

УДК 519.2:330.4(075.8)

ББК 60.6я73

О-66

Рецензенты:

кафедра «Системы управления экономическими объектами»
Московского государственного авиационного института (ГУ)
(заведующий кафедрой доктор экономических наук,
профессор *В.Д. Калачанов*);
доктор экономических наук, профессор *С.А. Смоляк*

Орлов А.И.

О-66 Организационно-экономическое моделирование : учебник : в 3 ч. / А.И. Орлов. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2009.

ISBN 978-5-7038-3276-9

Ч. 1 : Нечисловая статистика. – 2009. – 541 с. : ил.

ISBN 978-5-7038-3277-6

Впервые в систематизированном виде рассмотрена одна из четырех основных областей современной прикладной математической статистики — нечисловая статистика.

Проанализированы основные виды нечисловых данных, методология, процедуры и особенности их статистического анализа. Представлены статистические методы в пространствах произвольной природы, статистика нечисловых данных конкретных видов, статистика интервальных данных. Большое внимание уделено практическому применению методов и результатов нечисловой статистики.

Материал учебника соответствует курсам лекций, которые автор читает в МГТУ им. Н.Э. Баумана.

Для студентов, преподавателей и специалистов, заинтересованных в применении современных статистических методов в технике, экономике, управлении, медицине, социологии и иных областях, для разработчиков таких методов и соответствующего программного обеспечения, а также для исследователей в области прикладной и математической статистики, анализа данных, методов оптимизации, математического и организационно-экономического моделирования.

УДК 519.2:330.4(075.8)

ББК 60.6я73

© Орлов А.И., 2009

© МГТУ им. Н.Э. Баумана, 2009

© Оформление. Издательство

МГТУ им. Н.Э. Баумана, 2009

ISBN 978-5-7038-3277-6 (ч. 1)

ISBN 978-5-7038-3276-9

Оглавление

| | |
|---|-----|
| ПРЕДИСЛОВИЕ | 5 |
| ВВЕДЕНИЕ. Нечисловая статистика — основа высоких статистических технологий | 15 |
| В1. О развитии статистических методов | 15 |
| В2. Структура нечисловой статистики | 42 |
| Литература | 57 |
| Глава 1. Нечисловые статистические данные | 61 |
| 1.1. Количественные и категоризованные данные | 61 |
| 1.2. Основы теории измерений | 69 |
| 1.3. Виды нечисловых данных | 78 |
| 1.4. Вероятностные модели порождения нечисловых данных | 98 |
| 1.5. Нечеткие множества — частный случай нечисловых данных ... | 121 |
| 1.6. Сведение нечетких множеств к случайным | 136 |
| 1.7. Данные и расстояния в пространствах произвольной природы | 149 |
| 1.8. Аксиоматическое введение расстояний | 156 |
| Литература | 169 |
| Контрольные вопросы и задачи | 173 |
| Темы докладов, рефератов, исследовательских работ | 174 |
| Глава 2. Статистические методы в пространствах произвольной природы | 176 |
| 2.1. Эмпирические и теоретические средние | 176 |
| 2.2. Законы больших чисел | 187 |
| 2.3. Экстремальные статистические задачи | 198 |
| 2.4. Одношаговые оценки | 202 |
| 2.5. Непараметрические оценки плотности | 214 |
| 2.6. Статистики интегрального типа | 224 |
| 2.7. Методы восстановления зависимостей | 247 |
| 2.8. Методы классификации | 258 |
| 2.9. Методы шкалирования | 284 |
| Литература | 294 |
| Контрольные вопросы и задачи | 299 |
| Темы докладов, рефератов, исследовательских работ | 301 |
| Глава 3. Статистика нечисловых данных конкретных видов | 302 |
| 3.1. Инвариантные алгоритмы и средние величины | 302 |
| 3.2. Теория случайных толерантностей | 311 |
| 3.3. Метод проверки гипотез по совокупности малых выборок | 323 |
| 3.4. Теория люсианов | 336 |

