

УДК 519.2: 005.521:633.1:004.8

UDC 519.2: 005.521:633.1:004.8

01.00.00 Физико-математические науки

Physics and mathematical sciences

**ПРИКЛАДНАЯ СТАТИСТИКА -
СОСТОЯНИЕ И ПЕРСПЕКТИВЫ****APPLIED STATISTICS – THE STATE AND
THE PROSPECTS**

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор
РИНЦ SPIN-код: 4342-4994
*Московский государственный технический
университет им. Н.Э. Баумана, Россия, 105005,
Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru*

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci.,
professor
*Bauman Moscow State Technical University,
Moscow, Russia*

Прикладная статистика - наука о том, как обрабатывать статистические данные. Как самостоятельная научно-практическая область она развивается весьма быстро. В ее состав входят многочисленные широко и глубоко развитые научные направления. Те, кто применяет прикладную статистику и другие статистические методы, обычно ориентированы на конкретные области исследования, т.е. не являются специалистами по прикладной статистике. Поэтому представляется полезным провести критический анализ современного состояния прикладной статистики и обсудить тенденции развития статистических методов. Большая практическая значимость прикладной статистики оправдывает целесообразность проведения работ по развитию ее методологии, в которых эта область научной и прикладной деятельности рассматривалась бы как целое. Дана краткая информация об истории прикладной статистики. На основе наукометрии прикладной статистики констатируем, что каждый специалист владеет лишь небольшой частью накопленных в этой области знаний. Обсуждаются пять актуальных направлений, в которых развивается современная прикладная статистика, т.е. пять "точек роста": непараметрика, робастность, бутстреп, статистика интервальных данных, нечисловая статистика. Подробнее рассмотрены основные идеи нечисловой статистики. В течение последних более чем 60 лет в России наблюдается огромный разрыв между государственной статистикой и научным сообществом специалистов по статистическим методам

Applied Statistics - the science of how to analyze the statistical data. As an independent scientific-practical area it develops very quickly. It includes numerous widely and deeply developed scientific directions. Those who use the applied statistics and other statistical methods, usually focused on specific areas of study, ie, are not specialists in applied statistics. Therefore, it is useful to make a critical analysis of the current state of applied statistics and discuss trends in the development of statistical methods. Most of the practical importance of applied statistics justifies the usefulness of the work on the development of its methodology, in which the field of scientific and applied activities would be considered as a whole. We have given some brief information about the history of applied statistics. Based on Scientometrics of Applied Statistics we state that each expert has only a small part of accumulated knowledge in this area. We discuss five topical areas in which modern applied statistics develops, ie five "points of growth": nonparametric, robustness, bootstrap, statistics of interval data, and statistics of non-numerical data. We discuss some details of the basic ideas of a non-numerical statistics. In the last more than 60 years in Russia, there has been a huge gap between official statistics and the scientific community of experts on statistical methods

Ключевые слова: МАТЕМАТИКА,
ПРИКЛАДНАЯ СТАТИСТИКА,
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА,
СТАТИСТИЧЕСКИЕ МЕТОДЫ, ТОЧКИ РОСТА,
НЕПАРАМЕТРИЧЕСКАЯ СТАТИСТИКА,
РОБАСТНОСТЬ, БУТСТРЕП, СТАТИСТИКА
ИНТЕРВАЛЬНЫХ ДАННЫХ, СТАТИСТИКА
НЕЧИСЛОВЫХ ДАННЫХ

Keywords: MATHEMATICS, APPLIED
STATISTICS, MATHEMATICAL STATISTICS,
STATISTICAL METHODS, GROWTH POINTS,
NONPARAMETRIC STATISTICS,
ROBUSTNESS, BOOTSTRAP, STATISTICS OF
INTERVAL DATA, STATISTICS OF NON-
NUMERICAL DATA

1. Введение

Прикладная статистика - наука о том, как обрабатывать статистические данные. Как самостоятельная научно-практическая область она развивается весьма быстро. В ее состав входят многочисленные широко и глубоко развитые научные направления. Те, кто применяет прикладную статистику и другие статистические методы, обычно ориентированы на конкретные области исследования, т.е. не являются специалистами по прикладной статистике. Поэтому представляется полезным провести критический анализ современного состояния прикладной статистики и обсудить тенденции развития статистических методов.

2. Что дает прикладная статистика народному хозяйству?

Такой вопрос часто задают специалисты различных областей науки, отраслей народного хозяйства, не владеющие методами прикладной статистики. В ответ нами была написана статья [1], в которой приводились многочисленные примеры успешного использования методов прикладной математической статистики. при решении практических задач. Перечень примеров можно продолжать практически безгранично. Например, можно сослаться на обобщающую монографию В.Г. Горского [2], на диссертацию А.Н. Гуды [3]. По данным Института информации Гарфилда (США) каждая из основополагающих книг В.В. Налимова [4, 5] цитировалась не менее 1000 раз (см. также монографию [6, с.270, 274, 373]). Практически в любом номере журнала "Заводская лаборатория. Диагностика материалов" есть работы, в которых те или иные методы прикладной статистики применяются для решения прикладных задач. Не раз публиковались в этом журнале и обобщающие статьи по вопросам применения прикладной статистики [7 - 10].

Итак, бесспорно совершенно, что методы прикладной статистики успешно применяются в различных отраслях народного хозяйства, практически во всех областях науки. Согласно докладу [11, с.157-158], в 1988 г. затраты в СССР на статистический анализ данных оценивались в 2 миллиарда рублей ежегодно.

Большая практическая значимость прикладной статистики оправдывает целесообразность проведения работ по развитию ее методологии, в которых эта область научной и прикладной деятельности рассматривалась бы как целое, "с высоты птичьего полета". Чтобы иметь возможность обсуждения тенденций развития прикладной статистики и других статистических методов, кратко рассмотрим их историю.

3. Об истории прикладной статистики

Типовые примеры раннего этапа применения статистических методов описаны в Ветхом Завете (см., например, Книгу Чисел - четвертую книгу Моисееву). С математической точки зрения они сводились к подсчетам числа попаданий значений наблюдаемых признаков в определенные градации. В дальнейшем результаты стали представлять в виде таблиц и диаграмм, как это и сейчас делает Федеральная служба государственной статистики (Росстат). Надо признать, что по сравнению с Ветхим Заветом есть прогресс - в Библии не было таблиц. Однако в работах Росстата нет продвижения по сравнению с работами российских статистиков конца девятнадцатого - начала двадцатого века (типовой монографией тех времен можно считать книгу [12], которая в настоящее время ещё легко доступна).

Сразу после возникновения теории вероятностей (Паскаль, Ферма, XXVII век) вероятностные модели стали использоваться при обработке статистических данных. Например, изучалась частота рождения мальчиков и девочек, было установлено отличие вероятности рождения мальчика от

0.5, анализировались причины того, что в парижских приютах эта вероятность не та, что в самом Париже, и т.д. Имеется достаточно много публикаций по истории теории вероятностей, однако в некоторых из них имеются неточные утверждения, что заставило академика Украинской АН Б.В.Гнеденко включить в шестое издание своего знаменитого курса [13] главу по истории математики случайного, выпущенную затем отдельным изданием [14].

Как установил Ф. Клейн, анализируя записные книжки К. Гаусса, тот в 1794 г. К. Гаусс разработал метод наименьших квадратов, один из наиболее популярных ныне статистических методов, и применил его при расчете орбиты астероида Церера - для борьбы с ошибками астрономических наблюдений [15]. Математические публикации К. Гаусса по методу наименьших квадратов появились лишь через 12 лет, когда близкие вопросы стали обсуждаться в литературе. В XIX веке заметный вклад в развитие практической статистики внес в развитие практической статистики внес бельгиец Ламбер Адольф Жак Кетле, на основе анализа большого числа реальных данных показавший устойчивость относительных статистических показателей, таких, как доля самоубийств среди всех смертей [16]. Важно отметить, что основные идеи статистического приемочного контроля и сертификации продукции глубоко анализировались российским академиком М.В. Остроградским и применялись в российской армии ещё в середине XIX в. [17, 18]. Статистические методы управления качеством, сертификации и классификации продукции сейчас весьма актуальны как с теоретической, так и с прикладной точек зрения [19, 20].

