А.А.БОРОВКОВ

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ОЦЕНКА ПАРАМЕТРОВ

ПРОВЕРКА ГИПОТЕЗ

А. А. БОРОВКОВ

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

•

ОЦЕНКА ПАРАМЕТРОВ ПРОВЕРКА ГИПОТЕЗ

Допущено Министерством высшего и среднего специального образования СССР в качестве учебного пособия для студентов математических и физических специальностей вузов

МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1984

22.172
Б 82
УДК 519.31

Боровков А. А. **Математическая статистика.**— Учебник.— М.: Наука. Главная редакция физико-математической литературы, 1984.— 472 с.

Книга написана на основании лекций, которые читались автором в течение многих лет на 3-м курсе математического факультета Новосибирского университета.

Книга охватывает широкий круг вопросов, включающий в себя основания современной математической статистики, предельные теоремы для эмпирических распределений и основных типов статистик, основы теории оценок и теории проверки гипотез. Излагаются, в частности, методы отыскания оптимальных и асимптотически оптимальных процедур.

Книга рассчитана на аспирантов и студентов математических и физических специальностей высших учебных заведений.

Илл. 5. Библ. 80 назв.

Александр Алексеевич Боровков

●

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Редакторы *А. А. Новиков, М. М. Горячая*
Технический редактор *Е. В. Морозова*
Корректор *Н. Б. Румянцева*

ИБ № 2239

Сдано в набор 20.04.83. Подписано к печати 03.01.84. Формат $60\times90^1/_{16}$. Бумага тип. №3. Обыкновенная гарнитура. Высокая печать. Условн. печ. л. 29,5. Условн. кр.-отт. 29,75. Уч.-изд. л. 31,17. Тираж 23 000 экз. Заказ № 594. Цена 1 р. 30 к.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

4 типография издательства «Наука».
630077, Новосибирск 77, Станиславского, 25

Б $\dfrac{1702060000-020}{053(02)-84}$ 72-83

ОГЛАВЛЕНИЕ

Глава 3

Теория проверки гипотез

ПРЕДИСЛОВИЕ

В основу этой книги положены лекции по математической статистике, которые автор в течение многих лет читал на 3-м курсе математического факультета Новосибирского университета. Со временем курс лекций неоднократно менялся в поисках варианта, который был бы по возможности более стройным, доступным и в то же время соответствовал бы современному состоянию этой науки. Были испробованы варианты, начиная с курса преимущественно рецептурного характера с изложением основных типов задач (построение оценок, критериев и изучение их свойств) и кончая курсом общего теоретико-игрового характера, в котором теория оценок и проверка гипотез представлялись как частные случаи единого подхода. Ограничения во времени (один семестр) не позволяли объединить эти дополняющие друг друга варианты, каждый из которых в отдельности обладал очевидными недостатками. В первом случае набор конкретных фактов мешал развитию общего взгляда на предмет. Второй вариант испытывал недостаток в простых конкретных результатах и оказывался загруженным многими новыми непростыми понятиями, трудными для усвоения. По-видимому, наиболее приемлемым является вариант, в котором изложение элементов теории оценок и теории проверки гипотез сочетается с последовательным проведением линии на отыскание оптимальных процедур.

Основу книги составляет объединение материала, читавшегося в разное время в лекциях, расширенное за счет разделов, присутствие которых диктуется самой логикой изложения. Главной целью являлось изложение современного состояния предмета в сочетании с максимально возможной доступностью и математической цельностью и стройностью.

Содержание книги распадается на 3 главы и Приложения.

В главе 1 изучаются свойства (в основном асимптотические) эмпирических распределений, лежащие в основе математической статистики.

В главах 2, 3 излагаются соответственно теория оценок и теория проверки статистических гипотез. Первые части каждой из этих глав посвящены описанию возможных подходов к решению поставленных задач и отысканию оптимальных процедур. Вторые части посвящены построению асимптотически оптимальных процедур.

