

ФМ

517.8
п-89
614225

Е. И. ПУСТЫЛЬНИК

Статистические
методы
анализа
и обработки
наблюдений

БИ

Физико-
Математическая
Библиотека
Инженера

Е. И. ПУСТЫЛЬНИК

614225
СТАТИСТИЧЕСКИЕ
МЕТОДЫ АНАЛИЗА
И ОБРАБОТКИ
НАБЛЮДЕНИЙ



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1968

ВОЛОГОДСКАЯ
областная библиотека

517.8

П 89

УДК 519.240

**Статистические методы анализа и обработки наблюдений,
Пустыльник Е. И.**

Книга посвящена изложению некоторых простых, но важных методов обработки числовых результатов наблюдений. Эти методы стали в последнее время обязательным этапом каждого исследования.

Настоящая книга является попыткой изложения большого числа классических методов анализа и обработки наблюдений на минимальной аналитической основе. В ней излагаются важнейшие оценки нормально распределенной случайной величины, сравнение и анализ серий наблюдений. Приводятся элементарные доказательства основных формул дисперсионного, корреляционного и регрессионного анализа, некоторые принципы планирования эксперимента и последовательного анализа. В начале книги излагается необходимый минимум сведений из теории вероятностей и общей теории наблюдений. В книге илл. 30, табл. 33.

Евгений Изиевич Пустыльник

**СТАТИСТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА
И ОБРАБОТКИ НАБЛЮДЕНИЙ**

М., 1968 г., 288 стр. с илл.

Редактор *И. М. Овчинникова*.

Техн. редактор *С. Я. Шкляр*. Корректоры *Л. Н. Боровина, А. Ф. Серкина*

Сдано в набор 20/X 1967 г. Подписано к печати 27/II 1968 г. Бумага 84×108^{1/32}. Тип. № 1. Физ. печ. л. 9. Условн. печ. л. 15,12. Уч.-изд. л. 14,42. Тираж 28 000 экз. Т-00079. Цена книги 1 р. 09 к. Заказ № 2138.

Издательство «Наука».

Главная редакция физико-математической литературы
Москва, В-71, Ленинский проспект, 15.

Ордена Трудового Красного Знамени

Первая Образцовая типография имени А. А. Жданова
Главполиграфпрома Комитета по печати при Совете Министров СССР
Москва, Ж-54, Валовая, 28.

2-2-3

Р 4-68

СОДЕРЖАНИЕ

Предисловие	5
Введение	7
§ 1. Случайные события	13
1.1. Понятие случайного события (13). 1.2. Вероятность случайного события (16). 1.3. Классическое определение вероятности (19). 1.4. Геометрическое определение вероятности (26). 1.5. Последовательность независимых испытаний (30).	
§ 2. Случайные величины.	33
2.1. Дискретные и непрерывные случайные величины (33). 2.2. Распределение случайной величины (35). 2.3. Числовые характеристики случайных величин (43).	
§ 3. Нормальное распределение.	52
3.1. Пример нормального распределения (задача о рассеянии снарядов) (52). 3.2. Свойства нормального распределения (58). 3.3. Использование нормального распределения для изучения других случайных величин (67).	
§ 4. Наблюдения.	76
4.1. Наблюдение как этап исследования (76). 4.2. Наблюдение как случайная величина (81). 4.3. Основная схема производства наблюдений (выборочный метод) (85). 4.4. Среднее и дисперсия выборки (89).	
§ 5. Основные задачи математической статистики	99
5.1. Параметры распределения (99). 5.2. Доверительные интервалы и доверительные вероятности (106). 5.3. Проверка статистических гипотез (111). 5.4. Односторонние и двусторонние критерии (116).	
§ 6. Оценка результатов наблюдений над нормально распределенной случайной величиной.	120
6.1. Оценка генерального среднего (120). 6.2. Оценка генеральной дисперсии (126). 6.3. Сравнение дисперсий (131). 6.4. Сравнение средних (138). 6.5. Проверка однородности наблюдений (146).	

§ 7. Анализ распределения наблюдений.	152
7.1. Проверка основной гипотезы (152). 7.2. Неравенство Чебышева и его использование (162). 7.3. Подбор плотности теоретического распределения (165). 7.4. Оценка вероятности случайного события (172). 7.5. Использование оценок вероятности для анализа распределения (179).	
§ 8. Дисперсионный анализ.	186
8.1 Постановка задачи (186). 8.2. Однофакторный дисперсионный анализ (189). 8.3. Двухфакторный дисперсионный анализ (196). 8.4. Планирование эксперимента при дисперсионном анализе (206).	
§ 9. Зависимость между случайными величинами	213
9.1. Корреляция (213). 9.2. Регрессия (221). 9.3. Вычисление и анализ приближенной регрессии (226). 9.4. Линейная регрессия (233). 9.5. Нелинейная регрессия (239). 9.6. Параболическая регрессия (243).	
§ 10. Некоторые вопросы планирования эксперимента.	252
10.1. Способы отбора (252). 10.2. Выбор числа наблюдений (257). 10.3. Последовательный анализ (263).	
Рекомендуемая литература	268
П р и л о ж е н и е. Математические таблицы	270

ПРЕДИСЛОВИЕ

За последнее время статистические методы, проникнув в самые отдаленные разделы техники, стали одним из основных приемов анализа и обработки опытных данных. Однако сила и могущество уже разработанных наукой статистических методов исследования все чаще приходят в противоречие с недостатком соответствующего образования у инженеров, которым эти методы приходится применять. Основы математической статистики самой жизнью превращаются в необходимейший элемент математического образования студентов технических вузов.

Настоящая книга представляет собой обработанную запись лекций, которые автор в течение ряда лет читал для студентов и научных сотрудников Воронежского технологического института. Основы математической статистики излагаются в ней непосредственно в применении к задачам обработки наблюдений, что и обусловило соответствующий отбор материала.

Неотъемлемой частью книг по математической статистике является обычно обширный раздел, посвященный теории вероятностей. Теория вероятностей, несомненно, обладает большой самостоятельной ценностью и прикладными возможностями. Тем не менее, изучение этой теории требует серьезной математической подготовки и, как правило, усложняет для рядовых инженерных работников путь к овладению основами математической статистики. Поэтому в данной книге теоретико-вероятностная часть по возможности сокращена до минимума.

Автор ставил задачей избежать и другой крайности некоторых статистических руководств — чрезмерной рецептурности изложения, отсутствия сколько-нибудь глубоких объяснений применяемых методов. В книге последовательно

применяется эвристический метод изложения, некоторые пункты специально посвящены общему анализу рассматриваемых в дальнейшем задач.

Желание сделать книгу доступной для широкого круга лиц привело к необходимости жертвовать строгостью математического изложения: громоздкие доказательства чаще всего заменены простыми, но не очень строгими объяснениями. Для чтения книги требуются лишь начальные сведения из математического анализа. Однако в самой книге изложение постоянно опирается на предыдущий материал, в связи с чем ее рекомендуется изучать в строгой последовательности, ничего не опуская (исключение составляет разве лишь п. 3.1, который при первом чтении можно опустить). Только ознакомившись со всей книгой целиком, ее можно в дальнейшем использовать как справочное пособие.

В Приложении к книге даются статистические таблицы, что избавляет читателя от необходимости привлекать другую литературу при анализе и обработке наблюдений по рецептам данной книги. Этой же цели служат помещенные в Приложении таблицы десятичных логарифмов, квадратов, кубов и квадратных корней действительных чисел.

Книга предназначена в первую очередь для студентов вузов и инженеров, интересующихся более применением статистических методов, нежели их строгим математическим обоснованием. Вместе с тем, она, по-видимому, может быть полезна в качестве введения и для лиц, желающих заняться математической статистикой более углубленно. Именно последним адресован приведенный в конце книги список рекомендуемой литературы с краткими аннотациями. Этот список не претендует не только на полноту, но и на охват хотя бы основной литературы по математической статистике; в нем отразились главным образом вкусы и интересы самого автора.

Автор будет весьма признателен читателям за критические замечания, направленные на улучшение книги.

ВВЕДЕНИЕ

Все процессы, происходящие в природе, являются результатом взаимодействия многих факторов. Для того чтобы изучить эти процессы и в дальнейшем ими управлять, необходимо выяснить, какую роль в рассматриваемом процессе играет каждый фактор в отдельности. Так, например, изучая движение тела, необходимо выяснить, какие силы приводят его в движение, какие тормозят; наконец, каким образом само движущееся тело влияет на эти силы. Все эти факторы необходимо выразить в каких-то количественных оценках — после этого на помощь исследователю приходит могущество математических методов.

Таким образом, математические методы изучения взаимодействующих факторов требуют умения выражать действие различных факторов количественно. Чтобы получить необходимые числовые данные, нужно произвести серию наблюдений. Наблюдение — это решающее звено всякого эксперимента, всякого исследования.

Однако даже самый тщательно подготовленный эксперимент не позволяет выделить интересующий нас фактор в чистом виде. Мы не в силах изолировать многие посторонние факторы: например, изучая падающие тела, мы не можем избежать действия вращения Земли; изучая химические реакции, мы никогда не имеем дела с чистыми веществами; изучая электронные процессы, не можем вести их в абсолютном вакууме и т. д. Наконец, нужно вспомнить о различных помехах, связанных с окружающей обстановкой — ведь даже шум идущего по улице автомобиля сказывается на проводимом в лаборатории эксперименте.