| | |
|---|------------|
| 3.5. Метод парных сравнений | 356 |
| 3.6. Статистика нечетких множеств | 365 |
| 3.7. Статистика нечисловых данных в экспертных оценках | 373 |
| Литература | 396 |
| Контрольные вопросы и задачи | 399 |
| Темы докладов, рефератов, исследовательских работ | 402 |
| Глава 4. Статистика интервальных данных | 404 |
| 4.1. Основные идеи статистики интервальных данных | 404 |
| 4.2. Интервальные данные в задачах оценивания | 416 |
| 4.3. Интервальные данные в задачах проверки гипотез | 457 |
| 4.4. Линейный регрессионный анализ интервальных данных | 461 |
| 4.5. Интервальный дискриминантный анализ | 496 |
| 4.6. Интервальный кластер-анализ | 499 |
| 4.7. Интервальные данные в инвестиционном менеджменте | 503 |
| 4.8. Статистика интервальных данных в прикладной статистике | 509 |
| Литература | 513 |
| Контрольные вопросы и задачи | 517 |
| Темы докладов, рефератов, исследовательских работ | 518 |
| Приложение 1. Теоретическая база нечисловой статистики | 519 |
| П1. Законы больших чисел | 519 |
| П2. Центральные предельные теоремы | 522 |
| П3. Теоремы о наследовании сходимости | 527 |
| П4. Метод линеаризации | 533 |
| П5. Принцип инвариантности | 535 |
| Литература | 538 |
| Приложение 2. Основные книги профессора А.И. Орлова | 539 |

Предисловие

В учебнике впервые систематически рассматривается одна из четырех основных областей современной прикладной статистики — нечисловая статистика. Она порождена в 70-х годах XX в. потребностями прикладных социально-экономических, технических и медико-биологических исследований. Основа ее математического аппарата заключается в использовании расстояний между объектами нечисловой природы и решений оптимизационных задач, а не операций суммирования данных, как в других областях статистики. В учебнике рассмотрены основные виды нечисловых данных и особенности их статистического анализа. Большое внимание уделено проблемам практического применения рассматриваемых методов и результатов.

Нечисловую статистику называют также статистикой нечисловых данных или статистикой объектов нечисловой природы. Она — сердцевина высоких статистических технологий, т.е. современной прикладной статистики. Ее рассматривают также как одну из четырех основных областей статистики. Три других — это статистика чисел (случайных величин), статистика векторов (многомерный статистический анализ), статистика функций (временных рядов и случайных процессов).

Какие данные называют нечисловыми? Описание технического, социально-экономического, медицинского объекта изучения часто удается представить в виде вектора, часть координат которого измерена по количественным шкалам, а часть — по качественным, имеющим конечное число градаций. Это наиболее распространенный тип нечисловых данных.

В общем случае под нечисловыми данными понимают элементы пространств, не являющихся линейными (вектор-

ными), в которых нет операций сложения элементов и их умножения на действительное число. Кроме результатов измерений по качественным признакам, примерами являются последовательности из 0 и 1, бинарные отношения (ранжировки, разбиения, толерантности); множества (в том числе плоские изображения и объемные тела); нечеткие (размытые, расплывчатые, *fuzzy*) числа и множества, их частный случай — интервалы; результаты парных сравнений и другие объекты, возникающие в прикладных исследованиях. Все эти виды нечисловых данных и вероятностные модели их порождения подробно рассмотрены в учебнике. Их обобщением, как и обобщением числовых данных (чисел, векторов, функций), являются элементы пространств произвольной природы.