Книга содержит также 7 Приложений. Они связаны с теми утверждениями в основном тексте, доказательство которых выходит за рамки основного изложения либо по своему характеру, либо по трудности.

Имеются еще библиографические замечания, не претендующие на полноту, но позволяющие проследить возникновение и развитие основных направлений математической статистики. При этом везде, где это возможно, оказывалось предпочтение ссылкам на монографии (как на более доступный вид литературы), а не на оригинальные статьи.

В настоящее время существует довольно много книг по математической статистике. Выделим из них следующие четыре, в которых представлено большое количество материала, отражающее современное состояние предмета,— это книги Г. Крамера [37], Э. Лемана [40], Ш. Закса [30] и И. А. Ибрагимова и Р. З. Хасьминского [31]. Из них наибольшее влияние на изложение настоящей книги оказали монографии [31] (некоторые идеи этой книги использовались в §§ 23—25, 27—29 гл. 2) и [40] (изложение §§ 5—8 гл. 3 близко по содержанию к соответствующим разделам [40]). Остальное изложение мало связано по строю с имеющимися книгами.

Наряду с известными результатами и подходами в настоящую книгу включены некоторые новые разделы, упрощающие изложение, ряд методологических улучшений, а также некоторые новые результаты и результаты, впервые публикуемые в монографической литературе.

Ниже приводится краткое описание методологической структуры книги (см. также оглавление и краткие предисловия к каждой из глав).

В §§ 1, 2 главы 1 вводятся понятия выборки, эмпирического распределения и устанавливается теорема Гливенко — Кантелли, которую можно рассматривать как фундаментальный факт, составляющий основу статистических выводов.

В § 3 вводится два типа статистик (статистики I и II типов), включающие в себя подавляющее большинство практически интересных статистик. Эти статистики определяются как значения $G(\mathbf{P}_n^*)$ функционалов G (удовлетворяющих некоторым условиям) от эмпирического распределения $\mathbf{P}_n^*$. Позже, в §§ 7, 8, устанавливаются предельные теоремы о распределении этих статистик. Это разгружает последующее изложение и освобождает от необходимости для каждой конкретной статистики проводить практически одни и те же рассуждения, не относящиеся при том к существу дела.

В § 5 собраны вспомогательные теоремы (названные в книге «теоремами непрерывности») о сходимости распределений и о сходимости их моментов. Это также разгружает последующее изложение.

В § 6 (он не обязателен при первом чтении) устанавливается, что эмпирическая функция распределения $F_n^*(t)$ есть условный пуассоновский процесс, и приводится формулировка теоремы (доказанной в приложении I) о сходимости процесса $\sqrt{n}(F_n^*(t) - F(t))$ к броуновскому мосту.

В § 10 вводятся сглаженные эмпирические распределения, позволяющие приближать не только само распределение, но и его плотность.

В § 3 главы 2, посвященной оценкам неизвестных параметров, вводится единый подход к построению оценок, названный «методом подстановки». Оп состоит в том, что оценку θ^* для параметра θ, представимого в виде функционала $\theta = G(\mathbf{P})$ от распределения $\mathbf{P}$ выборки, следует искать в виде $\theta^* = G(\mathbf{P}_n^*)$, где $\mathbf{P}_n^*$ — эмпирическое распределение. Все «разумные» оценки, используемые на практике, являются оценками подстановки. Оптимальность оценки достигается путем выбора подходящего функционала G. Если статистика $\theta^* = G(\mathbf{P}_n^*)$ является статистической I или II типов, то теоремы главы 1 позволяют немедленно установить состоятельность этих оценок и их асимптотическую нормальность. В §§ 4, 5 этот подход иллюстрируется на оценках, полученных с помощью метода моментов и метода минимального расстояния. С тех же позиций можно было бы рассматривать и оценки максимального правдоподобия (§ 6), но непосредственное их изучение позволяет получить более глубокие результаты, необходимые в дальнейшем.