Современный этап развития прикладной статистики можно отсчитывать с 1900 г., когда англичанин К. Пирсон основан журнал "*Biometrika*". Мы полагаем, что именно 1900 г. следует считать началом развития прикладной математической статистики как самостоятельной

науки, хотя вначале она выступала под другими именами (биометрика, актуальные исследования и др.). Первая треть XX в. прошла под знаком параметрической статистики. Изучались методы, основанные на анализе данных из параметрических семейств распределений, описываемых кривыми семейства Пирсона. Наиболее популярным было нормальное (гауссово) распределение. Для проверки гипотез использовались критерии Пирсона, Стьюдента, Фишера. Были предложены метод максимального правдоподобия, дисперсионный анализ, сформулированы основные идеи планирования эксперимента.

Разработанную в первой трети XX в. теорию будем называть параметрической статистикой, поскольку ее основной объект изучения - это выборки из распределений, описываемых одним или небольшим числом параметров. Наиболее общим является семейство кривых Пирсона, задаваемых четырьмя параметрами. Как правило, нельзя указать каких-либо веских причин, по которым конкретное распределение результатов наблюдений должно входить в то или иное параметрическое семейство. Исключения хорошо известны: если вероятностная модель предусматривает суммирование независимых случайных величин, то сумму естественно описывать нормальным распределением; если же в модели рассматривается произведение таких величин, то итог, видимо, приближается логарифмически нормальным распределением, и т.д. Однако в подавляющем большинстве реальных ситуаций подобных моделей нет, и приближение реального распределения с помощью кривых из семейства Пирсона или его подсемейств - чисто формальная операция. Именно из таких соображений критиковал параметрическую статистику академик С.Н.Бернштейн в 1927 г. в своем докладе на Всероссийском съезде математиков [21]. Однако эта теория, к сожалению, до сих пор остается основой преподавания статистических методов и продолжает использоваться основной массой прикладников, остающихся далекими от

новых веяний в статистике. Почему так происходит? Чтобы попытаться ответить на этот вопрос, обратимся к наукометрии.

4. Наукометрия прикладной статистики

Проведенный в 1980-е годы (в ходе работ по созданию Всесоюзной статистической ассоциации [22]) анализ прикладной статистики как области научно-практической деятельности показал, в частности, что актуальными для специалистов в настоящее время являются не менее чем 100 тысяч публикаций (подробнее см. статьи [8, 23, 24]). Реально же каждый из нас знаком с существенно меньшим количеством книг и статей. Так, в наиболее обширном на русском языке сочинении по прикладной статистике - известном трехтомнике Кендалла и Стьюарта [25 - 27] - приведено около 2 тысяч литературных ссылок. При всей очевидности соображений о многократном дублировании в публикациях ценных идей приходится признать, что каждый специалист по прикладной статистике владеет лишь небольшой частью накопленных в этой области знаний. Не удивительно, что приходится постоянно сталкиваться с игнорированием или повторением ранее полученных результатов, с уходом в тупиковые (с точки зрения практики) направления исследований, с беспомощностью при обращении к реальным данным, и т.д. Все это - одно из проявлений адапционного механизма торможения развития науки, о котором еще 30 лет назад писали В.В. Налимов и другие науковеды (см., например, [28]).

Традиционный предрассудок состоит в том, что каждый новый результат, полученный исследователем - это кирпич в непрерывно растущее здание науки, который непременно будет проанализирован и использован научным сообществом. Реальная ситуация - совсем иная. Основа профессиональных знаний исследователя и инженера закладывается в период обучения. Затем они пополняются в том узком направлении, в котором работает специалист. Следующий этап - их

тиражирование новому поколению. В результате вузовские учебники отстают от современного развития на десятки лет. Так, учебники по математической статистике, по нашей экспертной оценке, в основном соответствуют 40 - 60-м годам XX в. А потому тем же годам соответствует большинство вновь публикуемых исследований и тем более - прикладных работ. Одновременно приходится признать, что результаты, которые не вошли в учебники, независимо от их ценности почти все забываются.

Активно продолжается развитие тупиковых направлений. Психологически это понятно. Приведем пример из своего опыта. В свое время по заказу Госстандарта автор настоящей статьи разработал методы оценки параметров гамма-распределения [29]. Поэтому мне близки и интересны работы по оцениванию параметров по выборкам из распределений, принадлежащих тем или иным параметрическим семействам, понятия функции максимального правдоподобия, эффективности оценок, использование неравенства Рао - Крамера и т.д. К сожалению, я знаю, что это - тупиковая ветвь, поскольку реальные данные не подчиняются каким-либо параметрическим семействам, надо применять иные статистические методы, о которых речь пойдет ниже. Понятно, что специалистам по параметрической статистике, потратившим многие годы на совершенствование в своей области, психологически трудно согласиться с этим утверждением. В том числе и автору настоящей статьи.

5. Точки роста

Отечественная литература по прикладной статистике столь же необозрима, как и мировая. Только в разделе "Математические методы исследования" журнала "Заводская лаборатория. Диагностика материалов" с 1960-х годов опубликовано более 1000 статей. Не будем даже пытаться перечислять коллективы исследователей или основные монографии в этой области (впрочем, см. статью [24]). Отметим только одно издание. По

нашему мнению, наилучшей отечественной книгой по прикладной статистике является сборник статистических таблиц Л.Н. Большева и Н.В. Смирнова [30] с подробными комментариями, играющими роль учебника и справочника по классическим вопросам прикладной статистики.

С целью управления развитием статистической науки кратко рассмотрим "точки роста" прикладной статистики [31], т.е. те ее направления, которые представляются перспективными, но пока отодвинуты на задний план традиционными постановками. Двадцать пять лет назад при описании современного этапа развития статистических методов нами были выделены [32] пять актуальных направлений, в которых развивается современная прикладная статистика, т.е. пять "точек роста": непараметрика, робастность, бутстреп, статистика интервальных данных, нечисловая статистика. Обсудим их.

Непараметрическая статистика. В первой трети XX в., одновременно с параметрической статистикой, в работах Спирмена и Кендалла появились первые непараметрические методы, основанные на коэффициентах ранговой корреляции, носящих ныне имена этих статистиков. Но непараметрика, не делаящая нереалистических предположений о том, что функции распределения результатов наблюдений принадлежат тем или иным параметрическим семействам распределений, стала заметной частью статистики лишь со второй трети XX века. В 30-е годы появились работы А.Н. Колмогорова и Н.В. Смирнова, предложивших и изучивших статистические критерии, носящие в настоящее время их имена (история этих работ подробно описана в статьях [33, 34]). Эти критерии основаны на использовании так называемого эмпирического процесса - разности между эмпирической и теоретической функциями распределения, умноженной на квадратный корень из объема выборки. В работе А.Н. Колмогорова 1933 г. изучено предельное распределение супремума модуля эмпирического процесса,

называемого сейчас статистикой критерия Колмогорова. Затем Н.В. Смирнов исследовал супремум и инфимум эмпирического процесса, а также интеграл (по теоретической функции распределения) квадрата эмпирического процесса. Следует отметить, что встречающееся иногда в литературе [35] словосочетание "критерий Колмогорова-Смирнова", как подробно обосновано в [33, 34], некорректно, поскольку эти два статистика никогда не печатались вместе и не изучали один и тот же критерий. Корректно сочетание "критерий типа Колмогорова-Смирнова", применяемое для обозначения критериев, основанных на использовании супремума функций от эмпирического процесса.

После второй мировой войны развитие непараметрической статистики пошло быстрыми темпами [36, 37]. Большую роль сыграли работы Вилкоксона и его школы. К настоящему времени с помощью непараметрических методов можно решать практически тот же круг статистических задач, что и с помощью параметрических [35]. Все большую роль играют непараметрические оценки плотности, непараметрические методы регрессии и распознавания образов (дискриминантного анализа). В нашей стране непараметрические методы получили достаточно большую известность после выхода в 1965 г. первого издания упомянутого выше сборника статистических таблиц Л.Н. Большева и Н.В. Смирнова [30], содержащего подробные таблицы для основных непараметрических критериев. Тем не менее параметрические методы всё еще популярнее непараметрических, особенно среди тех прикладников, кто слабо знаком со статистическими методами. Неоднократно публиковались (см. сводки в [4, 38]) экспериментальные данные, свидетельствующие о том, что распределения реально наблюдаемых случайных величин, в частности, ошибок измерения, в подавляющем большинстве случаев отличны от нормальных (гауссовских). Тем не менее, теоретики продолжают строить и изучать

статистические модели, основанные на гауссовости, а практики - применять подобные методы и модели. Другими словами, "ищут под фонарем, а не там, где потеряли".

