = n\frac{Q_0}{\mu} = n\sqrt{\frac{2g}{\mu s}},\ n = 1, 2, \ \dots \qquad (16.36)$$

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 от 0 до T_0 оптимальный план меняется на всем интервале $[0; T_0]$.

Как происходит это изменение? При малых горизонтах планирования T делается лишь одна поставка (в момент времени $t = 0$), 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), и т.д.

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 (см. главу 3). Для решения проблемы горизонта планирования необходимо использование конкретной модели принятия решений, 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 числа 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* [17].

Ясно, что этот план не будет оптимальным для всех T, кроме заданных формулой (16.36). 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

$$\lim_{T\to\infty} \frac{f\left[T; y_{opt}(T)\right]}{f(T; y)} = 1,$$

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 16.1. План $y = y_0$ является асимптотически оптимальным.

Таким образом, для достаточно больших горизонтов планирования T планы $y_0(t)$, $0 \le t \le T$, все зубцы у которых имеют высоту Q_0, имеют издержки, приближающиеся к минимальным. Следовательно, эти планы

Вильсона, являющиеся сужениями одной и той же функции $y : [0; +\infty) \to \to [0; +\infty)$ на интервалы $[0; T)$ при различных T, можно использовать одновременно при всех достаточно больших T.

Замечание. Согласно [88] 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

$$\frac{f\left[T; y_{opt}(T)\right]}{f(T; y)} \le 1. \tag{16.37}$$

Найдем нижнюю границу для рассматриваемого отношения. При фиксированном T можно указать неотрицательное целое число n такое, что

$$\frac{nQ_0}{\mu} \le T < \frac{(n+1)Q_0}{\mu}.$$

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 $(0; T)$ плана, то очевидно, поскольку второй интервала — часть первого (или совпадает с ним), первые издержки больше вторых, т.е.

$$Tf\left[T; y_{opt}(T)\right] \ge \frac{nQ_0}{\mu} f\left(\frac{nQ_0}{\mu}; y_{opt}(T)\right).$$

Так как 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

$$\frac{nQ_0}{\mu} f\left(\frac{nQ_0}{\mu}; y_{opt}(T)\right) \ge \frac{nQ_0}{\mu} f\left(\frac{nQ_0}{\mu}; y_0(T)\right).$$

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

$$Tf\left[T; y_{opt}(T)\right] \ge \frac{nQ_0}{\mu}\sqrt{2\mu gs}. \tag{16.38}$$

Для общих издержек на интервалах $(0; T)$ и $[0; (n + 1)Q_0/\mu]$ при использовании плана y_0, очевидно, справедливо неравенство

$$Tf\left[T; y_{opt}(T)\right] \le \frac{(n+1)Q_0}{\mu} f\left(\frac{(n+1)Q_0}{\mu}; y_0(T)\right).$$

Следовательно,

$$Tf\left[T; y_0(T)\right] \le \frac{(n+1)Q_0}{\mu}\sqrt{2\mu gs}. \tag{16.39}$$

Из неравенств (16.38) и (16.39) вытекает, что

$$\frac{f\left[T; y_{opt}(T)\right]}{f(T; y_0)} \ge \frac{n}{n+1} = 1 - \frac{1}{n+1} \ge 1 - \frac{Q_0}{\mu T}.$$

Так как $\frac{Q_0}{\mu T} \to 0$ при $T \to \infty$, то, учитывая неравенство (16.37), из последнего неравенства выводим справедливость заключения теоремы 16.1. 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,

$$\lim_{\varepsilon\to\infty} \frac{f\left[T; y_0(T)\right]}{f\left[T; y_{opt}(T)\right]} = \lim_{\varepsilon\to\infty} \frac{f\left[t_k + \varepsilon; y_0(t_k + \varepsilon)\right]}{f\left[t_k + \varepsilon; y_{opt}(t_k + \varepsilon)\right]} = 1 + \frac{1}{2k}.$$

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 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 $\{1 + 1/(2k + 2)\}$ и т.д.

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 (16.34), 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 ранее, они имеют вид

$$f_1(Q) = \frac{\mu g}{Q} + \frac{sQ}{2},$$

где Q — объем партии.

Тогда

$$\frac{f_1(Q) - f_1(Q_0)}{f_1(Q_0)} = \frac{1}{2}\left(\frac{Q - Q_0}{Q}\right)\left(\frac{Q - Q_0}{Q_0}\right). \qquad (16.40)$$

Это тождество нетрудно проверить с помощью простых алгебраических преобразований.

Пример 16.2. Пусть используется план с $Q = 0{,}9Q_0$. Тогда

$$\frac{f_1(Q) - f_1(Q_0)}{f_1(Q_0)} = \frac{1}{2}\left(\frac{-0{,}1Q_0}{0{,}9Q}\right)\left(\frac{-0{,}1Q_0}{Q_0}\right) = \frac{0{,}01}{1{,}8} = 0{,}0056.$$

Таким образом, изменение объема партии на 10% привело к увеличению средних издержек лишь на 0,56%.

Пример 16.3. Пусть используемое значение объема поставки Q отличается от оптимального не более чем на 30%. На сколько могут возрасти издержки?

Из формулы (16.40) вытекает, что максимальное возрастание издержек будет в случае $Q = 0{,}7\,Q_0$. Тогда

$$\frac{f_1(Q) - f_1(Q_0)}{f_1(Q_0)} = \frac{1}{2}\left(\frac{-0{,}3Q_0}{0{,}7Q}\right)\left(\frac{-0{,}3Q_0}{Q_0}\right) = \frac{0{,}09}{1{,}4} = 0{,}0643.$$

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 30%) приводит к столь 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 и т.д.

Здесь следует сделать ***важное замечание****: 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

$$g(x) = g(x_0) + \frac{dg(x_0)}{dx}(x - x_0) + \frac{1}{2}\frac{d^2g(x_0)}{dx^2}(x - x_0)^2 + O\left[(x - x_0)\right].$$

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 $(x - x_0)^2$ — справедливо равенство

$$g(x) - g(x_0) = \frac{1}{2}\frac{d^2g(x_0)}{dx^2}(x - x_0)^2. \qquad (16.41)$$

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

$$C = \frac{1}{2}\frac{d^2 g(x_0)}{dx^2}.$$

Вернемся к классической модели управления запасами. Для нее надо рассматривать $f_1(Q)$ в роли $g(x)$. С помощью соотношения (16.41) 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

$$\frac{d^2 f_1(Q)}{dQ^2} = \frac{d}{dQ}\left(-\frac{\mu g}{Q^2} + \frac{s}{2}\right) = \frac{\mu g}{Q^3}.$$

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 (16.40) 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 [15; 89]. Ранее 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 (16.40)—(16.41) 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:

$$\Delta Q = \frac{\partial Q}{\partial \mu}\Delta\mu + \frac{\partial Q}{\partial g}\Delta g + \frac{\partial Q}{\partial s}\Delta s = \sqrt{\frac{g}{2\mu s}}\Delta\mu + \sqrt{\frac{\mu}{2gs}}\Delta g - \sqrt{\frac{\mu g}{2s^3}}\Delta s \quad (16.42)$$

(с точностью до бесконечно малых более высокого порядка).

Величину $\Delta\mu$ можно определить по фактическим данным о спросе, оценив величину отклонения реального спроса от линейного приближения [88], например, с помощью математического аппарата линейного регрессионного анализа [89]. Для определения значений параметров g и s необходимо проведение специальных 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 [88].

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. Согласно подходу [88] 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 формулы (16.42). После сокращения общего множителя получаем, что согласно принципу уравнивания погрешностей должно быть справедливо соотношение

$$\frac{|\Delta\mu|}{\mu}=\frac{|\Delta g|}{g}=\frac{|\Delta s|}{s}. \qquad (16.43)$$

Таким образом, относительные погрешности определения параметров модели должны совпадать.

В соотношении (16.43) используются истинные значения параметров, которые неизвестны. Поэтому целесообразно вначале вместо параметров использовать их грубые оценки, из соотношения (16.43) определить их примерную точность, затем провести исследования, уточняющие значения параметров. Эту процедуру естественно повторять 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, 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. 16.3.

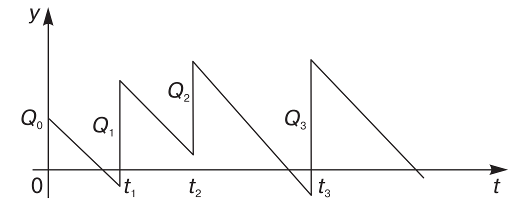


Рис. 16. 3. График изменения величины запаса на складе при возможности дефицита

Очевидно, рисунки 16.1 и 16.3 отличаются только тем, что на последнем рисунке зубцы графика могут опускаться ниже оси абсцисс, что соответствует сдвигу графика рис. 16.1 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

$$f_1(T,y)=f_1\left[y(t),0\le t\le T\right]=$$
$$=\frac{1}{T}\left\{s\int_0^T y(t)\chi\left[y(t)\ge 0\right]dt+h\int_0^T |y(t)|\chi\left[y(t)<0\right]dt+gn(T)\right\},$$

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. 16.4).