Сравнение оценок в главе 2 происходит на основе двух подходов: *среднеквадратического* (сравниваются $\mathsf{M}_\theta(\theta^* - \theta)^2$) и *асимптотического* (сравниваются дисперсии предельного распределения $\sqrt{n}(\theta^* - \theta)$ в классе асимптотически нормальных оценок). В параметрическом случае это позволяет выделить 3 типа оптимальных оценок: эффективные оценки в классах K_b с фиксированным смещением b, байесовские и минимаксные оценки. На основе тех же принципов выделяются классы *асимптотически* оптимальных оценок при асимптотическом подходе. Для построения эффективных оценок используются следующие традиционные методы: первый носит качественный характер и связан с принципом достаточности (§§ 12—14); второй метод основан на количественных соотношениях, вытекающих из неравенства Рао — Крамера (§ 16); третий связан с соображениями инвариантности (§§ 17, 19), позволяющими сузить класс рассматриваемых оценок. Отысканию асимптотически оптимальных оценок, а также изучению асимптотических свойств функции правдоподобия посвящены §§ 20—30. Параграф 20 содержит интегральное неравенство типа Рао — Крамера, которое позволяет, в частности, получить простые критерии асимптотической байесовости и минимаксности оценок, а также обосновать выделение некоторого подкласса оценок $\tilde{K}_0$, которым следует ограничиться в поисках асимптотически эффективных оценок. Это позволяет путем изучения асимптотических свойств оценок максимального правдоподобия в § 25 сразу установить асимптотическую байесовость и минимаксность названных оценок и их асимптотическую эффективность в $\tilde{K}_0$. Параграфы 21—24 носят вспомогательный характер. Интервальное оценивание параметров рассматривается в §§ 31, 32 и в § 8 главы 3.

Глава 3 посвящена проверке гипотез. В §§ 1, 2 рассматривается случай конечного числа простых гипотез. Выделяются (аналогично теории оценивания) три типа оптимальных критериев — наиболее мощные в подклассах, байесовские и минимаксные. Устанавливаются связи между этими критериями и находится их явный вид. При этом в основу рассмотрений положен байесовский принцип (а не лемма Неймана — Пирсона), что, на наш взгляд, упрощает изложение и делает его более прозрачным. В § 3 излагаются асимптотические подходы к расчету критериев для проверки двух простых гипотез и производится их сравнение. В § 4 рассматривается общая постановка задачи о проверке двух сложных гипотез и определяются классы оптимальных критериев (равномерно наиболее мощные, байесовские, минимаксные). Параграф 5 посвящен отысканию равномерно наиболее мощных критериев в тех случаях, когда это возможно. В §§ 6, 7 решается та же задача, но в классах критериев, суженных на основании соображений несмещенности и инвариантности. При этом в основу рассмотрений, как и в §§ 1, 2, положен байесовский подход. В § 8 с помощью полученных результатов строятся наиболее точные доверительные множества. В § 9 рассматриваются байесовские и минимаксные критерии. Параграфы 10, 13 посвящены критерию отношения правдоподобия. Он оказывается равномерно наиболее мощным во многих частных случаях и обладает свойством асимптотической байесовости при весьма широких предположениях. Изучение свойств асимптотической оптимальности критерия отношения правдоподобия продолжено в §§ 15—17. В § 11 установлена оптимальность этого критерия в задачах последовательного анализа. Параграфы 14, 15 посвящены отысканию асимптотически оптимальных критериев для проверки близких гипотез, и найден их простой явный вид для основных статистических задач.

Существенной особенностью книги является тот факт, что в ней рассмотрены лишь статистические задачи, связанные с использованием *одной* выборки; задачи с двумя и более выборками, а также общий теоретико-игровой подход к проблемам статистики будут изложены в отдельной книге, которая будет являться прямым продолжением и дополнением настоящей.

Книга имеет многоцелевое назначение. В ее полном объеме она, конечно, ближе к программе кандидатского минимума по специальности «математическая статистика», чем к учебному пособию для студентов. Однако в книге предусмотрена система мер, упрощающих первое чтение, которая делает книгу доступной и для студентов. Параграфы повышенной трудности или «более продвинутые» по своему содержанию отмечены звездочкой, их следует при первом чтении опускать, как и текст, набранный петитом. Кроме того, изложение технически более сложного случая, относящегося к многомерному параметру, почти везде выделено в отдельные разделы и параграфы, которые также можно пропускать.