Исторически нечисловые данные стали рассматриваться раньше, чем статистические данные в виде действительных чисел. Книга «Числа» Ветхого Завета содержит обширные сведения о численностях тех или иных совокупностей. Натуральные числа можно отнести к нечисловым данным, хотя их можно складывать, но умножение на действительное число выводит за пределы натурального ряда. Теория вероятностей также начиналась с моделирования нечисловых данных. Это результаты бросания игральных костей и вытаскивания шаров из урн. Однако к началу XX в. основное внимание статистиков переместилось на рассмотрение числовых случайных величин, моделирующих действительно-значные результаты наблюдений.

К 70-м годам XX в. развитие прикладных научных исследований в инженерном деле, социологии, экономике, менеджменте, психологии, медицине и других областях привело к необходимости разработки методов статистического анализа нечисловых данных. В СССР вокруг Всесоюзного семинара «Экспертные оценки и нечисловая статистика» сложился неформальный научный коллектив из нескольких десятков активных исследователей.

Сначала изучались методы анализа конкретных видов нечисловых данных, устанавливались связи между ними. Затем пришло понимание статистики нечисловых данных как самостоятельной области прикладной статистики со своей внутренней структурой и разнообразными связями между подходами и результатами, относящимися к тем или иным видам нечисловых данных.

Статистика нечисловых данных выделена нами как самостоятельная область прикладной статистики в 1979 г. За прошедшие с тех пор годы арсенал ее методов пополнился многими полезными новшествами. Но основные идеи выдержали проверку временем, что и оправдывает их изложение в настоящей книге.

О развитии нечисловой статистики. В 70-е годы XX в. в СССР возник неформальный научный коллектив исследователей, изучающих методы анализа нечисловых данных различных видов. Центром стал научный семинар «Экспертные оценки и нечисловая статистика» и одноименная комиссия в составе Научного совета АН СССР по комплексной проблеме «Кибернетика».

Вначале разбирались подходы предшественников: аксиоматическое введение расстояний между объектами нечисловой природы и нахождение среднего по Кемени, репрезентативная теория измерений, нечеткие множества Заде, парные сравнения по Дэвиду и др. Затем были проведены многочисленные самостоятельные исследования, установлены взаимосвязи между подходами и результатами для различных типов нечисловых данных, разработана общая теория статистического анализа нечисловых данных произвольной природы.

В итоге стало возможным говорить о новой области прикладной статистики — нечисловой статистике. Время ее окончательного формирования — первая половина 1980-х годов — время наибольшей организационной активности. Две Всесо-

юзные конференции — в Алма-Ате (1981 г.) и Таллине (1984 г.) — собрали по 300–500 участников.

Со второй половины 80-х годов XX в. статистика нечисловых данных (статистика объектов нечисловой природы) развивается стабильно. Много публикаций содержится в журналах «Заводская лаборатория», «Социология: методология, методы, математическое моделирование», в периодических сборниках «Статистические методы оценивания и проверки гипотез». Разделу нечисловой статистики — статистике интервальных данных — была посвящена Международная конференция по интервальным и стохастическим методам в науке и технике «Интервал–92» (г. Калининград Московской области, сентябрь 1992 г.).

Неформальный коллектив по нечисловой статистике включает в себя десятки российских исследователей, а если учитывать авторов одной-двух работ, то и сотни. Почти за 30 лет выпущено несколько десятков сборников и монографий, много статей опубликовано в научных журналах. Однако из-за отсутствия формальной инфраструктуры (в частности, Института нечисловой статистики в составе Российской академии наук) имеются лишь единичные методики и программные продукты, предназначенные для практического использования. В отличие от научных монографий почти отсутствуют учебники и учебные пособия, а также книги, содержащие введение и общий обзор нечисловой статистики.