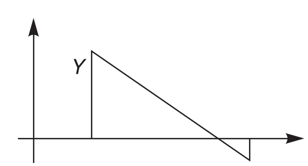


Рис. 16.4. Первый шаг построения оптимального плана в модели с дефицитом

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. 16.4 выражаются через Y, заданный интервал времени 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 этап 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 объема начальной поставки (в нулевой момент времени), равны между собой:

$$Q = Q_1 = Q_2 = Q_3 = \ldots = Q_0 = \frac{h}{s+h}Q. \qquad (16.44)$$

На третьем этапе среди указанного однопараметрического дискретного множества планов находим оптимальный план. 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, описываемый соотношением (16.44) 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 v, т.е. модель, в которой за время Δt поставляется продукция объемом $v\Delta t$ (при наличии в то же время постоянного спроса с интенсивностью μ, причем считается, что $v > \mu$). 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 v.

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, т.е. здесь используются только постоянные издержки. Представляется вполне естественным ввести линейный член, соответствующий возрастанию платы за доставку в зависимости от величины партии (переменные издержки) — далее 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 [87]. 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 ранее видов обобщений исходной модели.

При этом $a(1) = 0$, если поставки мгновенные, и $a(1) = 1$, если поставки являются линейными с интенсивностью υ, причем $v > \mu$.

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 ранее 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 $2 \times 2 \times 3 \times 3 = 36$ 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:

$$Q_0(\mu, v, s, a, h, b, g, g_1, g_2) = \sqrt{\frac{\mu g - \dfrac{(a-b)^2}{2(s+h)}\left(\dfrac{1}{1-\dfrac{\mu}{v}}\right)}{\dfrac{sh}{2(s+h)}\left(1-\dfrac{\mu}{v}\right)+\mu g_2}}.$$

В частности, план с $Q = Q_0(\mu, v, s, a, h, b, g, g_1, g_2)$ является асимптотически оптимальным.

Формула для $Q_0(\mu, v, s, a, h, b, g, g_1, g_2)$ 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. $v >> \mu$, 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

$$Q_0(\mu, +\infty, s, 0, h, 0, g, g_1, 0) = \sqrt{\frac{\mu g}{\dfrac{sh}{2(s+h)}}} = \sqrt{\frac{2\mu g(s+h)}{sh}}.$$

Эта формула уже была получена ранее 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.

Здесь важно отметить следующее: *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 расчетов, более сложные применяются для изучения точности, достигаемой с помощью более простых, согласно подходу, развитому в [15; 89].

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 Московской области. 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 о 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 заработную плату, начисления на нее); расходах на содержание охраны, эксплуатацию складских зданий и сооружений; расходах на текущий ремонт, тару, приемку, хранение, упаковку и реализацию товаров; величине амортизационных отчислений и др. Для расчета расходов на доставку новых партий товара использованы данные о расходах по завозу; плате за пользование вагонами и контейнерами сверхустановленных норм; расходах на содержание и эксплуатацию подъемно-транспортных механизмов; заработной плате работников, занятых в процессе доставки товара; канцелярских, почтовых и телеграфных расходах и др.

Полезным оказалось вытекающее из «принципа уравнивания погрешностей» соотношение (16.43). 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 Института материально-технического снабжения и по методике Центрального экономико-математического института. Для каждой из методик с помощью соотношения (16.43) были найдены 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 от 22,5 до 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.

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 [88].

Двухуровневая модель управления запасами. Создание любой 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). 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 ранее 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

$$f_1(T, y) = f_1\left[y(t), 0 \le t \le T\right] =$$

$$= \frac{1}{T}\left\{s\int_0^T y(t)\chi\left[y(t) \ge 0\right]dt + h\int_0^T |y(t)|\chi\left[y(t) < 0\right]dt + gn(T)\right\},$$

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 ранее.

Оптимизация состоит в определении значений нижнего уровня R и верхнего уровня $R + Q$, минимизирующих средние издержки.

В 1950-х годах 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 [88]. Этот принципиально важный теоретический результат стимулировал развитие исследований свойств двухуровневой модели. Однако окончательная теория была построена только в конце 1970-х гг.

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

$$X(T) = X_1 + X_2 + ... + X_{\tau(T)},$$

где X_j — величина j-й заявки.

Предполагается, что $X_1, X_2, ..., X_n, ...$ — 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 [88] 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 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 [87]. 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 $G(i, t)$, $0 \le t \le 1$, где t — отношение времени, прошедшего с начала i-го периода, к продолжительности его, $i = 1, 2, ..., m$.

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 $\delta(i, j)$, $j = 1, 2, ..., n(i)$, $i = 1, 2, ..., m$, с функциями распределения $F(i, j, t)$, $0 \le t \le 1$, 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, $i = 1, 2, ..., m$, все моменты поступления товара и требований $\tau(i)$, $\delta(i, j)$, $j = 1, 2, ..., n(i)$, 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 $s(i)$, $i = 1, 2, ..., m$. Издержки от дефицита складываются из потерь у каждого из потребителей; они пропорциональны величине и длительности дефицита с коэффициентами пропорциональности $h(i, j)$, $j = 1, 2, ..., n(i)$, $i = 1, 2, ..., m$.

Пусть $x(0)$ — начальный запас, $x(i)$ — количество продукта на базе в конце i-го периода, $i = 1, 2, ..., m$. Пусть $S(i) = \{s(i), c(j, i), h(i, j), G(i, t), F(i, j, t), 0 \le t \le 1, j = 1, 2, ..., n(i)\}$ — 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

$$Z(m) = f[x(0), x(1), S(1)] + f[x(1), x(2), S(2)] + ... + \\ + f[x(i-1), x(i), S(i)] + f[x(m-1), x(m), S(m)].$$

Необходимо минимизировать $Z(m) = Z[x(0), x(1), ..., x(i), ..., x(m)]$ 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 даны в [88].

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, $f_2[x(i-1), x(i), S(i)]$ — то же 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, что

$$f_1[x(i-1), x(i), S(i)] \equiv f_1[x(i-1), S(i)].$$

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 [88]. Ясно, что задача минимизации $Z(m)$ распадается на $m + 1$ задачу минимизации функций одной переменной:

$$f_1[x(0), S(1)] \to \min,$$

$$f_2[x(i), S(i)] + f_1[x(i), S(i+1)] \to \min, \qquad (16.45)$$

$$i = 1, 2, ..., m-1,$$

$$f_2[x(m), S(m)] \to \min$$

(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 $S(i) = S$, $i = 1, 2, ..., m$, оптимальный план имеет вид $\{a, b, b, ..., b, ..., b, c\}$, где a — решение первой из указанных в системе (16.45) задач, b — решение второй задачи и c — третьей.

Переход к этим задачам 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.

Рассмотренная модель дает хороший пример пользы математического анализа оптимизационной задачи принятия решений. Такой анализ позволяет решать задачу не стандартными методами, требующими больших вычислительных ресурсов, а с помощью специально разработанных алгоритмов, учитывающих специфику задачи и позволяющих на много порядков 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 [15, глава 5]. Моделирование с целью нахождения оптимальных решений было ранее продемонстрировано на примерах системы моделей, исходящих из классической модели Вильсона, двухуровневой модели, модели оптимизации объемов поставок на базу (склад).

Контрольные вопросы

1. На основе паутинообразной модели ответьте на вопрос: всегда ли цена и объем выпуска приближаются к равновесным?

2. Чем модель экономического роста отличается от модели экономического цикла?

3. Каковы переменные управления в модели влияния государственной финансовой политики на экономику США?

4. Какова роль блоков в создании макромодельных комплексов?

5. Чем экономико-математическая модель малого предприятия типа «поток проектов» отличается от модели типа «занятие ниши»?

6. 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?

7. Каково 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%?

8. Каким образом концепция асимптотически оптимального плана позволяет решить проблему горизонта планирования при принятии логистических решений?

9. В чем состоит основной вклад математики при разработке модели планирования оптимальных размеров поставок и начального запаса?

Темы докладов и рефератов

1. Использование модели межотраслевого баланса В. Леонтьева при планировании.
2. Отражение возможностей управления государственными расходами и налогами в макроэкономических моделях.
3. Место математической теории планирования эксперимента при использовании имитационных экономических моделей.
4. Соотношение экономической теории и эконометрики при построении экономико-математических моделей с целью принятия решений.
5. Экономико-математическое моделирование работы промышленного предприятия.
6. Применение теории массового обслуживания при моделирования работы организаций сферы массовых услуг (телефонных сетей, магазинов и др.).
7. Эконометрическая и экономико-математическая поддержка работы малого предприятия.
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 принятия решений при управления организацией.

ГЛАВА 17
ПРИНЯТИЕ РЕШЕНИЙ НА ОСНОВЕ МОДЕЛЕЙ ОБЕСПЕЧЕНИЯ КАЧЕСТВА

17.1. ОСНОВЫ СТАТИСТИЧЕСКОГО КОНТРОЛЯ КАЧЕСТВА

Одна из наиболее важных областей теории и практики принятия решений, приносящая к тому же наибольший доход, — это обеспечение качества, основанное на применении статистического моделирования.

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» широко известны. 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 [115, 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, а 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[1], 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.

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 явля-

[1] 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.

ется 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 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 [133]). В нашей стране внедрение 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. далее).