Преподаватели ВУЗов, уже знакомые хотя бы частично с предметом, могут выбирать из книги совокупность параграфов (вариантов может быть много), используя которые (не обязательно полностью) можно составить семестровый курс математической статистики. Вот один из вариантов: §§ 1, 3, 5 главы 1; §§ 2—4, 6—12, 14, 16 (21, 23—25), 31, 32 главы 2; §§ 1, 2, 4, 5, 12 (13, 16) главы 3. Параграфы, взятые в скобки, посвящены асимптотически оптимальным процедурам. Их следует в зависимости от подготовленности студентов либо максимально облегчить, либо опустить вовсе.

Чтение книги предполагает знакомство с курсом теории вероятностей в объеме учебника А. А. Боровкова [11]. Ссылки на эту книгу, в отличие от других ссылок, появляются в тех местах, которые предполагаются читателю известными, и служат, главным образом, для напоминаний.

Нумерация параграфов в книге в каждой главе самостоятельная, как и нумерация теорем (лемм, примеров и т. д.) в каждом параграфе. Для удобства чтения используются разные системы для ссылок на теоремы, леммы, примеры, формулы и т. д., в зависимости от того, как далеко от читаемого места они находятся. Если делается ссылка на теорему 1 или формулу (12) читаемого параграфа, то ссылки на них будут выглядеть так: теорема 1, формула (12). Если делается ссылка на теорему 1 и формулу (12) одного из предыдущих параграфов этой главы (например, § 13), то ссылки будут иметь такой вид: теорема 13.1, формула (13.12). Наконец, если ссылки относятся к другой главе, то будет появляться еще указатель номера главы (первая цифра). Например, теорема 2.13.1 означает теорему 1 § 13 главы 2, а формула (2.13.12) — формулу (12) § 13 главы 2. То же самое относится к обозначению параграфов. Ссылка на § 13 означает ссылку на § 13 этой главы, а ссылка на § 2.13 означает ссылку на § 13 главы 2.

Значок $\lhd$ означает окончание доказательства.

Для удобства читателя в конце книги помещены список основных обозначений и предметный указатель.

Написание этой книги было весьма трудным многоэтапным делом. Значительную помощь при подготовке первоначальных записок лекций к печати и по устранению недостатков в них мне оказал И. С. Борисов. Второй вариант рукописи прочел по моей просьбе К. А. Боровков, который представил мне длинный список полезных советов и замеченных им погрешностей. Он же существенно помог мне при «чистке» окончательного варианта рукописи. В поисках новой критики я обратился с просьбой ознакомиться с рукописью к А. И. Саханенко. Он также представил мне длинный список замечаний и предложений по улучшению изложения, многими из которых я воспользовался. Наиболее существенные изменения при этом коснулись доказательств в §§ 16, 21, 23, 29 главы 2, §§ 13—15 главы 3, Приложениях II, V (см. также библиографические замечания).

Много ценных замечаний, направленных на улучшение книги, было сделано Д. М. Чибисовым. Рукопись просматривали В. В. Юринский и А. А. Новиков и также сделали ряд полезных замечаний. Всем названным коллегам и всем тем, кто так или иначе помогал мне при работе над книгой, я хотел бы выразить здесь свою глубокую и искреннюю признательность и благодарность за их труд и содействие в написании книги.

А. А. Боровков

Сентябрь 1982 года

ВВЕДЕНИЕ

Настоящая книга посвящена изложению основ раздела математики, который называется *математическая статистика*. Последнюю для краткости называют иногда просто *статистикой*. Однако следует иметь в виду, что такое сокращение возможно лишь при достаточно хорошем взаимопонимании, поскольку сам по себе термин «статистика» соответствует обычно несколько иному понятию.