Устойчивость статистических процедур (робастность). Если в параметрических постановках на данных накладываются слишком жесткие требования - их функции распределения должны принадлежать определенному параметрическому семейству, то в непараметрических, наоборот, излишне слабые - требуется лишь, чтобы функции распределения были непрерывны. При этом игнорируется априорная информация о том, каков "примерный вид" распределения. Априори можно ожидать, что учет этого "примерного вида" улучшит показатели качества статистических процедур. Развитием этой идеи является теория устойчивости (робастности) статистических процедур, в которой предполагается, что распределение исходных данных мало отличается от некоторого параметрического семейства. С 60-х годов эту теорию разрабатывали П. Хубер [39], Ф. Хампель [40] и многие другие. Из монографий на русском языке, трактующих о робастности и устойчивости статистических процедур, самой ранней и наиболее общей была книга [41], следующей - монография [42]. Современное состояние отражено в [43, 44]. Частными случаями реализации идеи робастности (устойчивости) статистических процедур являются рассматриваемые ниже статистика интервальных данных и нечисловая статистика.

Имеется большое разнообразие моделей робастности в зависимости от того, какие именно отклонения от заданного параметрического семейства допускаются. Сначала наиболее популярной [39, 40] была модель выбросов, в которой исходная выборка "засоряется" малым числом "выбросов", имеющих принципиально иное распределение. Однако эта модель представляется "тупиковой", поскольку в большинстве случаев большие выбросы либо невозможны из-за ограниченности шкалы прибора,

либо от них можно избавиться, применяя статистики, построенные по центральной части вариационного ряда. Кроме того, в подобных моделях обычно считается известной частота засорения, что в сочетании со сказанным выше делает их малопригодными для практического использования. Более перспективной представляется модель, в которой расстояние между распределением каждого элемента выборки и базовым распределением не превосходит заданной малой величины.

Бутстреп (размножение выборок). Третье из упомянутых выше направлений - бутстреп - связано с интенсивным использованием возможностей вычислительной техники. Основная идея состоит в том, чтобы теоретическое исследование дополнить или даже заменить вычислительным экспериментом [45, 46]. Вместо описания выборки распределением из параметрического семейства строим большое число "похожих" выборок, т.е. "размножаем" выборку. Затем вместо оценивания характеристик и параметров и проверки гипотез на основе свойств теоретического распределения решаем эти задачи вычислительным методом, рассчитывая интересующие нас статистики по каждой из "похожих" выборок и анализируя полученные при этом распределения. Например, вместо того, чтобы теоретическим путем находить распределение статистики, доверительные интервалы и другие характеристики, моделируем много выборок, похожих на исходную, рассчитывают соответствующие значения интересующей исследователя статистики и изучают их эмпирическое распределение. Квантили этого распределения задают доверительные интервалы, и т.д.

Термин "бутстреп" мгновенно получил известность после статьи Б. Эфрона 1979 г. [47] по этой тематике. Новый подход сразу же стал обсуждаться в массе публикаций, в том числе и научно-популярных [48]. В журнале "Заводская лаборатория. Диагностика материалов" была помещена подборка статей по бутстрепу [49], выпущен сборник статей

Б.Эфрона [50]. Основная идея бутстрепа по Б. Эфрону состоит в том, что методом Монте-Карло (статистических испытаний) многократно извлекаются выборки из эмпирического распределения. Эти выборки, естественно, являются вариантами исходной, напоминают ее.

Сама по себе идея "размножения выборок" была известна гораздо раньше. Статья Б.Эфрона [47] называется так: "Бутстреп-методы: новый взгляд на метод складного ножа". Упомянутый "метод складного ножа" (jackknife) предложен М. Кенуем еще в 1949 г., за 30 лет до статьи Б. Эфрона. "Размножение выборок" при этом осуществляется путем исключения одного наблюдения. При этом для выборки объема n получаем n "похожих" на нее выборок объема $(n - 1)$ каждая. Если же исключать по 2 наблюдения, то число "похожих" выборок возрастает до $n(n - 1) / 2$ объема $(n - 2)$ каждая.

Преимущества и недостатки бутстрепа как статистического метода обсуждаются в [51]. Там же и в [23] приводится информация о ряде аналогичных методов. Необходимо подчеркнуть, что бутстреп по Эфрону [47 - 50] - лишь один из вариантов методов "размножения выборки" (resampling), и, на наш взгляд, не самый удачный. Метод "складного ножа" представляется более полезным. На его основе можно сформулировать следующую простую практическую рекомендацию.

Вы по выборке делаете какие-либо статистические выводы и хотите узнать также, насколько эти выводы устойчивы. Если у Вас есть другие (контрольные) выборки, описывающие то же явление, то Вы можете применить к ним ту же статистическую процедуру и сравнить результаты. А если таких выборок нет? Тогда Вы можете их построить искусственно. Берете исходную выборку и исключаете один элемент. Получаете похожую выборку. Затем возвращаете этот элемент и исключаете другой. Получаете вторую похожую выборку. Поступив так со всеми элементами исходной выборки, получаете столько выборок, похожих на исходную,

каков ее объем. Остается обработать их тем же способом, что и исходную, и изучить устойчивость получаемых выводов - разброс оценок параметров, частоты принятия или отклонения гипотез и т.д.

Можно изменять не выборку, а сами данные. Поскольку всегда имеются погрешности измерения, то реальные данные - это не числа, а интервалы (результат измерения плюс-минус погрешность). Нужна статистическая теория анализа таких данных.

Статистика интервальных данных. Перспективное и быстро развивающееся направление последних лет - математическая статистика интервальных данных. Речь идет о развитии методов математической статистики в ситуации, когда статистические данные - не числа, а интервалы, в частности, порожденные наложением ошибок измерения на значения случайных величин. Полученные результаты отражены, в частности, в дискуссии [52] и в докладах международной конференции ИНТЕРВАЛ-92 [53]. Статистика интервальных данных идейно связана с интервальной математикой, в которой в роли чисел выступают интервалы (см., например, [54]). Это направление математики является дальнейшим развитием известных правил приближенных вычислений, посвященных выражению погрешностей суммы, разности, произведения, частного через погрешности тех чисел, над которыми осуществляются перечисленные операции. Как видно из докладов [53], к настоящему времени удалось решить, в частности, ряд задач теории интервальных дифференциальных уравнений, в которых коэффициенты, начальные условия и решения описываются с помощью интервалов.

Одна из ведущих научная школа в области статистики интервальных данных - это школа проф. А.П. Воцинина (1937 - 2008), активно работающая с конца 70-х годов. Полученные результаты отражены в ряде монографий (см., в частности, [55 - 57]), статей [52, 58, 59], докладов [53], диссертаций. В частности, изучены проблемы регрессионного анализа,

планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности.

Рассмотрим другое - наше - направление в статистике интервальных данных, которое также представляется перспективным. В нем развиваются асимптотические методы статистического анализа интервальных данных при больших объемах выборок и малых погрешностях измерений. В отличие от классической математической статистики, сначала устремляется к бесконечности объем выборки и только потом - уменьшаются до нуля погрешности. В частности, с помощью такой асимптотики были сформулированы правила выбора метода оценивания параметров гамма-распределения в ГОСТ 11.011-83 [29]. Основные идеи статистики интервальных данных были разработаны нами при подготовке этого стандарта в начале 80-х, однако в научной печати появились лишь в начале 90-х [60, 61]. Общая схема исследования включает расчет нотны (максимально возможного отклонения статистики, вызванного интервальностью исходных данных) и рационального объема выборки (превышение которого не дает существенного повышения точности оценивания). Она применена к оцениванию математического ожидания, дисперсии, коэффициента вариации, параметров гамма-распределения и характеристик аддитивных статистик, при проверке гипотез о параметрах нормального распределения, в т.ч. с помощью критерия Стьюдента, а также гипотезы однородности с помощью критерия Смирнова. Разработаны подходы к рассмотрению интервальных данных в основных постановках регрессионного, дискриминантного и кластерного анализов. В частности, изучено влияние погрешностей измерений и наблюдений на свойства алгоритмов регрессионного анализа, разработаны способы расчета нотн и рациональных объемов выборок, введены и исследованы новые понятия многомерных и асимптотических нотн, доказаны соответствующие предельные теоремы. Начата разработка интервального

дискриминантного анализа, в частности, рассмотрено влияние интервальности данных на введенный нами показатель качества классификации. Изучено асимптотическое поведение оценок метода моментов и оценок максимального правдоподобия (а также более общих - оценок минимального контраста), проведено асимптотическое сравнение этих методов в случае интервальных данных. Найдены общие условия, при которых, в отличие от классической математической статистики, метод моментов дает более точные оценки, чем метод максимального правдоподобия. Основные результаты разработанной нами статистики интервальных данных приведены как развернутые главы в книгах [62 - 65].