Следовательно, каждое наблюдение дает нам лишь результат взаимодействия основного изучаемого фактора с многочисленными посторонними. Некоторые из этих факторов (например, вращение Земли) можно учесть, так как

они сами по себе достаточно хорошо изучены. Учет других факторов (например, наличие примесей в веществах) может быть очень громоздким. Он сильно затянет эксперимент, сделает его неоправданно дорогим. Наконец, многие факторы (помехи) бывают настолько неожиданными, что их вообще нельзя учесть. Сюда же нужно отнести и те факторы, о которых на данном этапе развития науки вообще ничего не известно.

Из сказанного выше можно сделать только один вывод: полное и точное описание какого-либо процесса возможно лишь в том случае, если известны все факторы, влияющие на этот процесс. Иными словами, такое описание вообще невозможно.

К счастью, оно и не нужно.

Большинство измеряемых на практике величин обладает свойством непрерывности, т. е. их значения сплошь заполняют некоторый числовой промежуток. Однако все применяемые при этом измерительные приборы обладают некоторым пределом точности (разрешающей способностью) — минимальной разницей в значениях двух величин, которую они в состоянии обнаружить. Этот предел обычно указывается на приборах, изготовленных в заводских условиях. Например, аналитические весы, взвешивающие с точностью до 0,1 мг, не смогут различить такие веса, как 12,52 и 12,54 мг, и в обоих случаях покажут 12,5 мг. В результате все дальнейшие вычисления, связанные с этими данными, также будут содержать некоторую неточность, даже если пользоваться абсолютно точными и полными формулами, описывающими исследуемый процесс.

С другой стороны, нужно учесть, что полученные данные не всегда удастся полностью использовать в дальнейшем — приходится округлять их, теряя добытую с таким трудом драгоценную точность. В этом отношении можно привести интересный пример. Некоторые математики XVIII—XIX веков увлекались вычислением числа π с высокой точностью. Математик Шенкс в шестидесятых годах прошлого столетия вычислил π с точностью до 707-го знака после запятой, потратив на это всю свою жизнь*). Однако подобная точность

*) В настоящее время с помощью электронных вычислительных машин найдено более двух тысяч знаков числа π .

еще нигде не была использована. Так, например, чтобы вычислить с точностью до микрона длину окружности с центром на Земле и радиусом до ближайшей звезды, т. е. $R=4,5$ световых лет, достаточно иметь число π с 25 знаками после запятой — даже в таком, заведомо бессмысленном вычислении 682 шенковских знака остаются лишними!

Любое увеличение точности при измерениях сильно усложняет эксперимент. Кроме того, добавление каждого лишнего знака усложняет вычисления на 10—15%. Поэтому нужно всегда хорошо знать ту точность, которая требуется от результата, не стремясь к излишней точности измерений и вычислений.

Итак, в наших наблюдениях всегда допускается некоторая «законная» неточность, величину которой можно рассчитать заранее. Благодаря этому мы можем не учитывать те посторонние факторы, действие которых намного меньше этой неточности; например, изучая движение тел на Земле, можно не учитывать силы тяготения между этими телами или кривизну Земли при малых перемещениях и т. д.

Однако и здесь возникают свои трудности. Рассмотрим для примера такой вопрос: нужно ли, изучая движение автомобиля, учитывать тепловое колебание молекул, из которых он состоит. Ответ напрашивается отрицательный. Но давайте внимательней присмотримся к движению молекул внутри твердого тела. Молекулы колеблются вокруг положения равновесия с достаточно большими скоростями, однако движение это хаотическое, так что молекулы, движущиеся в одну сторону, уравниваются молекулами, движущимися в противоположную сторону. При этом можно представить такое случайное стечение обстоятельств, когда большинство молекул двинется в одну сторону, не уравниваясь с другой стороны. Разумеется, в результате в эту же сторону резко переместится и сам автомобиль. Такая возможность вполне допустима теоретически. Как же быть: учитывать ее или нет?

Но ведь если такое стечение обстоятельств и возможно, то оно необычайно редко! — ответит читатель. Во всяком случае, оно ни разу не встречалось за всю автомобильную практику человечества.

Правильно! И в результате мы приходим к еще одному важному выводу: даже сильные отклонения не нужно

учитывать, если они достаточно редки. При этом, правда, мы рискуем однажды попасть именно на то самое несчастное стечение обстоятельств, однако риск здесь может быть не велик, а облегчение исследований будет достаточно большое.

По этой же причине удастся избежать детального исследования многочисленных непредвиденных (случайных) помех. Хотя действие каждой из них может оказаться вполне заметным, в общей массе они, как правило, уравнивают друг друга, лишь изредка давая заметный суммарный эффект.

Естественно, пренебрегать можно лишь теми отклонениями, которые действительно редки — в противном случае риск будет слишком велик, равносильен беспечности. Значит, эту меру риска надо оценивать, устанавливая допустимый предел для каждого конкретных обстоятельств. Иными словами, нужно научиться численно характеризовать, насколько редко то или иное отклонение.

Для того чтобы выяснить, является ли какое-то событие редким и насколько, необходимо провести очень большое число наблюдений, ибо нужно иметь возможность сравнивать это событие с другими. Частота или редкость познаются только в сравнении. Это можно подтвердить следующей юмористической историей: некий путешествующий англичанин дважды проезжал Париж и оба раза попадал в дождь; после этого он записал в дневнике: «Париж — ужасный город, здесь всегда идет дождь!».

Чем больше проведено наблюдений, тем лучше можно оценить редкость интересующего нас события. Но и этот процесс не может продолжаться бесконечно. Следовательно, при определении редкости события мы опять вынуждены идти на риск (например, упомянутый выше англичанин мог прожить в Париже целый месяц, а этот месяц мог оказаться небывало дождливым для Парижа). Эту долю риска опять нужно оценивать и т. д.

Подобные рассуждения быстро завели бы нас в тупик, если бы каждое случайное событие нужно было изучать заново. Оказывается, однако, что случайные, непредвиденные события в массе своей подчиняются некоторым общим неслучайным закономерностям.

Наука, изучающая закономерности массовых случайных событий, называется *теорией вероятностей*. Приме-

нение теории вероятностей к обработке больших совокупностей чисел называется *математической статистикой*.

Использование методов математической статистики в обработке наблюдений оказывается весьма плодотворным. Закономерности отклонений при наблюдениях изучены достаточно хорошо, составлены многочисленные таблицы. Это позволяет значительно сокращать объем наблюдений.

Но случайность остается случайностью, и никакие теории при наличии непредвиденных и случайных факторов не могут давать точные и однозначные ответы. Основная задача математической статистики при обработке наблюдений — оценить риск той или иной ошибки в полученном результате. Принять или не принять этот риск — дело исследователя. В том случае, если этот риск его не устраивает, он должен найти пути его уменьшения: применить более точную методику наблюдений, устранить наиболее заметные помехи и т. д. Хороший эффект дает увеличение числа параллельных наблюдений, так как при этом появляется больше шансов, что в среднем случайные помехи взаимно уничтожатся.

Статистические методы являются могучим оружием в руках умелого исследователя, однако они ни в коем случае не должны становиться самоцелью. В некоторых случаях результат можно оценить и не прибегая к детальному статистическому анализу. Кроме того, нужно помнить, что математика может дать исчерпывающий ответ только при исчерпывающих данных, ибо всегда, чем меньше исходных данных использует теория, тем грубее получаемый с ее помощью результат. Теория вероятностей отвлекается от всех физических особенностей изучаемого явления, кроме его случайности. Поэтому применение теории вероятностей к обработке наблюдений приводит нередко к очень грубым оценкам. Это отнюдь нельзя считать недостатком теории — причина здесь либо в недостаточном объеме исследуемых данных, либо в сильной нестабильности опыта. Значит, от результатов статистической обработки нельзя просто отмахнуться, если они неприемлемы — нужно обязательно найти пути к их улучшению.

Отметим еще одну особенность статистических методов обработки наблюдений. Применение этих методов почти всегда связано с большим объемом вычислений. В связи

с этим нужно стараться использовать любую возможность для облегчения счета: применять вычислительную технику (арифмометры, электрические счетные машины и т. п.), использовать правила приближенных вычислений, сокращенные формулы.

Каждому исследователю, применяющему статистические методы обработки наблюдений, нужно помнить следующее:

а) не допускать излишнего количества цифр в промежуточных вычислениях;

б) если найденная при вычислениях величина будет сравниваться с табличной, то число цифр в ней не должно превосходить числа цифр в табличной величине — в связи с этим можно округлять даже исходные данные;

в) везде, где это возможно, нужно избавляться от дробей, изменяя масштабы отсчета, переносить начало отсчета и т. д.;

г) все находимые при расчетах величины нужно проверять либо с помощью специальных приемов, либо проводя вместе с кем-нибудь параллельные вычисления;

д) не применять излишне тонкие методы анализа там, где результат ясен из грубых оценок.

§ 1. СЛУЧАЙНЫЕ СОБЫТИЯ

1.1. Понятие случайного события. Основу изучения различных процессов, происходящих в природе, составляет выяснение всевозможных причинно-следственных связей между отдельными явлениями путем эксперимента. Осуществив по своему желанию одно или несколько первоначальных явлений-причин (в дальнейшем они называются *факторами*), экспериментатор получает возможность изучать различные появляющиеся при этом явления-следствия.