Что представляет из себя предмет математической статистики? Можно приводить разные описательные «определения», которые в большей или меньшей степени отражают содержание этого раздела математики. Одно из самых простых и грубых основано на сравнении, связанном с понятием выборки из генеральной совокупности и с задачей о гипергеометрическом распределении, которая обычно рассматривается в начале курса теории вероятностей. Зная состав генеральной совокупности, там изучают распределения для состава случайной выборки. Это типичная *прямая задача* теории вероятностей. Однако часто приходится решать и *обратные задачи*, когда известен состав выборки и по нему требуется определить, какой была генеральная совокупность. Такого рода обратные задачи и составляют, образно говоря, предмет математической статистики.

Несколько уточняя это сравнение, можно сказать так: в теории вероятностей мы, зная природу некоторого явления, выясняем, как будут вести себя (как распределены) те или иные изучаемые нами характеристики, которые можно наблюдать в экспериментах. В математической статистике наоборот — исходными являются экспериментальные данные (как правило, это наблюдения над случайными величинами), а требуется вынести то или иное суждение или решение о природе рассматриваемого явления. Таким образом, мы соприкасаемся здесь с одной из важнейших сторон человеческой деятельности — процессом познания. Тезис о том, что «критерий истины есть практика» имеет самое непосредственное отношение к математической статистике, поскольку именно эта наука изучает методы (в рамках точных математических моделей), которые позволяют отвечать на вопрос, соответствуют ли практика, представленная в виде результатов эксперимента, данному гипотетическому представлению о природе явления или нет.

При этом необходимо подчеркнуть, что, как и в теории вероятностей, нас будут интересовать не те эксперименты, которые позволяют делать однозначные, детерминированные выводы о рассматриваемых в природе явлениях, а эксперименты, результатами которых являются случайные события. С развитием науки роль такого рода задач становится все больше и больше, поскольку с увеличением точности экспериментов становится все труднее избежать «случайного фактора», связанного с разного рода помехами и ограниченностью наших измерительных и вычислительных возможностей.

Математическая статистика является частью теории вероятностей в том смысле, что каждая задача математической статистики есть по существу задача (иногда весьма своеобразная) теории вероятностей. Однако сама по себе математическая статистика занимает и самостоятельное положение в табеле о науках. Математическая статистика может рассматриваться как наука о так называемом индуктивном поведении человека (и не только человека), в условиях, когда он должен на основании своего недерминированного опыта принимать решения с наименьшими для него потерями*).

Математическую статистику называют также теорией статистических решений, поскольку ее можно характеризовать как науку об оптимальных решениях (последующие два слова требуют пояснения), основанных на статистических (экспериментальных) данных. Точные постановки задач будут даны позже, в основном тексте книги. Здесь же мы ограничимся тем, что приведем три примера наиболее простых и типичных статистических задач.

Пример 1. Для многих изделий одним из основных параметров, которым характеризуется качество, является срок службы. Однако срок службы изделия (скажем, электролампы), как правило, случаен, и заранее определить его невозможно. Опыт показывает, что если процесс производства в известном смысле однороден, то сроки службы ξ_1, ξ_2, ... соответственно 1-го, 2-го и т. д. изделий можно рассматривать как независимые одинаково распределенные величины. Интересующий нас параметр, определяющий срок службы, естественно отождествить с числом $\theta = \mathsf{M}\xi_i$. Одна из стандартных задач состоит в выяснении, чему равно θ. Для того чтобы определить это значение, берут n готовых изделий и проверяют их. Пусть x_1, x_2, ..., x_n — сроки службы этих проверенных изделий. Мы знаем, что

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \xrightarrow[\text{п.н.}]{} \theta$$

*) Подробнее об этом см. [46].

при $n \to \infty$. Поэтому естественно ожидать, что число $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ при достаточно большом n окажется близким к θ и позволит в какой-то мере ответить на поставленные вопросы. При этом очевидно, что мы заинтересованы в том, чтобы требуемое число наблюдений n было по возможности наименьшим, а наша оценка числа θ по возможности более точной (завышение параметра θ, как и его занижение, приведут к материальным потерям).