В области асимптотической математической статистики интервальных данных российская наука имеет мировой приоритет. Развертывание работ по рассматриваемой тематике позволит закрепить этот приоритет, получить теоретические результаты, основополагающие в новой области математической статистики и необходимые для обоснованного статистического анализа почти всех типов данных. Со временем во все виды статистического программного обеспечения должны быть включены алгоритмы интервальной статистики, "параллельные" обычно используемым алгоритмам прикладной математической статистики. Это позволит в явном виде учесть наличие погрешностей у результатов наблюдений, сблизить позиции метрологов и статистиков.

Статистика объектов нечисловой природы как часть прикладной статистики. Согласно классификации статистических методов, принятой в [9, 62], прикладная статистика делится на следующие четыре области: статистика (числовых) случайных величин, многомерный статистический анализ, статистика временных рядов и случайных процессов, статистика объектов нечисловой природы. Первые три из этих областей являются классическими. Остановимся на четвертой, только еще

входящей в массовое сознание специалистов. Ее именуют также статистикой нечисловых данных или нечисловой статистикой [64].

Исходный объект в математической статистике - это выборка. В вероятностной теории статистики выборка - это совокупность независимых одинаково распределенных случайных элементов. Какова природа этих элементов? В классической математической статистике (той, что обычно преподают студентам) элементы выборки - это числа. В многомерном статистическом анализе - вектора. А в нечисловой статистике элементы выборки - это объекты нечисловой природы, которые нельзя складывать и умножать на числа. Другими словами, объекты нечисловой природы лежат в пространствах, не имеющих векторной структуры.

Примерами объектов нечисловой природы являются: значения качественных признаков, т.е. результаты кодировки объектов с помощью заданного перечня категорий (градаций); упорядочения (ранжировки) экспертами образцов продукции (при оценке её технического уровня и конкурентоспособности)) или заявок на проведение научных работ (при проведении конкурсов на выделение грантов); классификации, т.е. разбиения объектов на группы сходных между собой (кластеры); толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходства тематики научных работ, оцениваемого экспертами с целью рационального формирования экспертных советов внутри определенной области науки; результаты парных сравнений или контроля качества продукции по альтернативному признаку ("годен" - "брак"), т.е. последовательности из 0 и 1; множества (обычные или нечеткие), например, зоны, пораженные коррозией, или перечни возможных причин аварии, составленные экспертами независимо друг от друга; графы; слова, предложения, тексты; вектора, координаты которых - совокупность значений разнотипных признаков, например, результат

составления статистического отчета о научно-технической деятельности (форма №1-наука) или заполненная компьютеризированная история болезни, в которой часть признаков носит качественный характер, а часть - количественный; ответы на вопросы экспертной, маркетинговой или социологической анкеты, часть из которых носит количественный характер (возможно, интервальный), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты; и т.д. Интервальные данные тоже можно рассматривать как пример объектов нечисловой природы, а именно, как частный случай нечетких множеств.

С начала 70-х годов под влиянием запросов прикладных исследований в технических, медицинских и социально-экономических науках в России активно развивается нечисловая статистика. В создании этой сравнительно новой области прикладной математической статистики приоритет принадлежит российским ученым. Большую роль сыграл основанный в 1973 г. научный семинар "Экспертные оценки и анализ данных". В 60-е годы советское научное сообщество стало интересоваться методами экспертных оценок (об их истории и современном состоянии см. [66, 67]). Началось знакомство с конкретными математизированными теориями, связанными с этими методами. Речь идет о репрезентативной теории измерений, ставшей известной в нашей стране по статье П. Суппеса и Дж. Зинеса [68] и книге И. Пфанцагля [69], о теории нечеткости Л.А. Заде [70], теории парных сравнений, описанной Г. Дэвидом [71]. К этому кругу идей примыкают теория случайных множеств (см., например, книгу Ж. Матерона [72]) и методы многомерного шкалирования (описаны А.Ю. Терехиной [73] и В.Т. Перекрестом [74]). Но наибольшее влияние оказали идеи Дж. Кемени, который аксиоматически ввел расстояние между ранжировками (теперь оно именуется в литературе расстоянием Кемени) и предложил использовать в качестве средней величины решение оптимизационной задачи (теперь - медиана Кемени). Его небольшая книга

[75], написанная в соавторстве с Дж. Снеллом, породила большой поток исследований.

В течение 70-х годов на основе запросов теории экспертных оценок (а также социологии, экономики, техники и медицины) развивались конкретные направления нечисловой статистики. Были установлены связи между конкретными видами таких объектов, разработаны для них вероятностные модели [41, 64, 76]).

Следующий этап - выделение нечисловой статистики в качестве самостоятельного направления в прикладной статистике, ядром которого являются методы статистического анализа данных произвольной природы. Программа развития этого нового научного направления впервые была сформулирована в 1979 г. в статье [77]. Реализация этой программы была осуществлена в 80-е годы. Для работ этого периода характерна сосредоточенность на внутренних проблемах нечисловой статистики. Ссылки на конкретные монографии, сборники, статьи и иные публикации нескольких десятков авторов приведены в обзорах [78 - 80]. Отметим лишь сборник [81], специально посвященный нечисловой статистике, и диссертацию [82], относящуюся к непараметрической теории парных сравнений (теории люсианов).

К 90-м годам статистика объектов нечисловой природы с теоретической точки зрения была достаточно хорошо развита, основные идеи, подходы и методы были разработаны и изучены математически, в частности, доказано достаточно много теорем. Однако она оставалась недостаточно апробированной на практике. Это было связано как с ее сравнительной молодостью, так и с общеизвестными особенностями организации науки в 80-е годы, когда отсутствовали достаточные стимулы для тому, чтобы теоретики занялись широким внедрением своих результатов. И в 90-е годы наступило время от математико-статистических исследований перейти к применению полученных результатов на

практике. Следует отметить, что в нечисловой статистике, как и в других областях прикладной математической статистики и прикладной математики вообще, одна и та же математическая схема может с успехом применяться и в технических исследованиях, и в медицине, и в социологии, и для анализа экспертных оценок, а потому ее лучше всего формулировать и изучать в наиболее общем виде, для объектов произвольной природы.

6. Основные идеи нечисловой статистики

В чем её принципиальная новизна? Для классической математической статистики характерна операция сложения. При расчете выборочных характеристик распределения (выборочное среднее арифметическое, выборочная дисперсия и др.), в регрессионном анализе и других областях этой научной дисциплины постоянно используются суммы. Математический аппарат - законы больших чисел, Центральная предельная теорема и другие теоремы - нацелены на изучение сумм. В нечисловой же статистике нельзя использовать операцию сложения, поскольку элементы выборки лежат в пространствах, где нет операции сложения. Методы обработки нечисловых данных основаны на принципиально ином математическом аппарате - на применении различных расстояний в пространствах объектов нечисловой природы.

Кратко рассмотрим несколько идей, развиваемых в статистике объектов нечисловой природы для данных, лежащих в пространствах произвольного вида. Решаются классические задачи описания данных, оценивания, проверки гипотез - но для неклассических данных, а потому неклассическими методами.