Настоящая книга заполняет существенный пробел в литературе по нечисловой статистике, вводит в предмет, знакомит с нечисловой статистикой на современном научном уровне. Изложение материала доводится до переднего края ведущихся в настоящее время научных исследований. Постоянно в поле зрения находятся вопросы практического применения рассматриваемых подходов, методов, результатов. В частности, используется опыт разработки нашим коллективом автомати-

зированной рабочего места МАТЭК (математика в экспертизе) для организатора экспертного опроса. Отражены также работы по статистике нечисловых данных и ее применениям, за которые автору в 1992 г. присуждена ученая степень доктора технических наук (по научному докладу об опубликованных работах, т.е. без написания диссертации классического вида).

Чтобы в сравнительно небольшой книге охватить всю статистику нечисловых данных, приходится идти на жертвы. Мы отказываемся от разбора большинства доказательств, отсылая читателей к публикациям, содержащим эти доказательства. Примеры подобного стиля изложения — обзоры по статистике нечисловых данных в разделе «Математические методы исследования» журнала «Заводская лаборатория» (1990, № 3; 1995, № 3, 5; 1996, № 3).

Стиль книги. В любой использующей математику области есть три уровня исследований — методологический, теоретический и практический. На методологическом уровне излагаются общие подходы и формулируются основные результаты. На теоретическом уровне, грубо говоря, доказывают теоремы. В частности, выявление необходимых и достаточных условий регулярности обычно осуществляется в результате цепи работ этого уровня.

Например, на методологическом уровне Центральная предельная теорема теории вероятностей формулируется так: «При некоторых условиях регулярности распределение централизованной и нормированной суммы независимых случайных величин при росте числа слагаемых стремится к стандартному нормальному распределению». Около двухсот лет — от Муавра и Лапласа до Линдберга и Феллера — некоторые условия регулярности уточнялись в работах теоретического уровня.

Изложение материала в книге идет в основном на методологическом уровне. При спуске на теоретический уровень приводятся формулировки теорем, в основном без доказа-

тельств, но со ссылками на публикации, где они содержатся. Обоснованием для выбора такого варианта построения книги, кроме желания ограничить ее объем разумными рамками, послужило следующее представление о предпочтениях будущих читателей: большинство из них не извлечет пользы из того, что в некоторой формулировке можно заменить требование дифференцируемости определенной функции на требование ее непрерывности. Сказанное не означает, что автор отрицает целесообразность проведения научных работ, посвященных подобным ослаблениям условий регулярности. Просто им не место в книге для первого знакомства с нечисловой статистикой.

На практическом уровне исследований большое внимание уделяют конкретному объекту приложений — технической, социально-экономической или медицинской системе. Для достаточно информативного описания каждого такого исследования нужна отдельная монография, которая обычно и готовится в качестве отчета по работе. Поэтому мы ограничились краткими замечаниями о практическом применении различных методов нечисловой статистики. Суммарно эти замечания составляют существенную часть как авторского замысла, так и объема книги.

Содержание книги. Во введении обсуждены история и современное состояние статистических методов, прикладной статистики, место в ней статистики нечисловых данных. Проанализирована сложившаяся структура нечисловой статистики — сердцевины высоких статистических технологий.

Книга делится на главы, а главы — на разделы.

В *главе 1* изучены конкретные виды нечисловых статистических данных, соответствующие вероятностные модели. Сопоставлены количественные и категоризованные данные. Разобраны основы теории измерений. Большое внимание уделено нечетким множествам как частному виду нечисловых данных. Продемонстрирована возможность сведения теории

нечетких множеств к теории случайных множеств. Обсуждены статистические данные и необходимые для их анализа расстояния в пространствах произвольной природы; аксиоматический подход к введению расстояний и показателей различия в различных пространствах объектов нечисловой природы.