Иногда (очень редко) в процессе эксперимента удается сделать случайное открытие, т. е. обнаружить явление-следствие, о котором ранее ничего не было известно. Но, как правило, экспериментатор заранее намечает совокупность тех явлений-следствий, появления которых он ожидает и изучению которых посвящает свой эксперимент. Например, ожидая появления электрического тока в изучаемой системе, экспериментатор должен предусмотреть наличие приборов, регистрирующих и измеряющих этот ток.

Изучаемые явления-следствия могут интересовать экспериментатора с самых различных сторон. Так, если исследуется вещество, полученное в результате химических реакций, то могут быть отмечены такие его характеристики, как цвет, вкус, вес, объем и многое другое. Однако самое сложное явление всегда можно мысленно разбить на такие мелкие частные явления, относительно которых остается выяснить только одно: произошли они или не произошли. Например, рассмотрев в качестве отдельных явлений всевозможные цвета спектра, мы в дальнейшем, зная о каждом цвете, осуществился он или нет, немедленно можем указать совокупный цвет полученного в результате химической реакции вещества. Точно так же, измеряя вес вещества, мы в качестве отдельных частных явлений можем рассматривать всевозможные априорные значения этого веса.

Частные явления указанного типа могут иметь различную природу, но каждое интересует нас только с одной точки зрения: произошло оно или не произошло, осуществилось или не осуществилось. Явления, рассматриваемые только с той точки зрения, осуществились они или не осуществились, называются *событиями*. Применительно к событиям ставится следующая основная задача: предсказать, появится ли изучаемое событие при осуществлении некоторого наперед заданного комплекса факторов (явлений-причин).

Чтобы решить эту задачу, нужно, очевидно, знать влияние каждого фактора на интересующее нас событие. Например, составляя электрическую цепь из некоторых заданных элементов (сопротивления, приборы, провода и т. п.) и интересуясь, будет ли в этой цепи ток, мы о каждом заданном элементе (факторе) должны знать его отношение к электрическому току (событию) — проводник это или диэлектрик или, возможно, источник тока и т. д.

Событие, которое при заданном комплексе факторов обязательно произойдет, называется *достоверным* событием. Например, в замкнутой цепи, состоящей из хорошо соединенных проводников и исправного источника тока, появление электрического тока есть событие достоверное.

Событие, которое не может осуществиться при заданном комплексе факторов, называется *невозможным* событием. Так, невозможным событием является электрический ток в разомкнутой цепи (при отсутствии проводимости через воздух).

Суждение о достоверности или невозможности некоторого события является категорическим суждением — именно такие суждения принято считать окончательным результатом исследования. Отсюда возникает интерес к обратной задаче: указать такие комплексы факторов, при которых о заданном событии можно сделать категорические суждения (достоверность, невозможность).

Каждое событие является результатом действия большого числа факторов. Все их нужно знать для того, чтобы суждение о событии стало категорическим. Если же заданная совокупность факторов по отношению к событию неполна, то и категорическое суждение о событии становится невозможным. Получается следующая ситуация: заданные факторы благоприятствуют событию и, значит, оно может

произойти; с другой стороны, этих факторов недостаточно, чтобы гарантировать событие, и, значит, оно может и не произойти.

Событие, которое при заданном комплексе факторов может либо произойти, либо не произойти, называется *случайным событием* *).

С примерами случайных событий мы встречаемся на каждом шагу. Какой номер автобуса раньше подойдет к остановке, на которой мы ожидаем; какая будет завтра погода; какой стороной упадет подброшенная вверх монета — везде, где отсутствует полная информация, появляется случайность.

Понятия достоверного, невозможного и случайного события являются относительными, они связаны с заданным комплексом факторов. Достаточно этот комплекс изменить, как сразу может измениться характер события. Например, для замкнутой проводящей электрической цепи с исправным источником тока наличие электрического тока в цепи есть явление достоверное. Но достаточно лишить нас сведений хотя бы об одном элементе этой цепи, и наличие тока в сложившихся условиях станет событием случайным.

Таким образом, *случайность события связана с наличием факторов, влияющих на это событие, но не вошедших в заданный комплекс*. Эти факторы по отношению к заданному комплексу называются *случайными*. Случайным будет любой фактор, не вошедший в заданный комплекс факторов, даже если он хорошо изучен. Такое определение, разумеется, является только формальным — оно удобно для построения теории случайных событий. На практике чаще всего нет нужды объявлять случайными хорошо изученные факторы. В качестве случайных здесь рассматриваются обычно факторы, которые по тем или иным причинам невозможно (либо очень трудно) учесть. Эти факторы чаще всего не имеет смысла отделять друг от друга, поскольку связанные с ними причинно-следственные связи все равно не учитываются. Поэтому в дальнейшем мы, как правило, будем говорить об одном объединенном *факторе случайности*.

*) В дальнейшем там, где это не может вызвать недоразумений, вместо слов «случайное событие» употребляется просто термин «событие».

Случайность события никоим образом не связана с личными качествами исследователя — его способностями учитывать или предсказывать явления. Эта случайность связана с совершенно объективным фактом сужения комплекса факторов, обеспечивающих (или делающих невозможным) рассматриваемое событие.

Факторы, входящие в заданный комплекс, мы будем называть неслучайными или *основными*. Влияние таких факторов на исследуемое событие должно быть строго определено и неизменно — лишь тогда их можно включать в заданный комплекс.

1.2. Вероятность случайного события. Для того чтобы выяснить, произойдет или не произойдет некоторое событие при заданном комплексе основных факторов, нужно прежде всего осуществить этот комплекс. Каждое такое осуществление принято называть *испытанием*. Испытанием является, в частности, любой эксперимент, в результате которого производятся наблюдения. Ожидание автобуса, подбрасывание монеты в приводившихся примерах — тоже испытания.

Предсказать результат единичного испытания можно лишь для достоверных или невозможных событий. Случайность же события вообще не видна при единичном испытании: если событие произойдет, оно может показаться нам достоверным, если не произойдет — невозможным. Теория случайных событий может появиться лишь при большом числе испытаний, лишь для массовых событий.

Важным условием при этом является неизменность заданного комплекса основных факторов. События, происходящие при одном и том же комплексе основных факторов, называются *однородными*. Практика показывает, что события, сами по себе случайные, в большой массе при наличии однородности начинают подчиняться некоторым неслучайным закономерностям. Эти закономерности получили название *вероятностных*, а наука, изучающая вероятностные закономерности, стала называться *теорией вероятностей*.

Какой же характер имеют эти закономерности? Рассмотрим в качестве простейшего примера испытание — подбрасывание монеты. Событием пусть будет выпадение герба. Никто не возьмется предсказывать определенно, выпадет или не выпадет герб при одном подбрасывании. Но если это

испытание производится большое число раз, то мы с уверенностью можем ожидать, что герб выпадет примерно в половине случаев. Более того, если число выпадений герба сильно отличается от половины числа всех подбрасываний, то мы будем утверждать, что монета не совсем симметрична. Иными словами, такое совершенно случайное событие, как выпадение герба при подбрасывании монеты, становится эталоном качества монеты! Ясно, что такая закономерность не может быть случайной.

Примеры подобного рода можно легко продолжить. При большом числе бросаний игрального кубика, на гранях которого нанесены точки в количестве от 1 до 6, каждая грань должна выпадать в среднем один раз из шести. Вытаскивая карту из колоды, мы примерно в четверти всех случаев будем вытаскивать карту нужной масти и т. д.

Как мы видим, указанные закономерности не связаны с физической природой событий, а лишь с числом появлений этих событий при большом числе испытаний. Эти закономерности фактически опираются на следующую аксиому: если по каким-либо соображениям возможности осуществления некоторых событий одинаковы, то эти события и происходить должны одинаково часто. Так, у монеты есть две одинаковых грани, у кубика их шесть, у колоды карт — четыре масти с одинаковыми шансами быть вытащенными. События с одинаковыми возможностями осуществления мы будем называть в дальнейшем равновозможными.

Вероятностные закономерности наблюдаются и для событий с разными возможностями, хотя их и труднее предугадать. Например, если при бросании не совсем симметричной монеты из 1000 случаев герб выпал 600 раз, то мы вправе ожидать, что и в следующей тысяче подбрасываний число выпадений герба опять будет близко к 600, т. е. *частота* *) этого события остается почти неизменной.

Хотя указанная закономерность проявляется лишь при большом числе испытаний, она должна быть связана со свойствами самого события, его внутренними характеристиками; в противном случае каждая серия испытаний имела бы свои закономерности.

*) *Частотой* события называется отношение числа испытаний, в которых событие произошло, к числу всех проведенных испытаний.

Какими же свойствами должна обладать эта внутренняя характеристика события, управляющая вероятностными закономерностями? Желательно, чтобы это было число, одинаковое для равновозможных событий и в какой-то мере связанное с частотой. Рассматривать непосредственно частоту нельзя, так как частота сама есть случайная величина, зависящая от конкретной серии испытаний. Замечено, правда, что при очень большом числе испытаний частота почти перестает изменяться, приближаясь к некоторой величине, которую и можно принять за искомую характеристику. В таблице 1.1 приведены для примера результаты, полученные некоторыми экспериментаторами при бросании монеты. Как видно из таблицы, частота здесь приближается к числу $1/2$.

Т а б л и ц а 1.1

Экспериментатор	Число бросаний	Число выпадений герба	Частота
Бюффон	4 040	2 048	0,5080
К. Пирсон	12 000	6 019	0,5016
К. Пирсон	24 000	12 012	0,5005

Числовая характеристика случайного события, обладающая тем свойством, что для любой достаточно большой серии испытаний частота события лишь незначительно отличается от этой характеристики, называется *вероятностью* события.