Первой обсудим проблему определения средних величин. В рамках репрезентативной теории измерений удастся указать вид средних величин, соответствующих тем или иным шкалам измерения [41]. В классической

математической статистике средние величины вводят с помощью операций сложения (выборочное среднее арифметическое, математическое ожидание) или упорядочения (выборочная и теоретическая медианы). В пространствах произвольной природы средние значения нельзя определить с помощью операций сложения или упорядочения. Теоретические и эмпирические средние приходится вводить как решения экстремальных задач. Для теоретического среднего это - задача минимизации математического ожидания (в классическом смысле) расстояния от случайного элемента со значениями в рассматриваемом пространстве до фиксированной точки этого пространства (минимизируется указанная функция от этой точки). Для эмпирического среднего математическое ожидание берется по эмпирическому распределению, т.е. берется сумма расстояний от некоторой точки до элементов выборки и затем минимизируется по этой точке. При этом как эмпирическое, так и теоретическое средние как решения экстремальных задач могут быть не единственным элементом пространства, а состоять из множества таких элементов, которое может оказаться и пустым. Тем не менее удалось сформулировать и доказать законы больших чисел для средних величин, определенных указанным образом, т.е. установить сходимость эмпирических средних к теоретическим [64].

Оказалось, что методы доказательства законов больших чисел допускают существенно более широкую область применения, чем та, для которой они были разработаны. А именно, удалось изучить асимптотику решений экстремальных статистических задач, к которым, как известно, сводится большинство постановок прикладной статистики. В частности, кроме законов больших чисел установлена и состоятельность оценок минимального контраста, в том числе оценок максимального правдоподобия и робастных оценок. К настоящему времени подобные оценки изучены также и в статистике интервальных данных.

В статистике в пространствах произвольной природы большую роль играют непараметрические оценки плотности, используемые, в частности, в различных алгоритмах регрессионного, дискриминантного, кластерного анализов. В нечисловой статистике предложен и изучен ряд типов непараметрических оценок плотности в пространствах произвольной природы, в частности, доказана их состоятельность, изучена скорость сходимости и установлен примечательный факт совпадения наилучшей скорости сходимости в произвольном случае с той, которая имеет быть в классической теории для числовых случайных величин.

Дискриминантный, кластерный, регрессионный анализы в пространствах произвольной природы основаны либо на параметрической теории - и тогда применяется подход, связанный с асимптотикой решения экстремальных статистических задач - либо на непараметрической теории - и тогда используются алгоритмы на основе непараметрических оценок плотности.

Для проверки гипотез могут быть использованы статистики интегрального типа, в частности, типа омега-квадрат. Любопытно, что предельная теория таких статистик, построенная первоначально в классической постановке [83], приобрела естественный (завершенный, изящный) вид именно для пространств произвольного вида [84], поскольку при этом удалось провести рассуждения, опираясь на базовые математические соотношения, а не на те частные (с общей точки зрения), что были связаны с конечномерным пространством.

Представляют интерес результаты, связанные с конкретными областями статистики объектов нечисловой природы, в частности, со статистикой нечетких множеств [85], со случайными множествами [41] (следует отметить, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств [41, 86]), с непараметрической

теорией парных сравнений [82], с аксиоматическим введением метрик в конкретных пространствах объектов нечисловой природы [81, 87].

Для анализа нечисловых, в частности, экспертных данных весьма важны методы классификации. С другой стороны, наиболее естественно ставить и решать задачи классификации, основанные на использовании расстояний или показателей различия, в рамках статистики объектов нечисловой природы. Это касается как распознавания образов с учителем (другими словами, дискриминантного анализа), так и распознавания образов без учителя (т.е. кластерного анализа). Современное состояние дискриминантного и кластерного анализа с точки зрения нечисловой статистики отражено в работах в [88 - 90].

Статистические методы анализа нечисловых данных особенно хорошо приспособлены для применения в экономике, социологии и экспертных оценках, поскольку в этих областях от 50% до 90% данных являются нечисловыми [91].

7. Заключение

Выше рассмотрены пять "точек роста" прикладной статистики. Разумеется, они не исчерпывают все многообразие фронта научных исследований в этой области. В частности, решены отнюдь не все проблемы, поставленные в конце 70-х годов в т.н. "цахкадзорской тетради" [92]. Кроме того, мы почти не затрагивали разнообразные применения статистических методов в конкретных прикладных областях. Много интересных проблем есть в планировании экспериментов, особенно кинетических, при анализе проблем надежности, в статистических методах управления качеством продукции, в вопросах экологии и безопасности и др.

В течение последних более чем 60 лет в России наблюдается огромный разрыв между государственной статистикой и научным

сообществом специалистов по статистическим методам (подробнее об этом см. [24]). В учебнике по истории статистики [16] даже не упоминаются имена членов-корреспондентов АН СССР Н.В. Смирнова и Л.Н. Большева! Поэтому нет ничего удивительного в том, что тенденции развития современной прикладной математической статистики столь же мало обсуждаются отечественными авторами, как и ее история.

Литература

1. Орлов А.И. Что дает прикладная статистика народному хозяйству? // Вестник статистики. 1986. № 8. С.52 – 56.
2. Горский В.Г. Планирование промышленных экспериментов. - М.: «Металлургия», 1978. - 112 с.
3. Гуда А.Н. Модели, методы и средства анализа данных в затрудненных условиях. Автореф. дисс. докт. технич. наук. - Таганрог: Таганрогский государственный радиотехнический университет, 1997. 38 с.
4. Налимов В.В. Применение математической статистики при анализе вещества. - М.: Физматгиз, 1960. - 430 с.
5. Налимов В.В., Чернова Н.Л. Статистические методы планирования экстремальных экспериментов. - М.: Физматгиз, 1965. - 340 с.
6. Налимов В.В. Канатоходец. Воспоминания. - М.: Издательская группа "Прогресс", 1994. - 456 с.
7. Гнеденко Б.В., Орлов А.И. Роль математических методов исследования в кардинальном ускорении научно-технического прогресса // Заводская лаборатория. Диагностика материалов. 1988. Т.54. №1. С.1-4.
8. Орлов А.И. О современных проблемах внедрения прикладной статистики и других статистических методов. // Заводская лаборатория. Диагностика материалов. 1992. Т.58. №1. С.67-74.
9. Орлов А.И. Современная прикладная статистика // Заводская лаборатория. Диагностика материалов. 1998. Т.64. №3. С. 52-60.
10. Горский В.Г., Орлов А.И. Математические методы исследования: итоги и перспективы // Заводская лаборатория. Диагностика материалов. 2002. Т.68. №1. С.108-112.
11. Комаров Д.М., Орлов А.И. Роль методологических исследований в разработке методоориентированных экспертных систем (на примере оптимизационных и статистических методов) // Вопросы применения экспертных систем. - Минск: Центросистем, 1988. С.151-160.
12. Ленин В.И. Развитие капитализма в России. Процесс образования внутреннего рынка для крупной промышленности. - М.: Политиздат, 1986. - XII, 610 с.
13. Гнеденко Б.В. Курс теории вероятностей: Учебник. - Изд. 6-е, перераб. и доп. - М.: Наука, Гл. ред. физ.-мат. лит., 1988. - 448 с.
14. Гнеденко Б.В. Очерк по истории теории вероятностей. – М.: Едиториал УРСС, 2001. – 88 с.
15. Клейн Ф. Лекции о развитии математики в XIX столетии. Часть I. - М. - Л.: Объединенное научно-техническое издательство НКТП СССР, 1937. - 432 с.