Пример 2. Радиолокационное устройство в моменты времени $t_1, t_2, \ldots, t_n$ зондирует заданную часть воздушного пространства с целью обнаружения там некоторого объекта. Обозначим $x_1, \ldots, x_n$ значения отраженных сигналов, принятых устройством. Если в заданной части пространства интересующий нас объект отсутствует, то значения x_i можно рассматривать как независимые случайные величины, распределенные так же, как некоторая случайная величина ξ, природа которой обусловлена характером различных помех. Если же в течение всего периода наблюдений объект находился в поле зрения, то x_i будут наряду с помехами содержать «полезный» сигнал a, и значения x_i будут распределены как $\xi + a$. Таким образом, если в первом случае наблюдения x_i имели функцию распределения $F(x)$, то во втором случае их функция распределения будет иметь вид $F(x-a)$. По выборке $x_1, \ldots, x_n$ требуется решить, какой из этих двух случаев имеет место, т. е. существует в заданном месте интересующий нас объект или нет.

В этой задаче окажется возможным указать в известном смысле «оптимальное решающее правило», которое будет решать поставленную задачу с минимальными ошибками. Сформулированная задача может быть усложнена следующим образом. Сначала объект отсутствует, а затем, начиная с наблюдения с неизвестным номером θ, появляется. Требуется по возможности более точно определить момент θ появления объекта. Это так называемая «задача о разладке», имеющая и целый ряд других интерпретаций, важных для приложений.

Пример 3. Некоторый эксперимент производится сначала n_1 раз в условиях А и затем n_2 раз в условиях В. Обозначим $x_1, \ldots, x_{n_1}$ и $y_1, \ldots, y_{n_2}$ результаты этих экспериментов соответственно в условиях А и В. Спрашивается, сказывается ли изменение условий эксперимента на его результатах? Иными словами, если обозначить через $\mathbf{P}_A$ распределение x_i, $1 \leqslant i \leqslant n_1$, и через $\mathbf{P}_B$ — распределение y_i, $1 \leqslant i \leqslant n_2$, то вопрос состоит в том, выполнено соотношение $\mathbf{P}_A = \mathbf{P}_B$ или нет.

Например, если нужно установить, влияет ли некоторый препарат на развитие, скажем, растений или животных, то параллельно ставятся две серии экспериментов (с препаратом и без), результаты которых необходимо уметь сравнивать.

Часто возникают и более сложные задачи, когда аналогичный вопрос ставится для многих серий наблюдений, проведенных в различных условиях. Если результаты наблюдений зависят от условий, то бывает необходимым проверить тот или иной характер этой зависимости (так называемая задача о регрессии).

Пример 3 и названные более сложные проблемы относятся к классу статистических задач *с двумя и более выборками*. Такие задачи будут рассмотрены в отдельной книге (см. Предисловие).

Список примеров типичных статистических задач, разных по сложности и по своему существу, можно было бы продолжить. Однако общими для всех них будут следующие два обстоятельства:

1. Перед нами не было бы никаких проблем, если бы распределения результатов наблюдений, которые фигурируют в задачах, были нам известны.

2. В каждой из этих задач мы должны по результатам экспериментов принимать какое-то решение относительно распределения имеющихся наблюдений (отсюда и название «Теория статистических решений», упоминавшееся выше).

В связи с этими двумя замечаниями принципиальное значение для всего дальнейшего и, в частности, для решения приведенных в качестве примеров задач, приобретает следующий факт. Оказывается, по результатам наблюдений $x_1, \ldots, x_n$ над некоторой случайной величиной ξ можно при больших n сколь угодно точно восстановить неизвестное распределение $\mathbf{P}$ этой случайной величины. Аналогичное утверждение справедливо и для любого функционала $\theta = \theta(\mathbf{P})$ от этого неизвестного распределения.

Этот факт лежит в основе математической статистики. Ему и более точным постановкам задач будет посвящена глава 1.