Приведенное здесь определение вероятности называется *статистическим*. Оно позволяет вычислять вероятности таких событий, о структуре которых ничего неизвестно и частоту которых нельзя предсказать заранее. Например, только статистические данные за многие годы позволили найти вероятности рождения мальчиков и девочек. Оказалось, что эти вероятности отличны от $1/2$; вероятность рождения мальчиков равна примерно 0,52.

Статистическое определение вероятности не является достаточно строгим с точки зрения математики; из него даже не видно, всякое ли случайное событие имеет вероятность. В силу этого по статистическому определению трудно изу-

чать свойства вероятности. Непосредственно удается установить лишь следующие три факта:

- 1) вероятность достоверного события равна единице;
- 2) вероятность невозможного события равна нулю;
- 3) вероятность произвольного случайного события есть положительное число, не превосходящее единицы *).

Статистическое определение вероятности является самым широким по числу охватываемых событий. Оно ничего не требует от события, кроме принципиальной возможности проводить над ним сколь угодно большое число испытаний. Существуют другие, более удобные с формальной точки зрения определения вероятности, однако для них требуется знать структуру рассматриваемых событий.

Мы рассмотрим здесь два способа определения вероятности: классический и геометрический. С их помощью будут установлены некоторые важные для дальнейшего свойства вероятности.

1.3. Классическое определение вероятности. Будем обозначать различные случайные события прописными буквами латинского алфавита: A , B , C и т. д. Вероятности событий будем обозначать буквой P , указывая в скобках само событие — например, $P(A)$ есть вероятность события A . В этих обозначениях основное свойство вероятности (свойство 3 предыдущего пункта) запишется в виде неравенства

$$0 \leq P(A) \leq 1.$$

Будем предполагать, что задан некоторый неизменный комплекс основных факторов.

Изучая одновременно несколько случайных событий при одном и том же испытании, нетрудно заметить, что некоторые из этих событий находятся в неслучайной зависимости между собой. Скажем, появление события A всегда влечет за собой появление события B или, наоборот, события A и B не могут осуществляться одновременно. Детальное изучение таких зависимостей позволяет создать *алгебру событий* (которой мы в этой книге лишь слегка коснемся).

События A и B называются *несовместными*, если они не могут осуществиться оба в одном и том же испытании. Например, подброшенная монета не может выпасть

*) Отметим, что иногда вероятность выражается в процентах.

одновременно гербом и цифрой — значит, выпадение цифры и выпадение герба — это несовместные события.

Рассмотрим пример с бросанием игрального кубика. Выпадение четного числа очков и выпадение нечетного числа являются несовместными событиями. Если же рассмотреть такие события, как выпадение четного числа очков и выпадение числа очков, кратного трем, то они не являются несовместными. Действительно, при выпадении цифры 6 оба эти события осуществляются одновременно.

Выше мы определили понятие равновозможных событий. Предположим теперь, что все принципиально допустимые результаты испытания можно представить в виде совокупности равновозможных и попарно несовместных случайных событий (равновозможность событий оценивается обычно из каких-либо соображений симметрии). Каждое такое событие E_1, E_2, \dots назовем *исходом* испытания. Если, например, проводится испытание — бросание монеты, то исходами этого испытания можно считать выпадение герба и выпадение цифры. При бросании кубика исходом считают выпадение каждой из шести граней. Отметим, однако, что в последнем примере можно брать и другие исходы — например, выпадения четного и нечетного числа очков (два исхода). Таким образом, разбиение всех результатов испытания на исходы можно проводить не единственным способом, лишь бы сохранялась равновозможность исходов.

Обозначим число всех исходов через n . Тогда в силу равновозможности исходов вероятность каждого из них естественно положить равной $\frac{1}{n}$. Вероятность выпадения герба (так же как и выпадения цифры) равна $\frac{1}{2}$, вероятность выпадения каждой грани игрального кубика равна $\frac{1}{6}$.

Исходы испытания являются, вообще говоря, простейшими случайными событиями. Можно рассматривать более сложные события, объединяющие несколько исходов. Обычно говорят, что такому случайному событию одни исходы *благоприятствуют*, другие нет. Например, при бросании игрального кубика мы можем интересоваться таким случайным событием, как выпадение числа очков больше трех.

Этому событию благоприятствуют выпадения 4, 5 и 6 очков, т. е. три исхода из шести.

Согласно классическому определению, *вероятность случайного события равна отношению числа исходов, благоприятствующих событию, к числу всех возможных исходов*. Основное отличие классического определения от сформулированного в предыдущем пункте статистического в том, что здесь вероятность определяется до всяких испытаний, только исходя из структуры возникающих случайных событий (их разбиения на равновозможные исходы).

Рассмотрим несколько примеров вычисления классической вероятности.

Пример 1. Бросается игральный кубик. Какова вероятность того, что на его верхней грани выпадет число очков, кратное трем?

Из шести исходов интересующему нас событию благоприятствуют два: выпадение 3 и 6 очков. Таким образом, вероятность события

$$p = \frac{2}{6} = \frac{1}{3}.$$

Пример 2. В закрытой урне лежат 10 шаров одинакового размера, но разного цвета — 3 белых и 7 черных. На ощупь вытаскивается один шар. Какова вероятность того, что он окажется черным?

В данном испытании 10 исходов, соответствующих 10 шарам. Благоприятными будут лишь 7 исходов. Значит, искомая вероятность

$$p = \frac{7}{10}.$$

Задачи на вычисление классической вероятности оказываются иногда весьма сложными, так как трудно бывает подсчитать число равновозможных и благоприятных исходов. Да и само понятие равновозможности не всегда бывает очевидным. Некоторые задачи даже имеют несколько вполне правильных ответов, связанных с разными представлениями о равновозможности (симметрии) событий.

Понятие классической вероятности охватывает далеко не все случайные события, возникающие при испытании. После того, как выделены исходы испытания, можно рассматривать

лишь события, объединяющие в себе несколько исходов. Например, если при бросании кубика в качестве исходов взять выпадение четного и выпадение нечетного числа очков, то такое событие, как выпадение шести очков, вообще нельзя будет рассматривать. В связи с этим при выборе событий, считаемых в дальнейшем исходами, стараются их брать настолько «мелкими», чтобы охватить ими все необходимые события.

Отметим, что выбрать «самые мелкие» исходы, охватывающие *все* события, невозможно. Любое, самое «мелкое» событие можно еще «измельчить», добавив к нему несколько новых подробностей. Например, при бросании кубика в качестве «самых мелких» рассматриваются обычно выпадения отдельных граней. Эти исходы, однако, не смогут охватить такое событие: «кубик упал на ребро между гранями с двумя и тремя очками, прокатился еще через пять граней и остановился «шестеркой» вверх». Названное событие, как нетрудно заметить, является более мелким, чем просто выпадение «шестерки» — последняя может выпасть и при каких-нибудь других обстоятельствах.

Итак, прежде чем вычислять классическую вероятность, нужно твердо установить, какие события должны в дальнейшем играть роль исходов; это должно быть указано в условиях задачи. После того, как исходы испытания выделены, можно переходить к подсчету благоприятных для исследуемого события исходов.

Вычисление вероятности часто можно облегчить, если использовать связи между отдельными событиями. Говорят, что событие A распадается на *частные случаи* A_1 и A_2 , если события A_1 и A_2 несовместны, и событие A осуществляется в том и только том случае, когда осуществляется одно из событий A_1 или A_2 . Например, вытаскивая наудачу из урны с черными и белыми шарами два шара, мы можем рассматривать в качестве случайного события A вытаскивание двух шаров одинакового цвета. Это событие распадается на частные случаи: оба шара белые (A_1) или оба шара черные (A_2).

Пусть испытание имеет n исходов и событию A благоприятствуют m исходов. Если часть из этих m исходов, например m_1 , благоприятствует одному частному случаю A_1 , то остальные $m - m_1$ исходов должны благоприятствовать

другому частному случаю A_2 . Поэтому

$$P(A_1) = \frac{m_1}{n}, \quad P(A_2) = \frac{m - m_1}{n}.$$

Но тогда

$$P(A_1) + P(A_2) = \frac{m_1}{n} + \frac{m - m_1}{n} = \frac{m}{n} = P(A).$$

Мы доказали утверждение: *вероятность события A равна сумме вероятностей его частных случаев.*

Число частных случаев может быть и больше двух, лишь бы все эти случаи были попарно несовместны. Например, все исходы, благоприятствующие событию A , будут его частными случаями. Сформулированное выше утверждение остается справедливым для любого числа случаев.

Если события A и B несовместны, то их совместное осуществление есть невозможное событие и, значит, вероятность такого осуществления равна нулю. Если же события A и B могут произойти одновременно, то их совместное осуществление есть некоторое новое случайное событие, которое обычно обозначается через AB . Нетрудно заметить, что событие AB не может осуществляться чаще, чем каждое из событий A и B в отдельности. Поэтому

$$P(AB) \leq P(A), \quad P(AB) \leq P(B).$$

Чтобы привести пример совместного осуществления событий, рассмотрим следующее испытание: из колоды игральных карт наудачу вытаскивается одна карта. Будем считать, что осуществляется событие A , если вытащенная карта имеет пиковую масть. Будем считать, что осуществляется событие B , если вытащенная карта — дама. Тогда событие AB заключается в том, что вытащена пиковая дама.