16. Плошко Б.Г., Елисеева И.И. История статистики: Учеб. пособие. - М.: Финансы и статистика. 1990. - 295 с.
17. Остроградский М.В. Об одном вопросе, касающемся вероятностей / Полное собрание трудов. Т.3. – Киев: Издательство Академии наук УССР, 1961. – С.215 – 237.
18. Гнеденко Б.В. Математическая статистика и контроль качества. - М.: Знание, 1976. - 64 с.
19. Орлов А.И. Сертификация и статистические методы (обобщающая статья). // Заводская лаборатория. Диагностика материалов. 1997. Т.63. №3. С. 55-62.
20. Орлов А.И. Основные проблемы контроллинга качества // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 111. С. 20-52.
21. Бернштейн С.Н. Современное состояние теории вероятностей и ее приложений // Труды Всероссийского съезда математиков в Москве 27 апреля - 4 мая 1927 г. - М.-Л.: ГИЗ, 1928. - С.50-63.
22. Орлов А.И. Создана единая статистическая ассоциация // Вестник Академии наук СССР. 1991. № 7. С.152-153.
23. Орлов А.И. Первый Всемирный конгресс Общества математической статистики и теории вероятностей им. Бернулли // Надежность и контроль качества. 1987. № 6. С. 54-59.
24. Орлов А.И. О перестройке статистической науки и её применений // Вестник статистики. 1990. № 1. С.65 - 71.
25. Кендалл М., Стьюарт А. Теория распределений. - М.: Наука, 1966. 566 с.
26. Кендалл М., Стьюарт А. Статистические выводы и связи. - М.: Наука, 1973. 899 с.
27. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976. 736 с.
28. Налимов В.В., Мульченко З.М. Наукометрия. Изучение развития науки как информационного процесса. - М.: Наука, 1969. 192 с.
29. ГОСТ 11.011-83. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения. - М.: Изд-во стандартов. 1984. - 53 с.
30. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - М.: Наука, 1965 (1-е изд.), 1968 (2-е изд.), 1983 (3-е изд.).
31. Орлов А.И. Точки роста статистических методов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 103. С. 136–162.
32. Орлов А.И. Пути развития статистических методов: непараметрика, робастность, бутстреп и реалистическая статистика // Надежность и контроль качества. 1991. № 8. С. 3-8.
33. Орлов А.И. О критериях Колмогорова и Смирнова // Заводская лаборатория. Диагностика материалов. 1995. Т.61. №7. С.59-61.
34. Орлов А.И. Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат и ошибки при их применении // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 97. С. 32-45.
35. Холлендер М., Вулф Д. Непараметрические методы статистики. - М.: Финансы и статистика, 1985. - 518 с.

36. Орлов А.И. Современное состояние непараметрической статистики // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 106. С. 239 – 269.
37. Орлов А.И. Структура непараметрической статистики (обобщающая статья) // Заводская лаборатория. Диагностика материалов. 2015. Т.81. №7. С. 62-72.
38. Орлов А.И. Часто ли распределение результатов наблюдений является нормальным? // Заводская лаборатория. Диагностика материалов. 1991. Т.57. №7. С.64-66.
39. Хьюбер П. Робастность в статистике. - М.: Мир, 1984. - 304 с.
40. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. - М.: Мир, 1989. - 512 с.
41. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.
42. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: Статистическая обработка неоднородных совокупностей. - М.: Статистика, 1980. - 208 с.
43. Орлов А.И. Устойчивые математические методы и модели // Заводская лаборатория. Диагностика материалов. 2010. Т.76. №3. С.59-67.
44. Орлов А.И. Новый подход к изучению устойчивости выводов в математических моделях // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 100. С. 146-176.
45. Орлов А.И. Компьютерно-статистические методы: состояние и перспективы // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 103. С. 163 – 195.
46. Орлов А.И. Взаимосвязь предельных теорем и метода Монте-Карло // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 114. С. 27 – 41.
47. Efron B. Bootstrap Methods: Another Look at the Jackknife // Ann. Statist. 1979. V.7. № 1. P.1-26.
48. Диаконис П., Эфрон Б. Статистические методы с интенсивным использованием ЭВМ // В мире науки. 1983. № 7. С.60-73.
49. Подборка статей по бутстрепу / Заводская лаборатория. 1987. Т.53. № 10. С.76-99.
50. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. - М.: Финансы и статистика, 1988. 263 с.
51. Орлов А.И. О реальных возможностях бутстрепа как статистического метода // Заводская лаборатория. Диагностика материалов. 1987. Т.53. №10. С.82-85.
52. Дискуссия по анализу интервальных данных / Заводская лаборатория. Диагностика материалов. 1990. Т.56. № 7. С.75-95.
53. Сборник трудов Международной конференции по интервальным и стохастическим методам в науке и технике. Тт. 1,2. - М.: МЭИ, 1992.
54. Шокин Ю.И. Интервальный анализ. - Новосибирск: Наука, 1981. 284 с.
55. Воцинин А.П. Метод оптимизации объектов по интервальным моделям целевой функции. - М.: МЭИ, 1987. 109 с.
56. Воцинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности. - М.: МЭИ - София: Техника, 1989. 224 с.
57. Воцинин А.П., Акматбеков Р.А. Оптимизация по регрессионным моделям и планирование эксперимента. - Бишкек: Илим, 1991. 164 с.

58. Вощинин А.П., Скибицкий Н.В. Интервальный подход к выражению неопределенности измерений и калибровке цифровых измерительных систем // Заводская лаборатория. Диагностика материалов. 2017. Т.73. № 11. С. 66-71.
59. Скибицкий Н.В., Севальнев Н.В. Интервальные модели в задачах оптимального управления с дифференциальными связями // Заводская лаборатория. Диагностика материалов. 2015. Т.81. № 11. С. 73-80.
60. Орлов А.И. Комментарий IV к статье А.П. Вощинина, А.Ф. Бочкова, Г.Р. Сотирова «Метод анализа данных при интервальной нестатистической ошибке» // Заводская лаборатория. Диагностика материалов. 1990. Т.56. №7. С.86-89.
61. Orlov A.I. Interval statistics // Interval Computations. 1992. № 1(3). Pp. 44-52.
62. Орлов А.И. Прикладная статистика. — М.: Экзамен, 2006. — 671 с.
63. Орлов А.И. Теория принятия решений. — М.: Экзамен, 2006. — 574 с.
64. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч. 1. Нечисловая статистика. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. — 541 с.
65. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с.
66. Орлов А.И. Организационно-экономическое моделирование: учебник : в 3 ч. Ч.2. Экспертные оценки. М.: Изд-во МГТУ им. Н.Э. Баумана, 2011. 486 с.
67. Орлов А.И. Теория экспертных оценок в нашей стране // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013. № 93. С. 1-11.
68. Суппес П., Зинес Дж. Основы теории измерений // Психологические измерения. -М.: Мир,1967. С. 9-110.
69. Пфанцагль И. Теория измерений. - М.: Мир, 1976. 166 с.
70. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. - М.: Мир, 1976. 168 с.
71. Дэвид Г. Метод парных сравнений. - М.: Статистика, 1978. 144 с.
72. Матерон Ж. Случайные множества и интегральная геометрия. - М.: Мир, 1978. 318 с.
73. Терехина А.Ю. Анализ данных методами многомерного шкалирования. - М.: Наука, 1986. 168 с.
74. Перекрест В.Т. Нелинейный типологический анализ социально-экономической информации: Математические и вычислительные методы. - Л.: Наука, 1983. 176 с.
75. Кемени Дж., Снелл Дж. Кибернетическое моделирование: Некоторые приложения. - М.: Советское радио, 1972. 192 с.
76. Тюрин Ю.Н., Литвак Б.Г., Орлов А.И., Сатаров Г.А., Шмерлинг Д.С. Анализ нечисловой информации. - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1981. - 80 с.
77. Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки // Экспертные оценки. Вопросы кибернетики. Вып.58. - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1979. С.17-33.
78. Орлов А.И. Статистика объектов нечисловой природы (Обзор) // Заводская лаборатория. Диагностика материалов. 1990. Т.56. №3. С.76-83.
79. Орлов А.И. Тридцать лет статистики объектов нечисловой природы (обзор) // Заводская лаборатория. Диагностика материалов. 2009. Т.75. №5. С.55-64.
80. Орлов А.И. О развитии статистики объектов нечисловой природы // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013. № 93. С. 41-50.

81. Анализ нечисловой информации в социологических исследованиях. - М.: Наука, 1985. - 220 с.
82. Рыданова Г.В. Некоторые вопросы статистического анализа случайных бинарных векторов. Автореф. дисс. канд. физ.-мат. наук. - М.: МГУ, 1988. 16 с.
83. Орлов А.И. Асимптотическое поведение статистик интегрального типа // Доклады АН СССР. 1974. Т.219. № 4. С. 808-811.
84. Орлов А.И. Асимптотическое поведение статистик интегрального типа // Вероятностные процессы и их приложения. - М.: МИЭМ, 1989. С.118-123.
85. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980. - 64 с.
86. Орлов А.И. Теория нечетких множеств – часть теории вероятностей // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2013. № 92. С. 51-60.
87. Орлов А.И. Расстояния в пространствах статистических данных // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета, 2014. № 101. С. 227 – 252.
88. Орлов А.И. Заметки по теории классификации // Социология: методология, методы, математические модели. 1991. № 2. С. 28-50.
89. Орлов А.И. Математические методы теории классификации // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 95. С. 23 – 45.
90. Орлов А.И. Базовые результаты математической теории классификации // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2015. № 110. С. 219–239.
91. Титма М.Х., Тоодинг Л.М. Математические методы в арсенале социолога // Социологические исследования. 1986. № 4. С. 123-128.
92. Загоруйко Н.Г., Орлов А.И. Некоторые нерешенные математические задачи прикладной статистики // Современные проблемы кибернетики (прикладная статистика). - М.: Знание, 1981. С. 53-63.