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 Ю.К. Беляева и Я.П. Лумельского можно смело назвать классическими. Был выпущен и длинный ряд 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—1987 гг. (перечень стандартов и описание ошибок приведены в работе [81]). 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 [68]).

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 1992г. 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, приемочный (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, т.е. 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. далее) 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 [81].

По оценкам, полученным в работе [87], применение современных статистических методов позволяет в среднем вдвое сократить трудозатраты на контрольные операции (как известно, они составляют 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 млрд дол. 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 дол. на 1 вложенный доллар. Вот почему мы получили такую серьезную поддержку от высшего руководства», — сообщает Билл Виггенхорн, ответственный за подготовку специалистов фирмы «Моторола».

По подсчетам профессора Массачусетского технологического института Фримена [14], только статистический приемочный контроль давал промышленности США 4 млрд дол. в 1958 г. (это более 22 млрд дол. в ценах 2003 г.), т.е. 0,8% валового внутреннего продукта (ВВП).

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 числа 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.

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. превышена предельно допустимая концентрация (ПДК) 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

$$X = X_1 + X_2 + ... + X_n. \qquad (17.1)$$

Из формулы (17.1) и Центральной предельной теоремы теории вероятностей вытекает, что при увеличении объема выборки n распределение X сближается с нормальным распределением. Известно, что распределение X имеет вид

$$P(X = k) = C_n^k p^k (1-p)^{n-k}, \qquad (17.2)$$

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 (17.2) 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.

$$P(Y=k)=\frac{C_D^k C_{N-D}^{n-k}}{C_N^N}. \qquad (17.3)$$

Известно, что биномиальная и гипергеометрическая модели весьма близки, когда объем генеральной совокупности (партии), по крайней мере, в 10 раз превышает объем выборки, т.е.

$$P(X=k)=P(Y=k), \qquad (17.4)$$

если объем выборки мал по сравнению с объемом партии. При этом в качестве p в формуле (17.4) 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 (17.4) показывает, 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 дефектное только одно? 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

$$f(p) = P(X = 0) = (1 - p)^n. \tag{17.5}$$

Для плана $(n, 1)$ оперативная характеристика, как легко видеть, такова:

$$f(p) = P(X = 0) + P(X = 1) = (1 - p)^n + np(1 - p)^{n-1}. \tag{17.6}$$

Оперативные характеристики для конкретных планов статистического контроля не всегда имеют такой простой вид, как в случае формул (17.5) и (17.6). 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 (17.5) 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 $p_{\text{пр}}$ и браковочный $p_{\text{бр}}$, определяемые из уравнений:

$$f(p_{\text{пр}}) = 1 - \alpha, \;\; f(p_{\text{бр}}) = \beta. \tag{17.7}$$

Таким образом, если входной уровень дефектности не превосходит $p_{\text{пр}}$, 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 B 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 $\alpha = 0{,}05$ (т.е. 5%) и $\beta = 0{,}1$ (т.е. 10%).

В качестве примера вычислим приемочный и браковочный уровни дефектности для плана $(n, 0)$. Из формул (17.5) и (17.7) вытекает, что

$$(1 - p_{пр})^n = 1 - \alpha, \quad p_{пр} = 1 - (1 - \alpha)^{1/n}.$$

Поскольку риск поставщика α мал, то из известного соотношения математического анализа

$$\sqrt[m]{1-\alpha} = 1 - \frac{\alpha}{n} + 0\left(\frac{\alpha^2}{n^2}\right)$$

вытекает приближенная формула

$$p_{пр} \approx \frac{\alpha}{n}.$$

Для браковочного уровня дефектности имеем

$$p_{бр} = 1 - \beta^{1/n}.$$

При практическом применении методов статистического приемочного контроля аналитическими формулами, являющимися полезными 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. 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):

$$\text{ПСВУД} = \max_{0 \leq p \leq 1} f_1(p).$$

Пример 17.1. 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:

$$\frac{df_1(p)}{dp} = \left[p(1-p)^n\right]' = (1-p)^n + pn(1-p)^{n-1} =$$
$$= (1-p)^{n-1}(1-p-pn) = (1-p)^{n-1}\left[1-(n+1)p\right] = 0.$$

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,

$$\text{ПСВУД} = p_n(1-p_n)^n = \frac{1}{n+1}\left(1-\frac{1}{n+1}\right)^n. \qquad (17.8)$$

По выражению (17.8) могут быть проведены конкретные расчеты. Однако оно довольно громоздко. Его можно упростить, используя один замечательный предел из курса математического анализа, а именно:

$$\lim_{n\to\infty}\left(1-\frac{1}{n+1}\right)^{n+1} = \mathrm{e}^{-1} = \frac{1}{2{,}718281828\ldots} \approx 0{,}368. \qquad (17.9)$$

Сравнивая соотношения (17.8) и (17.9), 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 (17.9), приближается к 0,368 при росте объема выборки. Поэтому получаем простую асимптотическую формулу

$$\text{ПСВУД} = \frac{0{,}368}{n}.$$

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).

17.2. 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

$$\Phi\left(\frac{c-np_{\text{пр}}}{\sqrt{np_{\text{пр}}(1-p_{\text{пр}})}}\right) = 1-\alpha,\ \Phi\left(\frac{c-np_{\text{бр}}}{\sqrt{np_{\text{бр}}(1-p_{\text{бр}})}}\right) = \beta, \qquad (17.10)$$

откуда с помощью элементарных преобразований получаем, что

$$p_{\text{пр}} = \frac{c}{n} - \frac{1}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}\Phi^{-1}(1-\alpha),$$
$$p_{\text{бр}} = \frac{c}{n} - \frac{1}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}\Phi^{-1}(\beta). \qquad (17.11)$$

Так как величины c/n и p близки друг к другу, то при переходе от формулы (17.10) к формуле (17.11) 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 $\alpha = 0{,}05$, $\beta = 0{,}10$, то в предыдущие формулы следует подставить $\Phi^{-1}(0{,}95) = 1{,}64$ и $\Phi^{-1}(0{,}10) = -1{,}28$. Итак, итоговые формулы для приемочного и браковочного уровней дефектности имеют вид

$$p_{\text{пр}} = \frac{c}{n} - \frac{1{,}64}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)},\ \ p_{\text{бр}} = \frac{c}{n} + \frac{1{,}28}{\sqrt{n}}\sqrt{\frac{c}{n}\left(1-\frac{c}{n}\right)}.$$

Перейдем к задаче синтеза. Пусть заданы приемочный и браковочный уровни дефектности. Требуется построить одноступенчатый план, имеющий эти характеристики. Из формул (17.10) следует, в частности, что

$$\begin{gathered} c - np_{\text{пр}} = \Phi^{-1}(1-\alpha)\sqrt{np_{\text{пр}}(1-p_{\text{пр}})}, \\ c - np_{\text{бр}} = \Phi^{-1}(\beta)\sqrt{np_{\text{бр}}(1-p_{\text{бр}})}. \end{gathered} \qquad (17.12)$$

Вычитая из первого уравнения второе, получаем, что

$$np_{\text{бр}} - np_{\text{пр}} = \sqrt{n}\left[\Phi^{-1}(1-\alpha)\sqrt{p_{\text{пр}}(1-p_{\text{пр}})} - \Phi^{-1}(\beta)\sqrt{p_{\text{бр}}(1-p_{\text{бр}})}\right].$$

Следовательно, оценка n^* необходимого объема выборки имеет вид:

$$n^* = \left[\frac{\Phi^{-1}(1-\alpha)\sqrt{p_{\text{пр}}(1-p_{\text{пр}})} - \Phi^{-1}(\beta)\sqrt{p_{\text{бр}}(1-p_{\text{бр}})}}{p_{\text{бр}} - p_{\text{пр}}}\right]^2.$$

Для стандартных значений рисков $\alpha = 0{,}05$, $\beta = 0{,}10$ имеем

$$n^* = \left[\frac{1{,}64\sqrt{p_{\text{пр}}(1-p_{\text{пр}})} + 1{,}28\sqrt{p_{\text{бр}}(1-p_{\text{бр}})}}{p_{\text{бр}} - p_{\text{пр}}}\right]^2. \qquad (17.13)$$

С помощью уравнений (17.12) нетрудно найти оценку c^* приемочного числа, заменив неизвестный объем выборки на его оценку n^*. Будем использовать оценку

$$c^* = n^* p_{\text{бр}} + \Phi^{-1}(\beta)\sqrt{n^* p_{\text{бр}}(1-p_{\text{бр}})}.$$

Для стандартного значения $\beta = 0{,}10$ имеем

$$c^* = n^* p_{\text{бр}} - 1{,}28\sqrt{n^* p_{\text{бр}}(1-p_{\text{бр}})}. \qquad (17.14)$$

Итак, по формуле (17.13) можно рассчитать оценку объема выборки, затем по формуле (17.14) 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 задачу — по заданным приемочному и браковочному уровням дефектности подобрать такой одноступенчатый план контроля, чтобы его оперативная характеристика $f(p)$ удовлетворяла неравенствам

$$f(p_{\text{пр}}) \geq 1-\alpha,\ \ f(p_{\text{бр}}) \leq \beta.$$

Поэтому при практической работе корректировка асимптотических результатов должна быть направлена на выполнение указанных неравенств.