Возникает естественный вопрос, можно ли по вероятностям событий A и B найти вероятность их совместного осуществления AB . Оказывается, это возможно не для всяких событий A и B .

Случайные события A и B называются *независимыми*, если появление или отсутствие одного из них никак не сказывается на вероятности другого события. Независимыми будут, например, события A и B в последнем примере, так как вероятность вытащить даму в каждой масти одинакова. Если бы в нашей колоде была лишняя дама пик, то событие

A (пиковая масть) было бы более благоприятным для вытягивания дамы, чем остальные масти, и значит, события A и B уже не были бы независимыми. Наилучший пример независимых событий дает повторение испытаний при одном и том же комплексе основных факторов — действительно, любые два события в разных испытаниях будут при этом независимыми (в связи с этим и сами испытания с одним и тем же комплексом основных факторов обычно называют *независимыми*). Так, независимыми будут выпадения герба в двух подбрасываниях монеты, выпадения 3 и 5 очков в двух бросаниях кубика и т. д.

Рассмотрим два независимых события A и B , которым благоприятствуют соответственно m_A и m_B исходов. Для простоты можно считать, что эти события происходят в двух различных испытаниях с числом исходов n в каждом. Тогда

$$P(A) = \frac{m_A}{n}, \quad P(B) = \frac{m_B}{n}.$$

Поскольку испытания независимы, каждому исходу первого испытания может соответствовать любой из n исходов второго. Следовательно, общее число исходов двух испытаний будет $n \cdot n = n^2$. Для того чтобы осуществилось событие AB , нужно, чтобы в первом испытании осуществился один из m_A исходов, благоприятствующих событию A , а во втором испытании осуществился один из m_B исходов, благоприятствующих событию B . При этом, как и прежде, каждому из m_A исходов первого испытания может соответствовать любой из m_B исходов второго испытания. Таким образом, всего событию AB благоприятствуют $m_A \cdot m_B$ исходов и, значит,

$$P(AB) = \frac{m_A m_B}{n^2} = \frac{m_A}{n} \cdot \frac{m_B}{n} = P(A) P(B).$$

В результате мы доказали, что *вероятность совместного осуществления независимых событий равна произведению вероятностей этих событий*.

Найдем, например, вероятность того, что при двух бросаниях игрального кубика два раза выпадет 5 очков. Вероятность одного выпадения «пятерки» равна $\frac{1}{6}$; значит, искомая

вероятность равна

$$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}.$$

Можно изучать совместное осуществление и для нескольких событий. При этом нужно рассматривать события *независимые в совокупности*, т. е. вероятность каждого из этих событий не должна зависеть от появления или отсутствия остальных событий в любых сочетаниях. В этом случае сформулированное выше утверждение о вероятности совместного осуществления независимых событий остается в силе.

Рассмотрим теперь пример, показывающий, как все найденные выше утверждения облегчают вычисление вероятности сложного события.

Пример 3. Одновременно брошены два игральных кубика. Какова вероятность, что общая сумма выпавших очков будет не меньше 11?

Разобьем заданное событие A на три частных случая: на первом кубике выпало 5 очков, на втором 6 (случай A_1), на первом кубике выпало 6 очков, на втором 5 (случай A_2) и, наконец, на обоих кубиках выпало по 6 очков (случай A_3). Вероятность выпадения 5 очков, так же как и 6 очков, для каждого кубика равна $\frac{1}{6}$, а так как оба кубика бросаются независимо друг от друга, то вероятность каждого случая есть $\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$. В результате

$$P(A) = P(A_1) + P(A_2) + P(A_3) = \frac{3}{36} = \frac{1}{12}.$$

Рассмотрим в заключение еще один вид связи между событиями. События A и \bar{A} называются *противоположными*, если они являются частными случаями достоверного события. Иными словами, события A и \bar{A} противоположны, если одно из них появляется тогда и только тогда, когда другое отсутствует. Противоположными событиями являются, например, выпадение герба и выпадение цифры при бросании монеты, попадание и промах при стрельбе в цель, выпадение «шестерки» и выпадение числа очков меньше 6 при бросании кубика. Мы видим, что фразы: «событие A не

произошло» и «произошло событие, противоположное событию A » означают абсолютно одно и то же.

Из определения противоположного события немедленно вытекает, что

$$P(A) + P(\bar{A}) = 1.$$

Переход к противоположному событию нередко облегчает вычисление вероятности.

П р и м е р 4. Найти вероятность того, что при пяти подбрасываниях монеты хотя бы один раз выпадет цифра.

Допустим, мы стали бы вычислять вероятность заданного события A непосредственно. Цифра может выпасть 1, 2, 3, 4 или 5 раз, и все это удовлетворяет условиям задачи. Кроме того, эти выпадения могут происходить в различных по порядку испытаниях и притом в разнообразных сочетаниях. Таким образом, событие A оказывается весьма сложным по своей структуре.

Рассмотрим теперь противоположное событие \bar{A} . Оно состоит в том, что цифра не выпадет ни разу, т. е. все пять раз подряд выпадет герб. Иначе говоря, событие \bar{A} равносильно совместному осуществлению пяти взаимно независимых событий с вероятностью $\frac{1}{2}$ каждое. Поэтому

$$P(\bar{A}) = \left(\frac{1}{2}\right)^5 = \frac{1}{32}.$$

А отсюда уже нетрудно вычислить вероятность события A :

$$P(A) = 1 - P(\bar{A}) = \frac{31}{32}.$$

1.4. Геометрическое определение вероятности. Классическое определение вероятности применимо лишь тогда, когда число всех равновозможных исходов конечно; в противном случае вероятность каждого такого исхода была бы равна нулю и по ним нельзя было определять вероятности сложных событий. Однако иногда удается и для бесконечного числа возможных событий использовать принцип «равновозможности», т. е. выделить такую характеристику допустимых случайных событий, которая указывала бы на одинаковые шансы событий произойти в одном и том же испытании.

Рассмотрим в качестве примера следующее испытание — в некотором квадрате случайным образом выбирается точка. Какова вероятность, что эта точка окажется внутри области D (рис. 1)?

Проводя это же испытание для двух разных областей в одном квадрате, мы, естественно, должны считать, что у случайной точки больше шансов оказаться в той области, у которой вообще «больше» точек, т. е. в области, имеющей

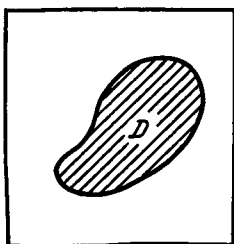


Рис. 1.

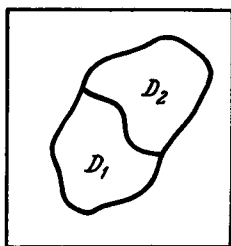


Рис. 2.

большую площадь. При равных же площадях шансы областей одинаковы. Иными словами, в качестве мерил «равновозможности» выступает здесь площадь.

Такое понятие равновозможности приводит к естественному определению вероятности того, что случайная точка попадет в область D (событие D):

$$P(D) = \frac{S_D}{S},$$

где S_D — площадь области D , S — площадь всего квадрата. Это определение вероятности называется *геометрическим*.

В рассматриваемом испытании каждому событию соответствует некоторая область. Это облегчает изучение свойств вероятности и различных связей между событиями. Например, тот факт, что событие D распадается на частные случаи D_1 и D_2 , геометрически соответствует тому, что область D распадается на две части D_1 и D_2 (рис. 2). При этом событию D_1 соответствует попадание случайной точки в область D_1 , событию D_2 — попадание в область D_2 . Несовместность событий означает, что соответствующие им области не пересекаются; совместное осуществление двух событий есть

попадание в пересечение соответствующих областей (рис. 3). Невозможному событию не соответствует ни одна область, достоверному событию соответствует весь квадрат. Отсюда ясно, что области, соответствующие двум противоположным событиям, не пересекаются и вместе заполняют весь квадрат (рис. 4).

Геометрическое определение вероятности является весьма специфичным. Однако его можно использовать и при

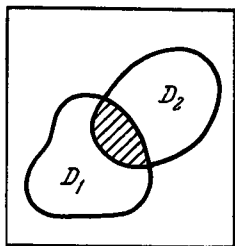


Рис. 3.

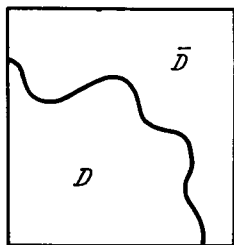


Рис. 4.

решении задач с негеометрическим содержанием, если использовать метод прямоугольных координат. Рассмотрим, например, следующую задачу, называемую обычно «задачей о встрече».

Два лица, A и B , условились встретиться в определенном месте между 7 и 8 часами вечера, причем тот, кто приходит первым, ждет другого 20 минут, после чего уходит. Чему равна вероятность их встречи, если моменты их прихода случайны и независимы друг от друга?

Обозначим через x и y время прихода A и B в минутах (сверх 7 часов). Тогда условие их встречи запишется в виде неравенства

$$|x - y| \leq 20. \quad (1.1)$$

Будем рассматривать x и y как координаты точки на плоскости. Всему рассматриваемому часу будет тогда соответствовать квадрат со сторонами 60 (рис. 5). Точками «встречи» будут все точки области, заштрихованной на рис. 5, ибо только у точек этой области координаты удовлетворяют неравенству (1.1). Поэтому искомая вероятность может быть вычислена согласно геометрическому определению и равна

отношению площадей:

$$p = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

Геометрическое определение вероятности пригодно не только для плоскости, но и для прямой или пространства. В первом случае основой испытания служит некоторый отрезок, случайным событиям соответствуют части этого отрезка, и вероятность вычисляется как отношение длины этой части к длине всего отрезка. Во втором случае испытание проводится в некотором кубе, случайным событиям соответствуют различные тела, расположенные в этом кубе, и вероятность равна отношению соответствующих объемов.