References

1. Orlov A.I. Chto daet prikladnaja statistika narodnomu hozjajstvu? // Vestnik statistiki. 1986. № 8. S.52 – 56.
2. Gorskij V.G. Planirovanie promyshlennyh jeksperimentov. - М.: «Metallurgija», 1978. - 112 s.
3. Guda A.N. Modeli, metody i sredstva analiza dannyh v zatrudnennyh uslovijah. Avtoref. diss. dokt. tehnic. nauk. - Taganrog: Taganrogskij gosudarstvennyj radiotekhnicheskij universitet, 1997. 38 s.
4. Nalimov V.V. Primenenie matematicheskoj statistiki pri analize veshhestva. - М.: Fizmatgiz, 1960. - 430 s.
5. Nalimov V.V., Chernova N.L. Statisticheskie metody planirovanija jekstremal'nyh jeksperimentov. - М.: Fizmatgiz, 1965. - 340 s.
6. Nalimov V.V. Kanatohodec. Vospominanija. - М.: Izdatel'skaja gruppa "Progress", 1994. - 456 s.
7. Gnedenko B.V., Orlov A.I. Rol' matematicheskikh metodov issledovanija v kardinal'nom uskorenii nauchno-tehnicheskogo progressa // Zavodskaja laboratorija. Diagnostika materialov. 1988. T.54. №1. S.1-4.

8. Orlov A.I. O sovremennyh problemah vnedrenija prikladnoj statistiki i drugih statisticheskikh metodov. // Zavodskaja laboratorija. Diagnostika materialov. 1992. T.58. №1. S.67-74.
9. Orlov A.I. Sovremennaja prikladnaja statistika // Zavodskaja laboratorija. Diagnostika materialov. 1998. T.64. №3. S. 52-60.
10. Gorskij V.G., Orlov A.I. Matematicheskie metody issledovanija: itogi i perspektivy // Zavodskaja laboratorija. Diagnostika materialov. 2002. T.68. №1. S.108-112.
11. Komarov D.M., Orlov A.I. Rol' metodologicheskikh issledovanij v razrabotke metodoorientirovannyh jekspertnyh sistem (na primere optimizacionnyh i statisticheskikh metodov) // Voprosy primeneniya jekspertnyh sistem. - Minsk: Centrosistem, 1988. S.151-160.
12. Lenin V.I. Razvitie kapitalizma v Rossii. Process obrazovanija vnutrennego rynka dlja krupnoj promyshlennosti. - M.: Politizdat, 1986. - XII, 610 s.
13. Gnedenko B.V. Kurs teorii verojatnostej: Uchebnik. - Izd. 6-e, pererab. i dop. - M.: Nauka, Gl. red. fiz.-mat. lit., 1988. - 448 s.
14. Gnedenko B.V. Oчерk po istorii teorii verojatnostej. – M.: Editorial URSS, 2001. – 88 s.
15. Klejn F. Lekcii o razvitii matematiki v HIIH stoletii. Chast' I. - M. -L.: Ob#edinенное nauchno-tehnicheskoe izdatel'stvo NKTP SSSR, 1937. - 432 s.
16. Ploshko B.G., Eliseeva I.I. Istoriya statistiki: Ucheb. posobie. - M.: Finansy i statistika. 1990. - 295 s.
17. Ostrogradskij M.V. Ob odnom voprose, kasajushhemsja verojatnostej / Polnoe sobranie trudov. T.3. – Kiev: Izdatel'stvo Akademii nauk USSR, 1961. – S.215 – 237.
18. Gnedenko B.V. Matematicheskaja statistika i kontrol' kachestva. - M.: Znanie, 1976. - 64 s.
19. Orlov A.I. Sertifikacija i statisticheskie metody (obobshhajushhaja stat'ja). // Zavodskaja laboratorija. Diagnostika materialov. 1997. T.63. №3. S. 55-62.
20. Orlov A.I. Osnovnye problemy kontrollinga kachestva // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 111. S. 20-52.
21. Bernshtejn S.N. Sovremennoe sostojanie teorii verojatnostej i ee prilozhenij // Trudy Vserossijskogo s#ezda matematikov v Moskve 27 aprelja - 4 maja 1927 g. - M.-L.: GIZ, 1928. - S.50-63.
22. Orlov A.I. Sozdana edinaja statisticheskaja asociacija // Vestnik Akademii nauk SSSR. 1991. № 7. S.152-153.
23. Orlov A.I. Pervyj Vsemirnyj kongress Obshhestva matematicheskoy statistiki i teorii verojatnostej im. Bernulli // Nadezhnost' i kontrol' kachestva. 1987. № 6. S. 54-59.
24. Orlov A.I. O perestrojke statisticheskoy nauki i ejo primenenij // Vestnik statistiki. 1990. № 1. S.65 - 71.
25. Kendall M., St'juart A. Teorija raspredelenij. - M.: Nauka, 1966. 566 s.
26. Kendall M., St'juart A. Statisticheskie vyvody i svjazi. - M.: Nauka, 1973. 899 s.
27. Kendall M., St'juart A. Mnogomernyj statisticheskij analiz i vremennye rjady. - M.: Nauka, 1976. 736 s.
28. Nalimov V.V., Mul'chenko Z.M. Naukometrija. Izuchenie razvitija nauki kak informacionnogo processa. - M.: Nauka, 1969. 192 s.
29. GOST 11.011-83. Prikladnaja statistika. Pravila opredelenija ocenok i doveritel'nyh granic dlja parametrov gamma-raspredelenija. - M.: Izd-vo standartov. 1984. - 53 s.

30. Bol'shev L.N., Smirnov N.V. Tablicy matematicheskoy statistiki. - M.: Nauka, 1965 (1-e izd.), 1968 (2-e izd.), 1983 (3-e izd.).
31. Orlov A.I. Točki rosta statističeskikh metodov // Politematičeskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 103. S. 136–162.
32. Orlov A.I. Puti razvitija statističeskikh metodov: neparametrika, robnost', butstrep i realističeskaja statistika // Nadežnost' i kontrol' kachestva. 1991. № 8. S. 3-8.
33. Orlov A.I. O kriterijah Kolmogorova i Smirnova // Zavodskaja laboratorija. Diagnostika materialov. 1995. T.61. №7. S.59-61.
34. Orlov A.I. Neparametričeskie kriterii soglasija Kolmogorova, Smirnova, omega-kvadrat i oshibki pri ih primenenii // Politematičeskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 97. S. 32-45.
35. Hollender M., Vulf D. Neparametričeskie metody statistiki. - M.: Finansy i statistika, 1985. - 518 s.
36. Orlov A.I. Sovremennoe sostojanie neparametričeskoy statistiki // Politematičeskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 106. S. 239 – 269.
37. Orlov A.I. Struktura neparametričeskoy statistiki (obobshhajushhaja stat'ja) // Zavodskaja laboratorija. Diagnostika materialov. 2015. T.81. №7. S. 62-72.
38. Orlov A.I. Chasto li raspredelenie rezul'tatov nabljudenij javljaetsja normal'nym? // Zavodskaja laboratorija. Diagnostika materialov. 1991. T.57. №7. S.64-66.
39. H'juber P. Robnost' v statistike. - M.: Mir, 1984. - 304 s.
40. Hampel' F., Ronchetti Je., Rausseu P., Shtajel' V. Robnost' v statistike. Podhod na osnove funkcij vlijanija. - M.: Mir, 1989. - 512 s.
41. Orlov A.I. Ustojčivost' v social'no-jekonomičeskikh modeljah. - M.: Nauka, 1979. - 296 s.
42. Smoljak S.A., Titarenko B.P. Ustojčivye metody ocenivanija: Statističeskaja obrabotka neodnorodnyh sovokupnostej. - M.: Statistika, 1980. - 208 s.
43. Orlov A.I. Ustojčivye matematičeskie metody i modeli // Zavodskaja laboratorija. Diagnostika materialov. 2010. T.76. №3. S.59-67.
44. Orlov A.I. Novyj podhod k izucheniju ustojčivosti vyvodov v matematičeskikh modeljah // Politematičeskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 100. S. 146-176.
45. Orlov A.I. Komp'juterno-statističeskie metody: sostojanie i perspektivy // Politematičeskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 103. S. 163 – 195.
46. Orlov A.I. Vzaimosvjaz' predel'nyh teorem i metoda Monte-Karlo // Politematičeskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 114. S. 27 – 41.
47. Efron B. Bootstrap Methods: Another Look at the Jackknife // Ann. Statist. 1979. V.7. № 1. P.1-26.
48. Diakonis P., Jefron B. Statističeskie metody s intensivnym ispol'zovaniem JeVM // V mire nauki. 1983. № 7. S.60-73.
49. Podborka statej po butstrepu / Zavodskaja laboratorija. 1987. T.53. № 10. S.76-99.
50. Jefron B. Netradicionnye metody mnogomernogo statističeskogo analiza. - M.: Finansy i statistika, 1988. 263 s.
51. Orlov A.I. O real'nyh vozmožnostjah butstrepa kak statističeskogo metoda // Zavodskaja laboratorija. Diagnostika materialov. 1987. T.53. №10. S.82-85.