В *главе 2* получили развитие статистические методы анализа данных произвольного вида, лежащих в метрическом пространстве или в пространстве с мерой различия. Эмпирические и теоретические средние приходится определять как решения экстремальных статистических задач, и законы больших чисел оказываются частными случаями утверждений об асимптотическом поведении решений таких задач. Другие классы частных случаев подобных утверждений связаны с теорией одношаговых оценок параметров распределения вероятностей (они имеют преимущества по сравнению с оценками максимального правдоподобия) и с оптимизационными постановками основных задач прикладной статистики, в том числе задач восстановления зависимостей, классификации, шкалирования и снижения размерности. Для описания распределений нечисловых данных разработаны непараметрические оценки плотности, используемые в регрессионном, дискриминантном и кластерном анализе. В предельной теории статистик интегрального типа найден ряд необходимых и достаточных условий.

Глава 3 посвящена статистическому анализу конкретных видов нечисловых данных. В рамках репрезентативной теории измерений получены характеристики средних величин свойством устойчивости результата сравнения средних относительно той или иной группы допустимых преобразований шкалы. Изучены случайные толерантности. Метод проверки гипотез по совокупности малых выборок применен в теории люсианов — конечных последовательностей испытаний Бернулли с различными вероятностями успеха. Люсианы находят применение в теории парных сравнений. Рассмотрены основ-

ные вопросы статистики нечетких множеств. Обсуждено использование нечисловой статистики в теории и практике экспертных оценок — области исследований, во многом стимулировавшей развитие основных идей статистического анализа нечисловых данных.

Глава 4 посвящена основным подходам и результатам статистики интервальных данных, быстро развивающейся в последние годы. Для интервальных данных решен ряд задач оценивания и проверки гипотез. Построены интервальные аналоги регрессионного, дискриминантного и кластерного анализов. Интервальные данные применены в инвестиционном менеджменте. Рассмотрена роль статистики интервальных данных в прикладной статистике.

В *приложении 1* включены некоторые вопросы, относящиеся к теоретической базе нечисловой статистики. Рассмотрены классические законы больших чисел, центральные предельные теоремы, метод линеаризации и принцип инвариантности. Теоремы о наследовании сходимости сравнительно малоизвестны и могут представлять особый интерес. В *приложении 2* содержится информация об авторе, позволяющая читателям лучше понять происхождение идей, изложению которых посвящена настоящая книга.

Нумерация формул, определений, теорем, таблиц, рисунков — своя в каждом разделе. Литература приводится по главам в порядке первого упоминания. Списки литературы включают основные публикации по нечисловой статистике, а также работы, на которые даются ссылки в тексте. Они не претендуют на полноту хотя бы потому, что перечень известных автору публикаций по рассматриваемой тематике по объему превысил бы настоящую книгу в несколько раз.

Для кого эта книга? Она предназначена для широкого круга читателей — студентов и преподавателей, специалистов, занимающихся прикладной наукой, и математиков. Для ее чтения достаточно знаний в объеме вводного курса математической

статистики, включающего основные задачи описания данных, оценивания и проверки гипотез.

Эта книга — учебник для студентов различных специальностей, прежде всего технических, управленческих и экономических, для слушателей институтов повышения квалификации, структур послевузовского (в том числе второго) образования, в частности программ *MBA* («Мастер делового администрирования»), для преподавателей вузов. Учебник будет полезен инженерам, менеджерам, экономистам, социологам, биологам, медикам, психологам, историкам и другим специалистам, самостоятельно повышающим свой научный уровень, — короче, всем научным и практическим работникам, связанным с анализом данных.

Учебник может быть использован при изучении дисциплин, полностью или частично посвященных методам анализа нечисловых результатов наблюдений (измерений, испытаний, опытов). Типовые названия таких курсов — «Прикладная статистика», «Эконометрика», «Анализ данных», «Статистический анализ», «Теория принятия решений», «Управленческие решения», «Экономико-математическое моделирование», «Прогнозирование», «Хемометрия», «Математические методы в социологии» и т.п. Учебник необходим студентам специальности «Менеджмент высоких технологий», особенно при изучении учебной дисциплины «Организационно-экономическое моделирование».

Книга полезна широкому кругу специалистов, заинтересованных в применении современных статистических методов анализа нечисловых данных в любой предметной области; необходима разработчикам таких методов и соответствующего программного обеспечения, т.е. специалистам по прикладной статистике.