Пример 17.2. Пусть $p_{\text{пр}} = 0{,}02$, $p_{\text{бр}} = 0{,}09$. Тогда по формуле (17.13) оценка объема выборки

$$n^* = \left[\frac{1{,}64\sqrt{0{,}02(1-0{,}02)} + 1{,}28\sqrt{0{,}09(1-0{,}09)}}{0{,}09-0{,}02}\right]^2 =$$

$$= \left[\frac{1{,}64\times 0{,}14 + 1{,}28\times 0{,}286}{0{,}07}\right]^2 = \left[\frac{0{,}2296+0{,}3661}{0{,}07}\right]^2 =$$

$$= \left[\frac{0{,}5957}{0{,}07}\right]^2 = 8{,}51^2 = 72{,}42.$$

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 (17.14):

$$c^* = 72\times 0{,}09 - 1{,}28\sqrt{72\times 0{,}09\times 0{,}91} = 6{,}48 - 1{,}28\times 2{,}428 = 3{,}37.$$

Полученное число не является целым, поэтому в качестве приемочного числа надо взять ближайшее целое, т.е. 3.

Если объем выборки округлить до 73, то аналогично получим

$$c^{**} = 73\times 0{,}09 - 1{,}28\sqrt{73\times 0{,}09\times 0{,}91} = 6{,}57 - 1{,}28\times 2{,}445 = 3{,}44.$$

При округлении снова получаем 3.

С помощью первого из уравнений (17.12) можно построить оценку c^* на основе приемочного уровня дефектности:

$$\begin{aligned} c^* &= n^* p_{\text{пр}} + \Phi^{-1}(1-\alpha)\sqrt{n^* p_{\text{пр}}(1-p_{\text{пр}})} = \\ &= n^* p_{\text{пр}} + 1{,}64\sqrt{n^* p_{\text{пр}}(1-p_{\text{пр}})}. \end{aligned}$$

Подставив конкретные значения, получим практически ту же оценку, что и раньше:

$$c^* = n^* p_{\text{пр}} + 1{,}64\sqrt{n^* p_{\text{пр}}(1 - p_{\text{пр}})} =$$
$$= 72 \times 0{,}02 + 1{,}64\sqrt{72 \times 0{,}02 \times 0{,}98} = 3{,}39.$$

Итак, в результате асимптотических расчетов найден одноступенчатый план.

17.3. 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 («Правила приемки», «Методы контроля» и др.) имеют различные недостатки и неточности, что может послужить в дальнейшем причиной возникновения арбитражных ситуаций (т.е. ситуаций, 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).

Здесь необходимо особо отметить следующее. До сих пор постоянно говорилось о контроле единиц и партий продукции. Однако нет никакого принципиального отличия данного контроля от контроля 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 используют 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. 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 [14; 87].

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 не важно, каковы будут результаты по остальным 59 единицам продукции, окажутся они годными или дефектными — четыре 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 трех 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 применяют, 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 [34] 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 число пар точек, отстоящих друг от друга на определенное расстояние, и выяснить, есть связь (т.е. значим ли ранговый коэффициент корреляции Спирмена — см. главу 15) 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 ранее подходы для контроля ее качества (подробнее см. [34]).

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. 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 [81]. 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 не намного 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

$$n = 2{,}30/p_{бр}. \tag{17.15}$$

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, 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 (17.15), 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%) 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 [51; 118]. В частности, упомянем методы «всеобщего» (или 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 число 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 $\Phi(t)$, где $\Phi(t)$ — функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. Поскольку $\Phi(4) = 0{,}999968329$, 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:

- сплошной контроль (затраты C_1N);
- пополнение партии дополнительными изделиями в числе $D_0(4)$ — затраты $C_0D_0(4)$.

Вторая стратегия лучше (экономически выгоднее), если

$$C_1N > C_0D_0(4) = C_0\left[Np + 4\sqrt{Np(1-p)}\right]. \qquad (17.16)$$

Поделим обе части на C_0N, получим равносильное неравенство

$$\frac{C_1}{C_0} > p + 4\frac{\sqrt{p(1-p}}{\sqrt{N}}.$$

Поскольку $p(1-p)$ не превосходит $^1/_4$ при всех p, то из неравенства

$$C_1/C_0 > p + 2/N^{1/2} \qquad (17.17)$$

вытекает неравенство (17.16). Ясно, что в случае если

$$C_1/C_0 > p,$$

неравенство (17.17), а потому и неравенство (17.16) выполняются при достаточно больших объемах партии, а именно при

$$N > [2C_0/(C_1 - C_0 p)]^2.$$

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 (17.16) выполняется при

$$0{,}1 > 0{,}01 + 4\ (0{,}01 \times 0{,}99)^{1/2}/N^{1/2},$$

т.е. при $N > 19$. Расхождение более чем на порядок (в 26 раз) объясняется заменой при переходе от формулы (17.16) к формуле (17.17) 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 ранее 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, произведя затраты $C_0 D_0(4)$, общие затраты (в среднем на одну выпущенную партию) будут равны

$$C_{\text{общ}} = 0{,}95[C_1 n + C_0 D_0(4)].$$

Во втором случае фактически проводится сплошной контроль с исправлением дефектов и затратами $C_1 N$. Суммарные затраты при использовании выборочного контроля равны

$$C_{\text{сум}} = 0{,}95[C_1 n + C_0 D_0(4)] + 0{,}05 C_1 N.$$

Выборочный контроль 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 (17.16) показывает, что если контроль не является разрушающим, то выборочный контроль менее выгоден, чем сплошной. По сравнению с формулой (17.16) 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 невыгодно.

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

$$WNp < C_1N,\ p < C_1/W, \tag{17.18}$$

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 (17.18) участвует математическое ожидание WNp. Реальные потери могут быть больше, но не намного. Как и ранее, 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

$$WD_0(4) < C_1N,\ p + 4[p(1-p)]^{1/2}/N^{1/2} < C_1/W. \tag{17.19}$$

Аналогично выводу неравенства (17.17) заключаем, что неравенство (17.19) наверняка будет выполнено, если

$$p + 2/N^{1/2} < C_1/W. \tag{17.20}$$

Пусть $C_1/W = 0{,}1$, выпускается партия объема $N = 1600$. Тогда согласно неравенству (17.20) 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

$$y\{C_1n + W(N-n)p\} + (1-y)C_1N < WNp.$$

Преобразуем это неравенство к виду

$$yn\{C_1 - Wp\}(1-y)^{-1} + C_1N < WNp. \tag{17.21}$$

Если выполнено неравенство $p < C_1/W$, то второе слагаемое в левой части неравенства (17.21) больше правой части этого неравенства, в то время как первое слагаемое в левой части (17.21) положительно. Следовательно, неравенство (17.21) неверно и введение выборочного контроля нецелесообразно — как и в разобранном ранее случае метода пополнения партий.

Ранее был 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 было показано, 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 — *Total quality management* (TQM)», хотя бы потому, что они незаменимы при использовании разрушающих методов контроля.

Наиболее перспективным представляется использование полученных результатов в рамках концепции контроллинга (см., например, [24]). Итак, 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.

17.4. СТАТИСТИЧЕСКИЕ МЕТОДЫ УПРАВЛЕНИЯ КАЧЕСТВОМ

Как уже отмечалось, в России все расширяется 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.

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, можно 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. Следует отметить, что в литературе имеются различные мнения по поводу понятия «конкурентоспособность». В частности, нельзя согласиться с крайне упрощенным подходом, 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. Справедливо отмечается, что фирма, обладающая патентом или новой научно-технической разработкой, имеет более высокий «излишек производителя» по сравнению с другими фирмами. Технический уровень продукции, например, — 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 [132]). 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 [74; 81].

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 Институте высоких статистических технологий и эконометрики сложилось пятичленное деление по первому основанию (в скобках указаны наименования диалоговых систем, рассмотренных, в частности, в работе [68]):

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 13 лет (с 1991 по 2004 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 «Общее 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. 17.1, показывает, что статистические методы широко применяются на всех этапах жизненного пути продукции.

Таблица 17.1

Применение статистических методов на различных этапах жизненного цикла продукции по ИСО 9004—87

Номер этапа	Вид статистических методов					Специальные модели
	а	б	в	г	д	
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 16 данного учебника или главу 5 монографии [88]).

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 ори-

ентированный на экспертное оценивание ГРАНТ, на анализ интервальных данных — РЕСТ);

- прогнозирование (ЛИСАТИС и его развитие ЭКОСТАТ, а также относящиеся к временны́м рядам разделы пакета АВРОРА-РС — анализ временны́х 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 [88, глава 5]), массового обслуживания и др. Такие модели и их программное обеспечение, как правило, разрабатывают для конкретного предприятия, и потому они хорошо приспособлены к особенностям этого предприятия.

«Шесть сигм» — система внедрения контроллинга и его статистических инструментов. Подробно система «Шесть сигм» рассмотрена в подразделе 14.2. Система трудоемка, на ее внедрение нужны годы. Но и эффект велик [127]. 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.

17.5. ОБНАРУЖЕНИЕ РАЗЛАДКИ С ПОМОЩЬЮ КОНТРОЛЬНЫХ КАРТ[1]

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 временны́е 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 временны́х рядов, на предмет наискорейшего обнаружения момента спонтанного изменения (разладки) вероятностных характеристик контролируемых процессов. Во многих практических задачах 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 [51; 118].