В заключение пункта остановимся еще на одной особенности геометрической вероятности. При классическом определении нулевую вероятность имело только невозможное событие. Иначе обстоят дела при геометрическом определении. Действительно, рассмотрим внутри основного квадрата отрезок какой-нибудь прямой. Случайная точка вполне может попасть на этот отрезок и, следовательно, такое событие не является невозможным. Однако площадь отрезка равна нулю, т. е. вероятность этого события, согласно геометрическому определению, равна нулю. Таким образом, при геометрическом определении вероятности *события с нулевой вероятностью могут осуществляться*. Переходя к противоположному событию, получим, что *события с вероятностью 1 могут не быть достоверными*.

Указанное свойство геометрической вероятности не противоречит общему статистическому определению вероятности, данному в прошлом пункте. Действительно, согласно общему определению, частота приближается к вероятности при увеличении числа испытаний, но она не обязана в точности совпадать с вероятностью и, в частности, для событий с нулевой вероятностью частота может отличаться от нуля.

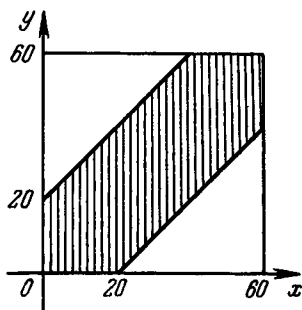


Рис. 5.

1.5. Последовательность независимых испытаний. Давая общее статистическое определение вероятности, мы требовали, чтобы при проведении большого числа испытаний частота события лишь незначительно отличалась от его вероятности. В классическом и геометрическом определениях этого условия не было, значит, его нужно проверить. Такую проверку нетрудно осуществить для последовательности независимых испытаний, так как при этом легко вычисляется вероятность появления той или иной частоты.

Допустим, проводятся n независимых испытаний (при неизменном комплексе основных факторов). Вероятность появления события A в одном отдельно взятом испытании равна p . Прежде чем рассматривать возможную частоту события A , нужно научиться решать следующую задачу: найти вероятность того, что в течение указанных n испытаний событие A осуществится ровно k раз.

Тот факт, что событие A осуществилось ровно k раз, сам по себе представляет некоторое новое случайное событие, которое мы обозначим через A^k . Таким образом, наша первая задача заключается в вычислении вероятности $P(A^k)$. Событие A^k распадается на частные случаи A_1^k, A_2^k, \dots , каждый из которых требует, чтобы событие A осуществилось k раз в определенных по счету испытаниях. Например, осуществления события A в первых k испытаниях или в последних k испытаниях будут разными частными случаями.

Число этих частных случаев равно числу способов, которыми k появлений события A можно разместить среди всех n испытаний, т. е. числу сочетаний из n по k (обозначается C_n^k). Если воспользоваться понятием факториала *), то можно записать равенство

$$C_n^k = \frac{n!}{k!(n-k)!}.$$

Все испытания независимы и, кроме номеров, ничем не отличаются друг от друга. Поэтому все частные случаи A_1^k, A_2^k, \dots равновозможны, благодаря чему достаточно вычислить

*) Факториалом числа n называется произведение всех целых чисел от 1 до n ; обозначается $n!$.

вероятность одного из них. Вычислим, например, вероятность того, что событие A осуществилось в первые же k раз. При этом нужно учесть, что в остальных $n - k$ испытаниях событие A не должно осуществиться ни разу, т. е. все $n - k$ раз должно осуществиться противоположное событие \bar{A} с вероятностью $q = 1 - p$. Используя независимость испытаний, получим, что $P(A_1^k) = p^k q^{n-k}$.

Число частных случаев A_i^k , как уже указывалось, равно C_n^k , и все они равновозможны (имеют одинаковую вероятность). Поэтому

$$P(A^k) = P(A_1^k) + P(A_2^k) + \dots = C_n^k p^k q^{n-k}. \quad (1.2)$$

Полученной формулой удобно пользоваться при решении задач, связанных с повторением испытаний.

Пример. Вероятность попадания в мишень при одном выстреле равна 0,8. Найти вероятность того, что из 5 выстрелов 3 поразят мишень.

Здесь $n=5$, $k=3$, $p=0,8$ и $q=0,2$. Непосредственно по формуле находим, что

$$P(A^3) = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}{1 \cdot 2 \cdot 3 \cdot 1 \cdot 2} (0,8)^3 (0,2)^2 = 10 \cdot 0,512 \cdot 0,04 = 0,2048.$$

Частота события A определяется по формуле $\omega = \frac{n_A}{n}$, где n — число всех испытаний, n_A — число появлений события A . Поэтому $P(A^k)$ можно считать вероятностью того, что $\omega = k/n$. Из (1.2) вытекает тогда формула

$$P\left\{\omega = \frac{k}{n}\right\} = \frac{n!}{k!(n-k)!} p^k q^{n-k},$$

которая позволяет вычислять вероятность всех возможных значений частоты ω . Непосредственными расчетами нетрудно проверить, что при увеличении n вероятности большинства частот будут резко стремиться к нулю и лишь для нескольких частот, лежащих близ самой вероятной частоты, эти вероятности будут сохранять заметное значение.

Попытаемся вычислить самую вероятную (т. е. обладающую наибольшей вероятностью) частоту. Для этого изучим повнимательнее функцию

$$f(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}.$$

Составим отношение

$$\frac{f(k+1)}{f(k)} = \frac{k! (n-k)!}{(k+1)! (n-k-1)!} \cdot \frac{p^{k+1} q^{n-k-1}}{p^k q^{n-k}} = \frac{n-k}{k+1} \cdot \frac{p}{q}.$$

Мы видим, что это отношение убывает с возрастанием k . Пока оно больше 1, функция $f(k)$ возрастает, когда оно станет меньше 1, функция $f(k)$ начнет убывать. Следовательно, для некоторого k функция $f(k)$ имеет максимум, который и соответствует самой вероятной частоте. Если бы k могло быть нецелым, этот максимум достигался бы как раз тогда, когда

$$\frac{n-k}{k+1} \cdot \frac{p}{q} = 1,$$

т. е. при $k = np - q$. Вообще же говоря, максимум функции $f(k)$ будет при k , равном ближайшему к $np - q$ целому числу. Это k и будет самым вероятным числом появлений события A при n испытаниях. Из сказанного выше следует, что оно всегда удовлетворяет неравенству $|k - (np - q)| < 1$.

Найденному числу k соответствует самая вероятная частота $\omega = \frac{k}{n}$, удовлетворяющая неравенству

$$\left| \omega - \left(p - \frac{q}{n} \right) \right| < \frac{1}{n}.$$

Из этого неравенства немедленно вытекает, что при $n \rightarrow \infty$ вероятнейшая частота $\omega \rightarrow p$. Тем самым доказано утверждение: *для достаточно большого числа испытаний вероятнейшая частота события мало отличается от его вероятности.*

Полученное утверждение полностью согласуется с основным требованием к вероятности. В реальных испытаниях, правда, осуществляется не только самая вероятная частота, но и некоторые близкие к ней частоты. Можно, однако, доказать следующее более общее утверждение (теорема Бернулли): *каково бы ни было наперед заданное положительное число ε , вероятность того, что частота события отличается от его вероятности больше, чем на ε , стремится к нулю при неограниченном увеличении числа испытаний.* Это утверждение мы докажем ниже, в п. 3.3

§ 2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

2.1. Дискретные и непрерывные случайные величины. В предыдущем параграфе мы рассматривали случайные события произвольной природы. Перейдем теперь к более детальному рассмотрению событий, состоящих в появлении того или иного числа. Из ранее встречавшихся к таким событиям относятся, например, результаты бросания игрального кубика — выпадение 1, 2, 3, 4, 5 или 6 очков.

Интерес к изучению указанных событий связан в первую очередь с тем, что именно к ним относятся результаты большинства наблюдений. Случайным является число космических частиц, регистрируемых счетчиком за одну секунду. Случайными являются количества примесей в используемом при реакции веществе. Даже самый точный метод анализа вещества дает при повторениях некоторое расхождение в результатах (ошибка воспроизводимости), значит, и здесь каждый числовой результат есть случайное событие.

Рассмотренные примеры приводят нас к важному понятию *случайной величины*. Случайной величиной называется величина, принимающая в результате испытания числовое значение, которое принципиально нельзя предсказать, исходя из условий испытания.

Случайная величина обладает целым набором допустимых значений, но в результате каждого отдельного испытания принимает лишь какое-то одно из них. В отличие от изучаемых в математическом анализе переменных величин, изменяющих свое значение лишь при изменении условий испытания, случайная величина может принимать различные значения даже при неизменном комплексе основных факторов. Причина изменения случайной величины от испытания к испытанию кроется в неучитываемых нами факторах, которые мы выше называли случайными.

Чтобы в достаточной степени охарактеризовать случайную величину, нужно прежде всего задать набор ее допустимых значений. Эти допустимые значения могут быть как ограничены, так и неограничены в совокупности; в зависимости от этого и сама случайная величина называется *ограниченной* или *неограниченной*.