52. Diskussija po analizu interval'nyh dannyh / Zavodskaja laboratorija. Diagnostika materialov. 1990. T.56. № 7. S.75-95.
53. Sbornik trudov Mezhdunarodnoj konferencii po interval'nym i stohasticheskim metodam v nauke i tehnike. Tt. 1,2. - M.: MJeI, 1992.
54. Shokin Ju.I. Interval'nyj analiz. - Novosibirsk: Nauka, 1981. 284 s.
55. Voshhinin A.P. Metod optimizacii ob'ektov po interval'nym modeljam celevoj funkcii. - M.: MJeI, 1987. 109 s.
56. Voshhinin A.P., Sotirov G.R. Optimizacija v uslovijah neopredelennosti. - M.: MJeI - Sofija: Tehnika, 1989. 224 s.
57. Voshhinin A.P., Akmatbekov R.A. Optimizacija po regressionnym modeljam i planirovanie jeksperimenta. - Bishkek: Ilim, 1991. 164 s.
58. Voshhinin A.P., Skibickij N.V. Interval'nyj podhod k vyrazheniju neopredelennosti izmerenij i kalibrovke cifrovych izmeritel'nyh sistem // Zavodskaja laboratorija. Diagnostika materialov. 2017. T.73. № 11. S. 66-71.
59. Skibickij N.V., Seval'nev N.V. Interval'nye modeli v zadachah optimal'nogo upravlenija s differencial'nymi svjazzjami // Zavodskaja laboratorija. Diagnostika materialov. 2015. T.81. № 11. S. 73-80.
60. Orlov A.I. Kommentarij IV k stat'e A.P. Voshhinina, A.F. Bochkova, G.R. Sotirova «Metod analiza dannyh pri interval'noj nestatisticheskoj oshibke» // Zavodskaja laboratorija. Diagnostika materialov. 1990. T.56. №7. S.86-89.
61. Orlov A.I. Interval statistics // Interval Computations. 1992. № 1(3). Rr. 44-52.
62. Orlov A.I. Prikladnaja statistika. — M.: Jekzamen, 2006. — 671 s.
63. Orlov A.I. Teorija prinjatija reshenij. — M.: Jekzamen, 2006. — 574 s.
64. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch. 1. Nechislovaja statistika. – M.: Izd-vo MGTU im. N.Je. Baumana, 2009. — 541 s.
65. Orlov A.I., Lucenko E.V. Sistemnaja nechetkaja interval'naja matematika. Monografija (nauchnoe izdanie). – Krasnodar, KubGAU. 2014. – 600 s.
66. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie: uchebnik : v 3 ch. Ch.2. Jekspertnye ocenki. M.: Izd-vo MGTU im. N.Je. Baumana, 2011. 486 s.
67. Orlov A.I. Teorija jekspertnyh ocenok v nashej strane // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2013. № 93. S. 1-11.
68. Suppes P., Zines Dzh. Osnovy teorii izmerenij // Psihologicheskie izmerenija. - M.: Mir, 1967. S. 9-110.
69. Pfancagl' I. Teorija izmerenij. - M.: Mir, 1976. 166 s.
70. Zade L. Ponjatie lingvisticheskoj peremennoj i ego primenenie k prinjatiju priblizennyh reshenij. - M.: Mir, 1976. 168 s.
71. Djevid G. Metod parnyh sravnenij. - M.: Statistika, 1978. 144 s.
72. Materon Zh. Sluchajnye mnozhestva i integral'naja geometrija. - M.: Mir, 1978. 318 s.
73. Terehina A.Ju. Analiz dannyh metodami mnogomernogo shkalirovanija. - M.: Nauka, 1986. 168 s.
74. Perekrest V.T. Nelinejnyj tipologicheskij analiz social'no-jekonomicheskoi informacii: Matematicheskie i vychislitel'nye metody. - L.: Nauka, 1983. 176 s.
75. Kemeni Dzh., Snell Dzh. Kiberneticheskoe modelirovanie: Nekotorye prilozhenija. - M.: Sovetskoe radio, 1972. 192 s.
76. Tjurin Ju.N., Litvak B.G., Orlov A.I., Satarov G.A., Shmerling D.S. Analiz nechislovoj informacii. - M.: Nauchnyj Sovet AN SSSR po kompleksnoj probleme "Kibernetika", 1981. - 80 s.

77. Orlov A.I. Statistika ob#ektov nechislovoj prirody i jekspertnye ocenki // Jekspertnye ocenki. Voprosy kibernetiki. Vyp.58. - M.: Nauchnyj Sovet AN SSSR po kompleksnoj probleme "Kibernetika", 1979. S.17-33.
78. Orlov A.I. Statistika ob#ektov nechislovoj prirody (Obzor) // Zavodskaja laboratorija. Diagnostika materialov. 1990. T.56. №3. S.76-83.
79. Orlov A.I. Tridcat' let statistiki ob#ektov nechislovoj prirody (obzor) // Zavodskaja laboratorija. Diagnostika materialov. 2009. T.75. №5. S.55-64.
80. Orlov A.I. O razvitii statistiki ob#ektov nechislovoj prirody // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2013. № 93. S. 41-50.
81. Analiz nechislovoj informacii v sociologicheskikh issledovanijah. - M.: Nauka, 1985. - 220 s.
82. Rydanova G.V. Nekotorye voprosy statisticheskogo analiza sluchajnyh binarnyh vektorov. Avtoref. diss. kand. fiz.-mat. nauk. - M.: MGU, 1988. 16 s.
83. Orlov A.I. Asimptoticheskoe povedenie statistik integral'nogo tipa // Doklady AN SSSR. 1974. T.219. № 4. S. 808-811.
84. Orlov A.I. Asimptoticheskoe povedenie statistik integral'nogo tipa // Verojatnostnye processy i ih prilozhenija. - M.: MIJeM, 1989. S.118-123.
85. Orlov A.I. Zadachi optimizacii i nechetkie peremennye. - M.: Znanie, 1980. - 64 s.
86. Orlov A.I. Teorija nechetkih mnozhestv – chast' teorii verojatnostej // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2013. № 92. S. 51-60.
87. Orlov A.I. Rasstojanija v prostranstvah statisticheskikh dannyh // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta, 2014. № 101. S. 227 – 252.
88. Orlov A.I. Zametki po teorii klassifikacii // Sociologija: metodologija, metody, matematicheskie modeli. 1991. № 2. S. 28-50.
89. Orlov A.I. Matematicheskie metody teorii klassifikacii // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 95. S. 23 – 45.
90. Orlov A.I. Bazovye rezul'taty matematicheskoj teorii klassifikacii // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2015. № 110. S. 219–239.
91. Titma M.H., Tooding L.M. Matematicheskie metody v arsenale sociologa // Sociologicheskie issledovanija. 1986. № 4. S. 123-128.
92. Zagorujko N.G., Orlov A.I. Nekotorye nereshennye matematicheskie zadachi prikladnoj statistiki // Sovremennye problemy kibernetiki (prikladnaja statistika). - M.: Znanie, 1981. S. 53-63.