Специалистам по теории вероятностей и математической статистике эта книга может быть интересна и полезна, поскольку в ней описан современный взгляд на прикладную ма-

тематическую статистику, основные подходы и результаты в этой области, открывающие большой простор для дальнейших математических исследований.

Книга представляет интерес для исследователей — специалистов по вопросам управления, в том числе по принятию решений, методам оптимизации и математическому моделированию. Без нее не сможет обойтись ни один преподаватель прикладной или математической статистики, статистических методов для любой конкретной области применений, если он хочет, чтобы его лекционный курс был современным.

Благодарности. Автор благодарен за полезные дискуссии многочисленным коллегам по научным семинарам, по работе в Институте высоких статистических технологий и эконометрики МГТУ им. Н.Э. Баумана, в Российской ассоциации статистических методов и Российской академии статистических методов.

Автор признателен научному редактору В.И. Осипову, всем сотрудникам издательства за поддержку нашего научного направления и большую работу по подготовке рукописи к изданию.

С текущей научной информацией по статистическим методам можно познакомиться на сайте «Высокие статистические технологии» (<http://orlovs.pp.ru>) и 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 имени Н.Э. Баумана. Достаточно большой объем информации содержит еженедельник «Эконометрика», выпускаемый с июля 2000 г. (<http://subscribe.ru/catalog/science.humanity.econometrika>). Автор искренне благодарен разработчику сайта и редактору электронного еженедельника А.А. Орлову за многолетний энтузиазм.

Автор просит читателей сообщать свои вопросы и замечания по адресу издательства или непосредственно автору по электронной почте: prof-orlov@mail.ru.

Введение. Нечисловая статистика — основа высоких статистических технологий

В1. О развитии статистических методов

Четыре столетия статистики. Впервые термин «статистика» появился в 1602 г. в трагедии У. Шекспира «Гамлет» (акт 5, сцена 2). Смысл этого слова у Шекспира — знать, придворные. По-видимому, оно происходит от лат. *status*, что в оригинале означает «состояние» или «политическое состояние».

В течение следующих 400 лет термин «статистика» понимали и понимают по-разному. В работе [1] собрано более 200 определений этого термина, некоторые из них обсуждаются ниже.

Вначале под статистикой понимали описание экономического и политического состояния государства или его части. Например, к 1792 г. относится определение: «Статистика описывает состояние государства в настоящее время или в некоторый известный момент в прошлом». И в настоящее время деятельность государственных статистических служб (в нашей стране — Федеральная служба государственной статистики (Росстат)) вполне укладывается в это определение.

Однако постепенно термин «статистика» стал использоваться шире. По Наполеону Бонапарту, «Статистика — это бюджет вещей». Тем самым статистические методы признаны полезными не только для административного управления, но и на уровне отдельного предприятия. Согласно формулировке 1833 г., «цель статистики в представлении фактов в наиболее сжатой форме». Приведем еще два высказывания: «Статисти-

Глава 1. Нечисловые статистические данные

1.1. Количественные и категоризованные данные

Статистические методы — методы анализа данных, причем обычно достаточно большого количества данных. Статистические данные могут иметь различную природу. Исторически самыми ранними были два вида данных — сведения о числе объектов, удовлетворяющих тем или иным условиям, и числовые результаты измерений.

Первый из этих видов данных до сих пор главенствует в статистических сборниках Росстата. Такого рода данные часто называют *категоризованными*, поскольку о каждом из рассматриваемых объектов известно, в какую из нескольких заранее заданных категорий он попадает. Примером является информация Росстата о населении страны с разделением по возрастным категориям и полу. Часто при составлении таблиц жертвуют информацией, заменяя точное значение измеряемой величины на указание интервала группировки, в которую это значение попадает. Например, вместо точного возраста человека используют лишь один из указанных в таблице возрастных интервалов.