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

[1] Подраздел написан совместно с И.Н. Митрохиным, ОАО «Северсталь-авто», Москва, Россия.

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 выборку значений контролируемого показателя качества моделируют совокупностью независимых одинаково распределенных случайных величин с математическим ожиданием μ и дисперсией σ^2 [89]. При статистическом регулировании технологических процессов проверяют гипотезы:

$H_0 : \mu = \mu_0$ (технологический процесс налажен),

$H_1 : \mu = \mu_1$ (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:

$H_0 : \sigma = \sigma_0$ (технологический процесс налажен),

$H_1 : \sigma = \sigma_1$ (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 число единиц продукции при нормальном ходе производства, т.е. при надлежащем качестве сырья и при отлаженном оборудовании. При этих условиях мы по правилам прикладной статистики [77] 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* [118]. Методология построения контрольных карт Шухарта имеет ту особенность, что проводимый контроль не учитывает предыдущего поведения контролируемого параметра. В середине XX в. 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 кумсумм (или кумсумм-метода).

Кумсумма образуется следующим образом:

$$S_r = \sum_{i=1}^{r}(x_i - k) = S_{r-1} + (x_r - k), \qquad (17.22)$$

где x_i — значение используемой статистики в i-й момент времени; k — константа, представляющая собой некоторое эталонное значении.

Нанесенные на график в порядке появления кумулятивные суммы S_r образуют кумсумм-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 17.3. 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 (см. табл). Причем ПФ за первые 3 года считается стабильным. Проверим, не разлажен ли процесс выплаты ПФ в 2005 г., при условии, что норматив ПФ в течение всего контролируемого периода не изменяется. На рисунке 17.1 показана динамика ПФ.

Премиальный фонд по месяцам (млн руб.)

Год	Месяц											
	Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь
2002	128	134	143	124	147	127	131	150	120	135	129	150
2003	132	143	141	152	140	125	137	142	153	130	136	137
2004	122	136	153	135	125	139	143	132	146	126	139	136
2005	150	143	155	160	144	153	139	168	151	—	—	—

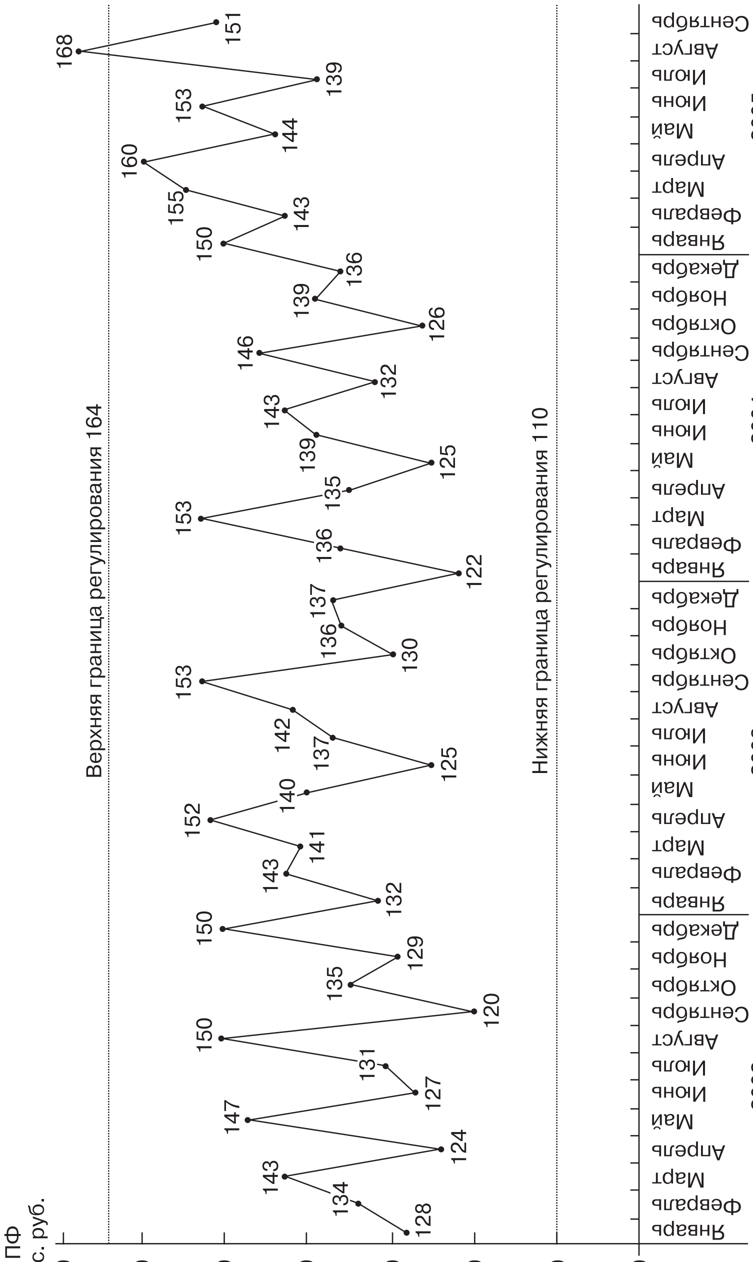


Рис. 17.1. Контрольная карта Шухарта для ПФ

На основании табличных данных слабо прослеживается изменение среднего уровня ПФ в 2005 г. Построим карту Шухарта. Делаем предположение, что с 2002 по 2004 г. норматив ПФ за месяц не менялся. Для расчета выборочного среднего арифметического значения и выборочной дисперсии воспользуемся данными 2002—2004 гг., поскольку в эти периоды ПФ считается стабильным, управляемым.

Среднее значение выборки рассчитываем по формуле

$$\overline{\text{ПФ}} = \frac{\text{ПФ}_1 + \text{ПФ}_2 + ... + \text{ПФ}_{36}}{36} =$$

$$= \frac{128 + 134 + ... + 136}{36} = 137 \text{ млн руб.}$$

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 среднеквадратичное 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 среднеквадратичных отклонения. Тогда допустимые значения премиального фонда будут лежать в диапазоне $\left[\overline{\text{МД}} - 3\sigma; \overline{\text{МД}} + 3\sigma\right]$, 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 ПФ (см. рис. 17.1). Анализ карты показывает, что в августе 2005 г. произошла разладка процесса, ПФ 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: $k = \overline{\text{ПФ}} = 137$ млн руб.

Далее рассчитываем кумулятивные суммы по формуле (17.22):

Значения кумулятивных сумм

Год	Месяц											
	Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь	Ноябрь	Декабрь
2002	–9	–11	–5	–17	–7	–17	–22	–9	–26	–27	–35	–21
2003	–26	–20	–15	0	4	–8	–8	–2	14	8	7	7
2004	–7	–8	8	7	–5	–2	4	–1	9	–2	1	0
2005	13	20	38	62	69	85	88	119	134			

Карта кумсумм представлена на рис. 17.2.

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 месяц 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 2002 г. = 134 млн руб.;

ПФ с декабря 2002 г. по август 2003 г. = 139 млн руб.;

ПФ с сентября 2003 г. по декабрь 2004 г. = 137 млн руб.;