Наиболее важной является классификация случайных величин в зависимости от числа их допустимых значений. Это число может быть самым разнообразным — от двух до бесконечности (оно не может равняться единице, так как при этом величина перестанет быть случайной).

Если число допустимых значений случайной величины конечно, то такая величина называется *конечнозначной*. Конечнозначная случайная величина обязательно ограничена.

Примеры конечнозначных случайных величин встречаются на каждом шагу даже в повседневной жизни. Именно к таким величинам относятся число зрителей в кинотеатре во время киносеанса, выручка магазина за день, номер обуви случайного прохожего и т. д. С конечнозначными случайными величинами можно встретиться и при наблюдениях, особенно, если результат может выражаться только в целых числах. Например, при изучении ядерных процессов регистрируется число вспышек на соответствующем фотоснимке.

Любое случайное событие, даже не имеющее числовой природы, можно сопоставить с конечнозначной случайной величиной следующего типа: каждому появлению события A ставится в соответствие число 1, неоявлению — число 0. Еще более важной случайной величиной такого типа является число появлений события A в n независимых испытаниях.

Перечисленные примеры показывают важность изучения конечнозначных случайных величин. Однако в практических приложениях, особенно в теории наблюдений, чаще приходится иметь дело с величинами, возможное число значений которых бесконечно. Поясним это на примере измерения давления газа. Это давление под влиянием множества случайных причин все время изменяется (хотя, быть может, и незначительно), поэтому оно является случайной величиной. Если в результате двух измерений получились значения

0,72 и 0,73 атм, то мы, по крайней мере принципиально, должны допустить возможность всех промежуточных значений между 0,72 и 0,73 атм.

Случайные величины с бесконечным числом допустимых значений имеют свою классификацию. Случайная величина называется *дискретной*, если между любыми двумя ее значениями заключено лишь конечное число других допустимых значений. Каждое значение дискретной величины отделено от соседних некоторыми промежутками, как бы оторвано от них, поэтому такую величину иногда называют *прерывной*. В качестве примера дискретной случайной величины можно рассмотреть число независимых испытаний, необходимых для того, чтобы некоторое случайное событие A появилось ровно n раз. Действительно, допустимыми значениями этой величины будут все целые значения, начиная с n , которые заведомо отстоят друг от друга по крайней мере на единицу.

Отметим, что определению дискретной величины удовлетворяют формально и конечнозначные величины (у которых вообще всех значений конечное число). В силу этого часто конечнозначные величины также называют дискретными.

Если значения случайной величины могут сплошь заполнять некоторый промежуток, как в приводившемся выше примере с измерением давления газа, то эти значения уже нельзя отделить друг от друга промежутками. Следовательно, случайная величина с таким набором допустимых значений уже не будет дискретной.

Недискретные случайные величины могут быть весьма сложными по своей структуре. Мы будем изучать лишь один класс таких величин — *непрерывные* случайные величины, строгое определение которых будет дано в следующем пункте.

2.2. Распределение случайной величины. Набор допустимых значений сам по себе очень слабо характеризует случайную величину. Приведем такой пример. Скорость движения молекулы газа постоянно меняется в связи с постоянными случайными столкновениями с другими молекулами. В силу этого она является случайной величиной с широким диапазоном допустимых значений. Если теперь

рассматривать газ при различных температурах, то допустимые значения скоростей молекул будут одни и те же, хотя состояния газа будут различными.

Для того чтобы полностью охарактеризовать случайную величину, необходимо не только указать, какие значения она может принимать, но и как часто, т. е. с какой вероятностью она принимает эти значения. Иными словами, нужно задать *распределение* этой случайной величины.

Легче всего это сделать для конечнозначной величины. Здесь можно непосредственно указать, с какой вероятностью она принимает каждое из своих допустимых значений. Описание совокупности значений случайной величины с указанием вероятности каждого значения называется *законом распределения* этой величины.

Пусть x_1, x_2, \dots, x_n — возможные значения случайной величины; через p_i обозначим вероятность принять значение x_i (в одном испытании). Тот факт, что случайная величина примет в результате испытания одно из своих значений, есть достоверное событие; при этом она не может одновременно принять более одного значения. Таким образом, значения x_1, x_2, \dots, x_n являются частными случаями достоверного события, в силу чего

$$p_1 + p_2 + \dots + p_n = 1.$$

Рассмотрим, например, распределение числа очков на верхней грани брошенного кубика. Здесь могут быть значения 1, 2, 3, 4, 5 и 6, а вероятности всех значений равны $\frac{1}{6}$. Мы получили пример так называемого *конечнозначного равномерного распределения*, т. е. распределения, при котором все допустимые значения конечнозначной случайной величины имеют одинаковые вероятности.

В более общем случае, чтобы задать закон распределения случайной величины, нужно просто выписать все ее значения и при каждом из них соответствующую вероятность. Обычно такую запись оформляют в виде таблицы, где верхняя строка содержит значения случайной величины, а нижняя — вероятности этих значений. Полученная таблица называется *рядом распределения* случайной величины.

Найдем в качестве примера закон распределения суммы очков, выпавших на двух одновременно брошенных играль-

ных кубиках. При бросании двух кубиков могут возникнуть 36 различных комбинаций выпавших очков, однако некоторые комбинации дадут одинаковые суммы. Подсчитывая число возможных комбинаций для каждого значения суммы, мы с помощью классического определения найдем все соответствующие вероятности. Так, два очка могут образоваться лишь одним способом — по одному очку на каждом кубике, значит, вероятность значения 2 равна $\frac{1}{36}$. Три очка могут уже образоваться двумя способами — одно очко на первом кубике, два на втором или два очка на первом кубике, одно на втором. В результате значение 3 будет иметь вероятность $\frac{2}{36} = \frac{1}{18}$. Продолжая вычисления, получим весь ряд распределения заданной величины, приведенный в таблице 2.1 (рекомендуем читателям самостоятельно проверить остальные значения этой таблицы).

Таблица 2.1

x	2	3	4	5	6	7	8	9	10	11	12
p	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Самой важной для приложений конечнозначной случайной величиной является число появлений некоторого события A с вероятностью p в последовательности n независимых испытаний. Эта величина может принимать любое значение $k=0, 1, 2, \dots, n$. Закон ее распределения был фактически найден в п. 1.5: если обозначить вероятность значения k через $P_n(k)$ (в п. 1.5 она обозначалась $P(A^k)$), то

$$P_n(k) = C_n^k p^k q^{n-k}. \quad (2.1)$$

Вероятности (2.1) полностью совпадают с членами разложения бинома Ньютона $(p+q)^n$, поэтому и само распределение называют *биномиальным*. Это же совпадение позволяет легко проверить, что сумма всех вероятностей в биномиальном распределении равна единице. Действительно,

$$P_n(0) + P_n(1) + \dots + P_n(n) = (p+q)^n = 1^n = 1.$$

Для дискретной случайной величины также можно задавать закон распределения, так как все ее значения можно отделить друг от друга и перенумеровать. Однако при этом соответствующие вероятности уже нельзя просто перечислить — нужно указать общее правило их нахождения. В качестве важного примера приведем *распределение Пуассона*. Случайная величина, подчиняющаяся закону Пуассона, принимает любое значение $k=0, 1, 2, \dots$ с вероятностями

$$P(k) = \frac{a^k e^{-a}}{k!},$$

где a — некоторый положительный параметр.

Значения дискретной случайной величины, так же как и конечнозначной, являются в совокупности частными случаями достоверного события. Поэтому сумма вероятностей этих значений (правильнее говоря, сумма ряда, составленного из этих вероятностей) должна быть также равна единице. Нетрудно проверить, что это условие выполнено для распределения Пуассона:

$$\sum_{k=0}^{\infty} \frac{a^k e^{-a}}{k!} = e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} = e^{-a} e^a = 1.$$

Распределение недискретной случайной величины уже нельзя задавать с помощью вероятностей отдельных значений. Число значений здесь так велико (их обычно даже нельзя перенумеровать), что для большинства из них вероятность принять это значение равна нулю. Здесь возникает та же ситуация, что и для геометрической вероятности (см. п. 1.4): событие может произойти, а вероятность его равна нулю.

Поскольку мы уже упомянули о геометрическом определении вероятности, рассмотрим его в качестве простейшего примера недискретной случайной величины. Возьмем за основу испытания отрезок $[0,1]$ некоторой числовой оси и будем в нем выбирать случайную точку. В качестве случайной величины будем рассматривать координату этой точки, обозначаемую через x . Согласно геометрическому определению, вероятность любой координаты равна нулю. Однако сам этот факт еще не определяет случайной величины — подобному условию удовлетворяют и многие другие величины.

Геометрическое определение позволяет находить вероятности более сложных событий — попаданий случайной точки в ту или иную часть отрезка. В нашем примере эта вероятность численно равна длине рассматриваемой части *). С точки зрения случайной величины такие события можно описать следующим образом: *изучается вероятность того, что в результате испытания значение случайной величины попадет в некоторую заранее намеченную совокупность чисел.*

Аналогичную задачу можно ставить для произвольной случайной величины. При этом наибольший интерес представляют совокупности чисел, определяемые одним или несколькими неравенствами. Иными словами, нужно уметь находить вероятность того, что случайная величина в результате испытания примет значение, удовлетворяющее одному или нескольким заданным неравенствам. Эта задача является особенно важной в теории наблюдений, где нужно знать вероятность того, что ошибка (случайная величина) не превзойдет некоторого допустимого предела.