Второй, наиболее распространенный вид данных — *количественные* данные, рассматриваемые как действительные числа. Таковы результаты измерений, наблюдений, испытаний, опытов, анализов. Количественные данные обычно описываются набором чисел (выборкой), а не таблицей.

Нельзя утверждать, что категоризованные данные соответствуют первому этапу исследования, а числовые — следующе-

Глава 2. Статистические методы в пространствах произвольной природы

2.1. Эмпирические и теоретические средние

Одна из основных статистических процедур — вычисление средних величин для тех или иных совокупностей данных. Законы больших чисел гласят: эмпирические средние сходятся к теоретическим. В классическом варианте: выборочное среднее арифметическое при определенных условиях сходится по вероятности при росте числа слагаемых к математическому ожиданию. На основе законов больших чисел обычно доказывают состоятельность различных статистических оценок. В целом эта тематика занимает заметное место в теории вероятностей и математической статистике.

Однако математический аппарат при этом основан на свойствах сумм случайных величин (векторов, элементов линейных пространств). Следовательно, он не пригоден для изучения вероятностных и статистических проблем, связанных со случайными объектами нечисловой природы. Это такие объекты, как бинарные отношения, нечеткие множества, вообще элементы пространств без векторной структуры. Объекты нечисловой природы все чаще встречаются в прикладных исследованиях. Много конкретных примеров приведено в главе 1. Поэтому необходимо научиться усреднять различные нечисловые данные, т.е. определять эмпирические и теоретические средние в пространствах произвольной природы. Кроме того, представляется полезным получение законов больших чисел в пространствах нечисловой природы.

Глава 3. Статистика нечисловых данных конкретных видов

От статистики в пространствах произвольной природы перейдем к обсуждению проблем анализа конкретных видов нечисловых данных. Начнем с задач обработки результатов измерений в шкалах, отличных от абсолютных.

3.1. Инвариантные алгоритмы и средние величины

Инвариантные алгоритмы и средние величины. Основное требование к алгоритмам анализа данных в теории измерений (см. гл. 1) формулируется так: *выводы, сделанные на основе данных, измеренных в шкале определенного типа, не должны меняться при допустимом преобразовании шкалы измерения этих данных.* Другими словами, выводы должны быть *инвариантны* по отношению к допустимым преобразованиям шкалы.

Таким образом, одна из основных целей теории измерений — борьба с субъективизмом исследователя при приписывании численных значений реальным объектам. Так, расстояния можно измерять в аршинах, метрах, микронах, милях, парсеках и других единицах измерения; массу (вес) — в пудах, килограммах, фунтах и других единицах измерения. Цены на товары и услуги можно указывать в юанях, рублях, тенге, гривнах, латах, кронах, марках, долларах США и иных валютах (при фиксированных курсах пересчета). **Выбор единиц измерения зависит от конкретного исследователя или от**

Глава 4. Статистика интервальных данных

В статистике интервальных данных элементы выборки — не числа, а интервалы. Это приводит к алгоритмам и выводам, принципиально отличающимся от классических. Настоящая глава посвящена основным идеям и подходам асимптотической статистики интервальных данных. Приведены результаты, связанные с основополагающими в рассматриваемой области прикладной математической статистики понятиями нотны и рационального объема выборки. Рассмотрен ряд задач оценивания характеристик и параметров распределения, проверки гипотез, регрессионного, кластерного и дискриминантного анализов.