ПФ с января 2005 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,

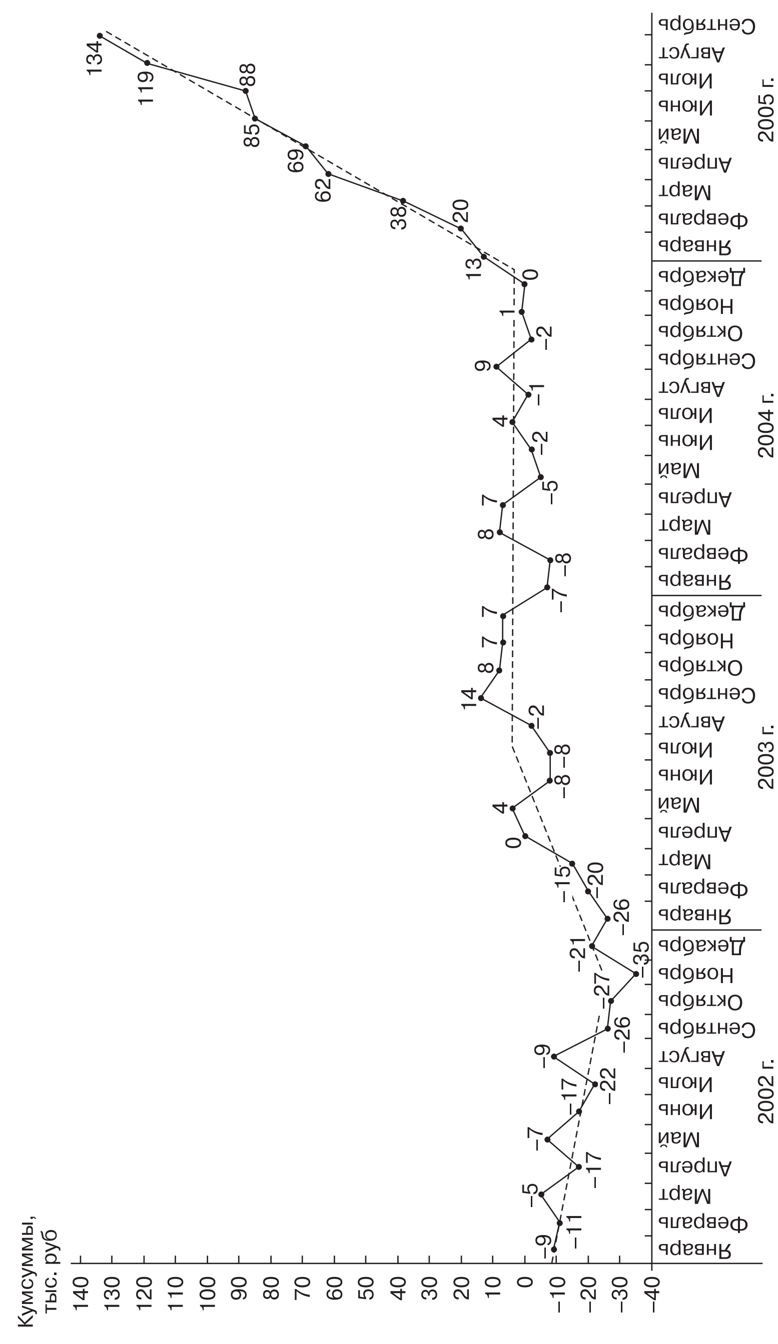


Рис. 17.2. Карта кумсумм

среднее квадратичное 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. Определяется СДС 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.

Эффективность плана контроля СДС определяет 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 обычно связанный с L_0 переводят в ограничение. Величина L_0 определяет положение границ регулирования.

Метод контрольных карт в России. Судьба метода контрольных карт в России драматична. В 1970-е г. 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 [81]. Эти стандарты были отменены. К сожалению, ряд литературных источников, как и отмененные ГОСТы, содержат 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 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 [77; 89]. 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 И.В. Никифорова (академический Институт проблем управления) и 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 [68].

Подведем итоги настоящей главы. В России активно разрабатываются теоретические, программные и практические вопросы статистических методов сертификации и управления качеством продукции. 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.

Контрольные вопросы

1. Какие решения необходимо принимать в связи с качеством продукции и сертификацией?
2. Почему необходимо использование выборочного контроля?
3. Как 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. Как могли быть утверждены государственные стандарты по статистическим методам управления качеством продукции, содержащие грубейшие ошибки?
6. Чем контрольные карты Шухарта отличаются от карт кумулятивных сумм?

Темы докладов и рефератов

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.

Заключение

Организационно-экономическое моделирование и его важнейшая часть — теория принятия решений — в настоящее время быстро развиваются. Происходит смена парадигмы (совокупности ценностей, методов, технических навыков и средств, принятых в научном сообществе в рамках устоявшейся научной традиции в определенный период времени) — переход от подходов 1950-х гг. к современным. Настоящий учебник написан в рамках новой парадигмы, и мы надеемся, что он будет полезным, по крайней мере, до середины XXI в.

Составление учебника — всегда отбор научных результатов. Кратко обсудим некоторые перспективные подходы теории и практики разработки и принятия управленческих решений, которые остались за пределами настоящего учебника.

Прежде всего обратим внимание на современные статистические методы. Их точки роста, методология, основные нерешенные проблемы обсуждаются в заключительных главах учебника «Прикладная статистика» [77]. Обратим внимание на концепцию высоких статистических технологий, а также на компьютерно-статистические методы — синтез информационных технологий и асимптотической математической статистики.

Центральной областью современной прикладной статистики является нечисловая статистика [76], а нечеткие множества — частный случай нечисловых данных. Они весьма перспективны с прикладной точки зрения. И с теоретической тоже — хотя уже более 30 лет назад установлено, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств [70; 88], дальнейшая проработка этой связи между двумя теориями сулит не только теоретические продвижения, но и создание новых эффективных методов разработки и принятия управленческих решений.

Неопределенности реальных явлений и процессов можно описывать с помощью вероятностно-статистических, нечетких и интервальных моделей. Статистика интервальных данных достаточно подробно изложена в [76; 77], поэтому в настоящем учебнике о ней нет ни слова. Подчеркнем высокую прикладную значимость этого раздела нечисловой статистики.

Процессы разработки и принятия решений реализуются в реальных ситуациях с достаточно высоким уровнем неопределенности. Велика роль нечисловой информации как на «входе», так и на «выходе» процесса принятия управленческого решения. Неопределенность и нечисловая природа управленческой информации должны быть отражены при анализе устойчивости экономико-математических методов и моделей. Для обоснованного практического применения математических моделей процессов управления промышленными предприятиями и основанных на них экономико-математических методов должна быть изучена их устойчивость по отношению к допустимым

отклонениям исходных данных и предпосылок моделей. Общая схема изучения проблем устойчивости разработана в [88], конкретные постановки проблем устойчивости в математических методах и моделях теории принятия решений разобраны, например, в главах 9, 10 настоящего учебника. Если же обратить внимание на необходимость изучения устойчивости по отношению к изменению данных (статистика интервальных и нечетких данных), объема данных (асимптотическая статистика) и особенно к изменению распределения данных (непараметрическая и робастная статистика), то приходится констатировать, что эта тематика пронизывает все содержание настоящего учебника.

От математических моделей перейдем к предметной области — менеджменту высоких технологий. Выделим четыре актуальных направления работ.

В настоящее время только создается единая теория риска. Даже единая классификация рисков пока не утвердилась [89, глава 14]. Однако можно с уверенностью прогнозировать, что в ближайшие годы на предприятиях и в организациях будут созданы службы управления риском во главе с одним из топ-менеджеров. Директор по рискам займет место рядом с финансовым директором, главным инженером и директором по маркетингу и сбыту.

Экологический менеджмент уже интегрирован в системы управления предприятиями (в соответствии со стандартами ИСО серии 14000). Но этого мало. Предсказываем, что социально-экологические проблемы управления будут все более актуальными в современных условиях [95].

Современные технологии управления сконцентрированы вокруг концепции контроллинга. Одна из основных функций контроллинга — информационно-аналитическая поддержка процессов принятия решений на предприятиях и в организациях. Службы контроллинга играют все более важную роль в совершенствовании процессов управления на базе интенсивного использования информационных технологий.

Неформальная информационная экономика будущего (НИЭБ) развивается нами как методологическая основа конкретных исследований в области организационно-экономического моделирования. Одна из ее целей — выявить основные черты экономики будущего на период (на 20—30 лет) стратегического планирования государства и крупных корпораций. Перспективные организационно-экономические механизмы управления производственно-хозяйственной деятельностью, разработки и принятия управленческих решений предлагаем конструировать на основе НИЭБ. Эта новая организационно-экономическая теория находится на стыке теории управления, экономики и прогностики.

Итак, теория принятия решений на основе организационно-экономического моделирования бурно развивается. Невозможно даже упомянуть все ее перспективные направления. В учебнике мы указали лишь на некоторые работы нашего коллектива — научно-учебного комплекса «Инженерный бизнес и менеджмент» Московского государственного технического университета им. Н.Э. Баумана.

Список литературы

1. *Абрамов Ю.А.* В поисках баланса интересов и ресурсов // Космос в фокусе политики, экономики, культуры / науч. ред. Л.В. Голованов. М. : Новости космонавтики, 2002.
2. *Адамар Ж.* Исследование психологии процесса изобретения в области математики. М. : УРСС, 2001.
3. *Алешин Д.Н.* Экономическое обоснование эффективности инвестиционных проектов на предприятиях на основе применения эконометрического метода интервальной оценки : автореф. дис. ... канд. экон. наук. М. : Изд-во МГТУ им. Н.Э. Баумана, 2001.
4. Анализ нечисловой информации / Ю.Н. Тюрин, Б.Г. Литвак, А.И. Орлов [и др.]. М. : Научный совет АН СССР по комплексной проблеме «Кибернетика», 1981.
5. Анализ нечисловой информации в социологических исследованиях / под ред. В.Г. Андреенкова, А.И. Орлова, Ю.Н. Толстовой. М. : Наука, 1985.
6. *Барский Б.В., Соколов М.В.* Средние величины, инвариантные относительно допустимых преобразований шкалы измерения // Заводская лаборатория. 2006. Т. 72. № 1. С. 59—66.
7. *Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. 3-е. изд. М. : Наука, 1983.
8. *Боровков А.А.* Математическая статистика : учеб. пособие для вузов. М. : Наука, 1984.
9. *Бурков В.Н., Заложнев А.Ю., Новиков Д.А.* Теория графов в управлении организационными системами. М. : Синтег, 2001.
10. *Ван дер Варден Б.Л.* Математическая статистика. М. : Иностранная литература, 1960.
11. *Воинов В.Г., Никулин М.С.* Несмещенные оценки и их применения. М. : Наука, 1989.
12. *Вольский В.И., Лезина З.М.* Голосование в малых группах. Процедуры и методы сравнительного анализа. М. : Наука, 1991.
13. *Гельфанд И.М., Розенфельд Б.И., Шифрин М.А.* Очерки о совместной работе математиков и врачей. 2-е изд., доп. М. : УРСС, 2004.
14. *Гнеденко Б.В.* Математика и контроль качества продукции. М. : Знание, 1978.
15. *Гнеденко Б.В., Коваленко И.Н.* Введение в теорию массового обслуживания. М. : Наука, 1966.
16. *Горский В.Г., Гриценко А.А., Орлов А.И.* Метод согласования кластеризованных ранжировок // Автоматика и телемеханика. 2000. № 3. С. 159—167.
17. *Гуськова Е.А.* Разработка организационно-экономических методов повышения эффективности деятельности промышленного предприятия на основе эконометрического подхода : автореф. дис. ... канд. экон. наук. М. : Изд-во МГТУ им. Н.Э. Баумана, 2004.
18. *Добротворский Н., Седов А.* Курс холодильника к кошельку: живем, как в 1985 году! // Комсомольская правда. 2003. 10 дек.

19. Доклад о мировом развитии 2004 г. Всемирный банк. М. : Весь Мир, 2004.
20. *Дылько Т.Н.* Проверка гипотез в экспертном оценивании // Вестник Белорусского государственного университета. Сер. 1 : Физика, математика и механика. 1988. № 2. С. 36—40.
21. *Дэвид Г.* Метод парных сравнений. М. : Статистика. 1978.
22. *Иванова Н.Ю., Орлов А.И.* Экономико-математическое моделирование малого бизнеса (обзор подходов) // Экономика и математические методы. 2001. Т. 37. № 2. С. 128—136.
23. Инженерная экономика : учебник / В.В. Кочетов, А.А. Колобов, И.Н. Омельченко ; под ред. А.А. Колобова, А.И. Орлова. М. : Изд-во МГТУ им. Н.Э.Баумана, 2005.
24. *Карминский А.М., Пересецкий А.А., Петров А.Е.* Рейтинги в экономике: методология и практика. М. : Финансы и статистика, 2005.
25. *Кемени Дж., Снелл Дж.* Кибернетическое моделирование : Некоторые приложения. М. : Советское радио, 1972.
26. *Кендалл М.Дж., Стьюарт А.* Многомерный статистический анализ и временны́е ряды. М. : Наука, 1976.
27. *Кендалл М.Дж., Стьюарт А.* Статистические выводы и связи. М. : Наука, 1973.
28. *Кендэл М.* Ранговые корреляции. М. : Статистика, 1975.
29. *Клейн Ф.* Лекции о развитии математики в XIX столетии. Ч. I. М. ; Л. : Объединенное научно-техническое изд-во НКТП СССР, 1937.
30. *Колмогоров А.Н.* Об определении среднего // Избранные труды : Математика и механика. М. : Наука, 1985.
31. *Колмогоров А.Н., Фомин С.В.* Элементы теории функций и функционального анализа. М. : Наука, 1972.
32. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях / А.М. Карминский, Н.И. Оленев, А.Г. Примак, С.Г. Фалько. М. : Финансы и статистика, 1998.
33. *Коростикова Т.* Цены вырастут в 5 раз // Аргументы и факты. 1994. № 16. С. 5.
34. *Кравченко Г.Г., Орлов А.И.* О статистическом приемочном контроле порошкообразных материалов // Надежность и контроль качества. 1991. № 2. С. 37—39.
35. *Крамер Г.* Математические методы статистики. М. : Мир, 1975.
36. *Крюкова Е.М.* Применение методов организационно-экономического прогнозирования в отрасли лома черных металлов // Заводская лаборатория. 2008. Т. 74. № 7. С. 67—72.
37. *Кудлаев Э.М., Орлов А.И.* Вероятностно-статистические методы исследования в работах А.Н. Колмогорова // Заводская лаборатория. 2003. Т. 69. № 5. С. 55—61.
38. *Леман Э.* Проверка статистических гипотез. М. : Наука, 1979.
39. *Леонтьев В.В.* Экономические эссе. Теория, исследования, факты и политика. М. : Политиздат, 1991.

40. *Литвак Б.Г.* Экспертиза в России // Заводская лаборатория. 2000. Т. 66. № 7. С. 61—66.
41. *Литвак Б.Г.* Экспертная информация. Методы получения и анализа. М. : Радио и связь, 1982.
42. *Литвак Б.Г.* Экспертные оценки и принятие решений. М. : Патент, 1996.
43. *Лумельский Я.П.* Статистические оценки результатов контроля качества. М. : Изд-во стандартов, 1979.
44. *Львов Д.С.* Реформы с позиции современной науки // Научные труды Международного союза экономистов и Вольного экономического общества России. Т.2. М. ; СПб. : [б.и.], 1995.
45. *Льюс Р., Галантер Е.* Психофизические шкалы // Психологические измерения. М. : Мир, 1967.
46. *Майстров Л.Е.* Теория вероятностей : исторический очерк. М. : Наука, 1967.
47. *Макконелл Кэмпбелл Р., Брю Стэнли Л.* Экономикс. Принципы, проблемы и политика : в 2 т. Т. 1. : пер. с англ. М. : Республика, 1995.
48. *Маркова Е.В., Никитина Е.П.* Математическая теория эксперимента: история, развитие, будущее // Заводская лаборатория. 2002. Т. 68. № 1. С. 112—118.
49. Математическое моделирование процессов налогообложения (подходы к проблеме) / под ред. В.Н. Жихарева, А.И. Орлова [и др.]. М. : ЦЭО Минобразования РФ, 1997.
50. Менеджмент : учеб. пособие / под ред. Ж.В. Прокофьевой. М. : Знание, 2000.
51. *Мердок Дж.* Контрольные карты : пер. с англ. М. : Финансы и статистика, 1986.
52. Методические рекомендации по оценке эффективности инвестиционных проектов и их отбору для финансирования. М. : Минэкономики РФ, 1994.
53. *Миркин Б.Г.* Проблема группового выбора. М. : Наука, 1974.
54. *Моисеев Н.Н.* Люди и кибернетика. М. : Молодая гвардия, 1984.
55. *Моисеев Н.Н.* Математические задачи системного анализа. М. : Наука, 1981.
56. *Муравьева В.С.* Точка встречи: асимптотическое распределение уровня качества и временного лага // Заводская лаборатория. 2008. Т. 74. № 41. С. 70—74.
57. *Муравьева В.С., Орлов А.И.* Непараметрическое оценивание точки пересечения регрессионных прямых // Заводская лаборатория. Диагностика материалов. 2008. Т. 74. № 1. С. 63—68.
58. *Муравьева В.С., Орлов А.И.* Организационно-экономические проблемы прогнозирования на промышленном предприятии // Управление большими системами. Вып. 17. М. : ИПУ РАН, 2007.
59. *Науман Э.* Принять решение — но как? : пер. с нем. М. : Мир, 1987.
60. *Нейлор Т.* Машинные имитационные эксперименты с моделями экономических систем. М. : Мир, 1975.

61. *Нейман Дж.фон, Моргенштейн О.* Теория игр и экономическое поведение. М. : Наука, 1970.
62. *Неуймин Я.Г.* Модели в науке и технике. История, теория, практика. Л. : Наука, 1984.
63. *Носовский Г.В. , Фоменко А.Т.* Царь славян. СПб. : Нева, 2004.
64. *Окстоби Дж.* Мера и категория. М. : Мир, 1974.
65. Организация и планирование машиностроительного производства (производственный менеджмент) : учебник / К.А. Грачева, М.К. Захарова, Л.А. Одинцова [и др.] ; под ред. Ю.В. Скворцова, Л.А. Некрасова. М. : Высшая школа, 2003.
66. *Орлов А.И.* «Шесть сигм» — новая система внедрения математических методов исследования // Заводская лаборатория. 2006. Т. 72. № 5. С. 50—53.
67. *Орлов А.И.* Вероятность и прикладная статистика — основные факты : справ. М. : КноРус, 2010.
68. *Орлов А.И.* Внедрение современных статистических методов с помощью персональных компьютеров // Качество и надежность изделий. № 5 (21). М. : Знание, 1992.
69. *Орлов А.И.* Грядущая смута 2012 года // Вестник Академии прогнозирования (исследований будущего). 2004. № 12. С. 42—45.
70. *Орлов А.И.* Задачи оптимизации и нечеткие переменные. М. : Знание, 1980.
71. *Орлов А.И.* Математические модели отдельных сторон обучения математике // Научно-методические статьи по математике. (Проблемы преподавания математики в вузах). Вып. 7. М. : Высшая школа, 1978.
72. *Орлов А.И.* Некоторые вероятностные вопросы теории классификации // Прикладная статистика : ученые записки по статистике. Т. 45. М. : Наука, 1983.
73. *Орлов А.И.* Непараметрический метод наименьших квадратов: учет сезонности // Статистические методы оценивания и проверки гипотез : межвуз. сб. науч. тр. Вып. 21. Пермь : Изд-во Пермского государственного университета, 2008.
74. *Орлов А.И.* О современных проблемах внедрения прикладной статистики и других статистических методов // Заводская лаборатория. 1992. Т. 58, № 1. С. 67—74.
75. *Орлов А.И.* Организационно-экономическое моделирование процессов управления промышленными предприятиями в условиях рисков инфляции. Стратегическое планирование и развитие предприятий / Материалы Девятого всероссийского симпозиума. М. : ЦЭМИ РАН, 2008.
76. *Орлов А.И.* Организационно-экономическое моделирование. Часть 1. Нечисловая статистика. М. : Изд-во МГТУ им. Н.Э. Баумана, 2009.
77. *Орлов А.И.* Прикладная статистика : учебник. М. : Экзамен, 2006.
78. *Орлов А.И.* Размытые цены. Нечисловая экономика и управление инвестиционным процессом // Российское предпринимательство. 2001. № 12. С. 103—108.