4.1. Основные идеи статистики интервальных данных

Перспективная и быстро развивающаяся область статистических исследований последних лет — математическая статистика интервальных данных. Речь идет о развитии методов прикладной математической статистики в ситуации, когда статистические данные — не числа, а интервалы, порожденные, в частности, наложением ошибок измерения на значения случайных величин. Полученные результаты отражены в выступлениях на проведенной в журнале «Заводская лаборатория» дискуссии [1] и в докладах Международной конференции «Интервал–92» [2]. Приведем основные идеи весьма перспективного для вероятностно-статистических методов и

Основные книги профессора А.И. Орлова

1. *Устойчивость* в социально-экономических моделях. М.: Наука, 1979. 296 с.
2. *Задачи* оптимизации и нечеткие переменные. М.: Знание, 1980. 64 с.
3. *Анализ* нечисловой информации (препринт) (совместно с Ю.Н. Тюриным, Б.Г. Литваком, Г.А. Сатаровым, Д.С. Шмерлингом). М.: Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1981. 80 с.
4. *ГОСТ 11.011–83*. Прикладная статистика. Правила определения оценок и доверительных границ для параметров гамма-распределения (в соавторстве). М.: Изд-во стандартов, 1984. 53 с. Переиздание: М.: Изд-во стандартов, 1985. 50 с.
5. *Анализ* нечисловой информации в социологических исследованиях: Сб. ст.; Под ред. А.И. Орлова и Ю.Н. Толстой. М.: Наука, 1985. 220 с.
6. *Пакет* программ анализа данных «ППАНД»: Учеб. пособие (совместно с И.Л. Легостаевой, О.М. Черномордиком и др.). М.: Сотрудничающий центр Всемирной организации здравоохранения по профессиональной гигиене, 1990. 93 с.
7. *Математическое* моделирование процессов налогообложения (подходы к проблеме) (совместно с В. Г. Кольцовым, Н.Ю. Ивановой и др.). М.: Изд-во ЦЭО Министерства общего и профессионального образования РФ, 1997. 232 с.
8. *Эконометрика*: Учеб. для вузов. М.: Экзамен, 2002 (1-е изд.), 2003 (2-е изд.), 2004 (3-е изд.). 576 с.
9. *Управление* промышленной и экологической безопасностью: Учеб. пособие (совместно с В.Н. Федосеевым, В.Г. Ларионовым, А.Ф. Козьяковым). М.: УРАО, 2002 (1-е изд.), 2003 (2-е изд.). 220 с.

10. *Менеджмент* в техносфере: Учеб. пособие (совместно с В.Н. Федосеевым). М.: Академия, 2003. 384 с.

11. *Принятие* решений. Теория и методы разработки управленческих решений. М.: ИКЦ «МарТ»; Ростов н/Д: Издательский центр «МарТ», 2005. 496 с.

12. *Прикладная* статистика. М.: Экзамен, 2006. 671 с.

13. *Теория* принятия решений. М.: Экзамен, 2006. 574 с.

14. *Менеджмент* высоких технологий (совместно с А.А. Колотовым, И.Н. Омельченко). М.: Экзамен, 2008. 621 с.

Учебное издание

Орлов Александр Иванович

**ОРГАНИЗАЦИОННО-ЭКОНОМИЧЕСКОЕ
МОДЕЛИРОВАНИЕ**

Часть 1

НЕЧИСЛОВАЯ СТАТИСТИКА

Технический редактор *Э.А. Кулакова*
Художник *Н.Г. Столярова*
Корректор *О.В. Калашикова*
Компьютерная верстка *Н.Ф. Бердавцевой*

Оригинал-макет подготовлен
в Издательстве МГТУ им. Н.Э. Баумана.

Санитарно-эпидемиологическое заключение
№ 77.99.60.953.Д.003961.04.08 от 22.04.2008 г.

Подписано в печать 05.02.2009.
Формат 84×108/32. Усл. печ. л. 28,56. Тираж 500 экз.

Заказ №

Издательство МГТУ им. Н.Э. Баумана.
E-mail: press@bmstu.ru
105005, Москва, 2-я Бауманская ул., д. 5.

Отпечатано в ГУП ППП «Типография «Наука».
121099, Москва, Шубинский пер., 6.

ISBN 978-5-7038-3277-6