79. *Орлов А.И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат // Заводская лаборатория. 1985. Т. 51. № 1. С. 60—62.
80. *Орлов А.И.* Связь между средними величинами и допустимыми преобразованиями шкалы // Математические заметки. 1981. Т. 30. № 4. С. 561—568.
81. *Орлов А.И.* Сертификация и статистические методы // Заводская лаборатория. 1997. Т. 63. № 3. С. 55—62.
82. *Орлов А.И.* Случайные множества с независимыми элементами (люсианы) и их применения // Алгоритмическое и программное обеспечение прикладного статистического анализа : ученые записки по статистике. Т. 36. М. : Наука, 1980.
83. *Орлов А.И.* Статистика объектов нечисловой природы и экспертные оценки // Вопросы кибернетики. Вып. 58. М. : Научный совет АН СССР по комплексной проблеме «Кибернетика», 1979.
84. *Орлов А.И.* Статистические методы прогнозирования // Малая российская энциклопедия прогностики. М. : Изд-во Института экономических стратегий, 2007.
85. *Орлов А.И.* Сценарии социально-экономического развития России в XXI в. // Обозреватель-Observer. 2000. № 10—11. С. 82.
86. *Орлов А.И.* Сценарии социально-экономического развития России до 2007 г. // Обозреватель-Observer. 1999. № 10 (117). С. 47—50.
87. *Орлов А.И.* Теория принятия решений : учебник. М. : Экзамен, 2006.
88. *Орлов А.И.* Устойчивость в социально-экономических моделях. М. : Наука, 1979.
89. *Орлов А.И.* Эконометрика : учебник для вузов. 4-е изд., испр. и доп. Ростов н/Д : Феникс, 2009.
90. *Орлов А.И.* Эконометрическая поддержка контроллинга // Контроллинг. 2002. № 1. С. 42—53.
91. *Орлов А.И.* Экспертные оценки // Заводская лаборатория. 1996. Т. 62. № 1. С. 54—60.
92. *Орлов А.И., Орлова Л.А.* Интервальная оценка инфляции по независимой информации // Российское предпринимательство. 2004. № 10. С. 44—49.
93. *Орлов А.И., Орлова Л.А.* Эконометрика в обучении контроллеров // Контроллинг. 2004. № 3 (11). С. 68—73.
94. *Орлов А.И., Раушенбах Г.В.* Метрика подобия : аксиоматическое введение, асимптотическая нормальность // Статистические методы оценивания и проверки гипотез : межвуз. сб. науч. тр. Пермь : Изд-во Пермского государственного университета, 1986.
95. *Орлов А.И., Федосеев В.Н.* Менеджмент в техносфере : учеб. пособие. М. : ИЦ «Академия», 2003.
96. *Орлов А.И., Гусейнов Г.А.* Математические методы в изучении способных к математике школьников // Исследования по вероятностно-статистическому моделированию реальных систем. М. : ЦЭМИ АН СССР, 1977.
97. *Панде П., Холп Л.* Что такое «Шесть сигм»? Революционный метод управления качеством : пер. с англ. М. : Альпина Бизнес Букс, 2004.
98. *Паркинсон С.Н.* Законы Паркинсона: сборник : пер. с англ. М. : Прогресс, 1989.
99. *Подиновский В.В.* Введение в теорию важности критериев в многокритериальных задачах принятия решений. М. : Физматлит, 2007.
100. *Подиновский В.В., Ногин В.Д.* Парето-оптимальные решения многокритериальных задач. М. : Наука, 1982.
101. *Пойа Д.* Математическое открытие. М. : Наука, 1970.
102. *Пфанцагль И.* Теория измерений. М. : Мир, 1976.
103. *Раушенбах Г.В.* Меры близости и сходства // Анализ нечисловой информации в социологических исследованиях. М. : Наука, 1986.
104. Рекомендации. Прикладная статистика. Методы обработки данных. Основные требования и характеристики / А.И. Орлов, Н.Г. Миронова, В.Н. Фомин, А.Н. Черчинцев. М. : ВНИИ стандартизации Госстандарта СССР, 1987.
105. *Ремарк Э.М.* Черный обелиск. М. : АО «ВИТА-ЦЕНТР», 1992.
106. *Ромашкина Г.Ф., Татарова Г.Г.* Коэффициент конкордации в анализе социологических данных // Социология: методология, методы, математические модели. 2005. № 20. С. 131—158.
107. Российская цивилизация: сквозь тернии к звездам : сборник. М. : Вече, 2003.
108. *Рыданова Г.В.* Некоторые вопросы статистического анализа случайных бинарных векторов : автореф. дис. ... канд. физ.-мат. наук. М. : Изд-во МГУ, 1987.
109. *Саати Т.* Принятие решений. Метод анализа иерархий. М. : Радио и связь, 1993.
110. *Самуэльсон П.* Экономика : М. : МГП «Алгон» — ВНИИСИ, 1992.
111. *Себер Дж.* Линейный регрессионный анализ. М. : Мир, 1980.
112. *Севастьянов Б.А.* Ветвящиеся процессы. М. : Наука, 1971.
113. *Селезнев В. Д., Денисов К.С.* Исследование свойств критериев согласия функции распределения данных с гауссовой методом Монте-Карло для малых выборок // Заводская лаборатория. 2005. Т. 71. № 1. С. 68—73.
114. *Сидельников Ю.В.* Системный анализ технологии экспертного прогнозирования. М. : Изд-во МАИ-ПРИНТ, 2007.
115. *Сидельников Ю.В.* Технология экспертного прогнозирования : учеб. пособие. 2-е изд., испр. М. : Доброе слово, 2004.
116. *Сидельников Ю.В., Танасова А.С.* Прогнозирование знака разности между ценой металла и форвардного контракта на него (на примере меди, алюминия, никеля) // Заводская лаборатория. 2006. № 11. С. 59—65.
117. *Смоляк С.А.* Интерполяция функций нескольких нечисловых переменных // Заводская лаборатория. 2007. Т. 72. № 3. С. 69—76.
118. Статистические методы повышения качества : пер. с англ. / под ред. Х. Кумэ. М. : Финансы и статистика, 1990.

119. *Суппес П., Зинес Дж.* Основы теории измерений // Психологические измерения. М. : Мир, 1967.
120. *Сычева Г.И., Колбачев Е.Б., Сычев В.А.* Оценка стоимости предприятия (бизнеса). Ростов н/Д : Феникс, 2003.
121. *Сэндидж Ч., Фрайбургер В., Ротцолл К.* Реклама: теория и практика : пер. с англ. М. : Прогресс, 1989.
122. *Толстова Ю.Н.* Измерение в социологии. М. : Инфра-М, 1998.
123. *Тюрин Ю.Н., Василевич А.П.* К проблеме обработки рядов ранжировок // Статистические методы анализа экспертных оценок : ученые записки по статистике. Т. 29. М. : Наука, 1977.
124. *Тюрин Ю.Н., Фигурнов В.Э.* О проверке датчиков псевдослучайных чисел // Заводская лаборатория. 1990. Т. 56. № 3. С. 89—92.
125. *Уилкс С.* Математическая статистика. М. : Наука, 1967.
126. *Файн В.Б., Дель М.В.* «Турнирный» метод ранжирования вариантов // Заводская лаборатория. 2005. Т. 71 № 7. С. 58—60.
127. *Фалько С.Г., Орлов А.И.* «Шесть сигм» как подход к совершенствованию бизнеса // Контроллинг. 2004. № 4 (12). С. 42—46.
128. *Фишберн П.* Теория полезности для принятия решений. М. : Наука, 1978.
129. *Фомин В.Н.* Нормирование показателей надежности. М. : Изд-во стандартов, 1986,
130. *Хан Д.* Планирование и контроль: концепция контроллинга : пер. с нем. М. : Финансы и статистика, 1997.
131. *Холлендер М., Вулф Д.* Непараметрические методы статистики. М. : Финансы и статистика, 1983.
132. Цели и принципы стандартизации / под ред. Т. Сандерса. М. : Изд-во стандартов, 1974.
133. *Шахнов И.Ф.* Некоторые модели квалиметрического анализа многофакторных объектов с бинарными факторами // Заводская лаборатория. 2005. Т. 71. № 5. С. 59—65.
134. *Шмален Г.* Основы и проблемы экономики предприятия. М. : Финансы и статистика, 1996.
135. *Шрейдер Ю.А.* Равенство, сходство, порядок. М. : Наука, 1971.
136. Экология / под ред. С.А. Боголюбова. М. : Знание, 1999.
137. Экономика предприятия : учебник / И.Э. Берзинь, С.А. Пикунова, Н.Н. Савченко, С.Г. Фалько ; под ред. С.Г. Фалько. М. : Дрофа, 2003.
138. *Фишер Рональд Э.* Использование множественных измерений в задачах таксономии // Современные проблемы кибернетики. М. : Знание, 1979.