Вероятность, с которой случайная величина удовлетворяет некоторым неравенствам, обозначается через $P\{ \}$, где в фигурных скобках выписываются заданные неравенства. Например, $P\{a \leq \xi < b\}$ означает вероятность того, что случайная величина ξ в результате испытания примет значение, удовлетворяющее неравенствам $a \leq \xi < b$.

Пусть ξ — случайная величина и x — произвольное действительное число. Вероятность того, что ξ примет значение меньшее, чем x , является функцией от x :

$$F(x) = P\{\xi < x\},$$

она называется *функцией распределения* случайной величины ξ .

Функцию распределения можно определить для любой случайной величины, независимо от числа ее допустимых значений, однако наиболее полезна эта функция для описания не дискретных случайных величин, распределение которых мы пока вообще не умеем описывать.

*) Существуют части отрезка, для которых нельзя определить длину, например, совокупность всех точек с рациональными координатами. Соответствующие таким частям события мы просто не будем рассматривать.

Как всякая вероятность, функция распределения удовлетворяет неравенствам

$$0 \leq F(x) \leq 1.$$

Если случайная величина ограничена, т. е. все ее допустимые значения лежат в некотором отрезке $[a, b]$, то $F(x) = 0$ для всех $x < a$ (невозможное событие) и $F(x) = 1$ для всех $x \geq b$ (достоверное событие).

Зная $F(x)$, нетрудно найти $P\{x_1 \leq \xi < x_2\}$ для любых x_1 и x_2 . Действительно, событие, состоящее в выполнении неравенства $\xi < x_2$, распадается на частные случаи: а) выполнено неравенство $\xi < x_1$; б) выполнено неравенство $x_1 \leq \xi < x_2$. Поэтому

$$P\{\xi < x_2\} = P\{\xi < x_1\} + P\{x_1 \leq \xi < x_2\},$$

откуда

$$P\{x_1 \leq \xi < x_2\} = P\{\xi < x_2\} - P\{\xi < x_1\},$$

и, следовательно,

$$P\{x_1 \leq \xi < x_2\} = F(x_2) - F(x_1).$$

Отсюда, в частности, вытекает, то $F(x_2) \geq F(x_1)$ — ведь вероятность $P\{x_1 \leq \xi < x_2\}$ не может быть отрицательной. Иными словами, *функция распределения любой случайной величины есть неубывающая функция.*

Неравенство $\xi \geq x$ является противоположным событием для неравенства $\xi < x$, поэтому

$$P\{\xi \geq x\} = 1 - P\{\xi < x\} = 1 - F(x).$$

Благодаря этому, по функции распределения случайной величины удастся вычислять вероятности всех основных неравенств, встречающихся при обработке наблюдений. Правда, в некоторых неравенствах у нас стоит нестрогий знак (\leq или \geq), а в других строгий ($<$ или $>$), и знаки эти не всегда можно заменять друг на друга. Рассмотрим, например, неравенства $\xi < x$ и $\xi \leq x$; нетрудно видеть, что

$$P\{\xi \leq x\} = P\{\xi < x\} + P\{\xi = x\}.$$

Поэтому $P\{\xi \leq x\} \neq P\{\xi < x\}$, если $P\{\xi = x\} \neq 0$, т. е. если случайная величина ξ принимает значение x с ненулевой

вероятностью. Как мы уже знаем, такие значения есть и у конечнозначной, и у дискретной величины — это как раз все их допустимые значения. При этих x функция распределения $F(x)$ имеет скачок. На рис. 6 приводится график функции распределения для числа выпадений герба при трех бросаниях монеты. При его составлении использовался тот факт, что для конечнозначной величины функция распределения $F(x)$ равна сумме вероятностей всех допустимых значений, меньших, чем x .

Если функция распределения непрерывна, то соответствующая случайная величина также называется *непрерывной*. Непрерывная случайная величина каждое свое значение принимает с нулевой вероятностью, поэтому для нее строгие и нестрогие знаки неравенства можно не различать.

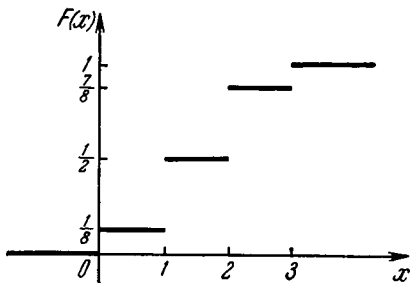


Рис. 6.

Непрерывные случайные величины будут основными в нашем дальнейшем изложении, ибо именно к ним в большинстве своем относятся результаты количественных наблюдений (см. об этом подробнее в § 4).

Функция распределения случайной величины скрадывает распределение вероятностей по отдельным значениям этой величины. Увеличиваясь от значения к значению, она является как бы функцией «накопленной вероятности». Вероятность $P\{x \leq \xi < x + \Delta x\}$ наглядней характеризует значение x , особенно при малых Δx . Однако, эта вероятность зависит еще и от Δx и стремится к нулю при неограниченном уменьшении Δx . Деля $P\{x \leq \xi < x + \Delta x\}$ на Δx , мы от последнего недостатка можем, вообще говоря, избавиться, но остается еще зависимость от Δx . И наконец, зависимость от Δx исчезает, если в полученном частном устремить Δx к нулю и рассмотреть функцию

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P\{x \leq \xi < x + \Delta x\}}{\Delta x}.$$

Эта функция называется *плотностью распределения* случайной величины. В ее определении участвует предел, который не всегда существует. Поэтому в дальнейшем будут рассматриваться только такие непрерывные величины, у которых можно определить плотность.

Плотность легко выразить через функцию распределения:

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x),$$

т. е. *плотность есть производная функции распределения*. Поэтому по формуле Ньютона — Лейбница

$$\int_a^b f(x) dx = F(b) - F(a).$$

А это значит, что с помощью плотности можно вычислять вероятность неравенства $a \leq \xi < b$:

$$P\{a \leq \xi < b\} = \int_a^b f(x) dx.$$

В частности,

$$F(x) = P\{-\infty < \xi < x\} = \int_{-\infty}^x f(x) dx.$$

Отсюда же можно вывести еще одно важное свойство плотности:

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

Действительно, попадание случайной величины в интервал $-\infty < \xi < \infty$ есть достоверное событие.

Наконец, отметим, что в силу самого своего определения плотность не может быть отрицательной.

Связь между плотностью и вероятностью хорошо иллюстрируется геометрически: вероятность того, что непрерывная случайная величина примет значение из интервала (a, b) , равна площади криволинейной трапеции, ограниченной осью x , прямыми $x=a$ и $x=b$ и графиком плотности $f(x)$ (рис. 7).

Выше мы рассматривали случайную величину, связанную с геометрическим определением вероятности — координату точки, случайно выбранной на отрезке $[0,1]$. Согласно

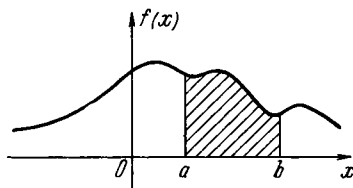


Рис. 7.

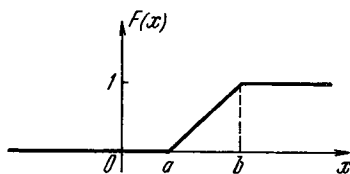


Рис. 8.

геометрическому определению, вероятность попадания случайной точки в интервал (a, b) будет здесь равна длине интервала $b-a$. Поэтому

$$P\{x \leq \xi < x + \Delta x\} = \Delta x,$$

если только $0 \leq x < 1 - \Delta x$. Отсюда

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta x}{\Delta x} = 1,$$

т. е. плотность постоянна для всех x из отрезка $[0,1]$ и равна единице. Если вместо отрезка $[0,1]$ за основу испытания взять какой-нибудь другой отрезок длины l , то получится плотность $f(x) = \frac{1}{l}$ для всех точек x из рассматриваемого отрезка. Вне этого отрезка плотность, очевидно, будет равна нулю.

Распределение, плотность которого постоянна на некотором отрезке, а вне его равна нулю, называется *непрерывным равномерным распределением*. На рис. 8 приведен график соответствующей функции распределения.

2.3. Числовые характеристики случайных величин.

В приложениях теории вероятностей нередко бывает, что распределение случайной величины известно с точностью до одного-двух неопределенных параметров. В этом случае для описания случайной величины достаточно лишь дополнительно задать несколько общих числовых характеристик распределения. Наиболее употребительны при этом такие

характеристики, как *математическое ожидание* и *дисперсия* случайной величины.

Рассмотрим вначале конечнозначную случайную величину ξ с рядом распределения

$$x_1, x_2, \dots, x_n,$$

$$p_1, p_2, \dots, p_n.$$

Математическим ожиданием этой величины называется число

$$M\xi = p_1x_1 + p_2x_2 + \dots + p_nx_n.$$

Если рассматривать x_1, x_2, \dots, x_n , как координаты точек, лежащих вдоль некоторого стержня, а p_1, p_2, \dots, p_n — как веса

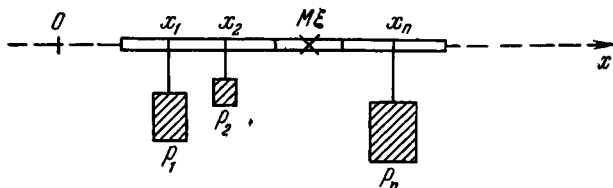


Рис. 9.

грузов, подвешенных в этих точках (рис. 9), то $M\xi$ будет совпадать с координатой центра тяжести образовавшейся системы. Поэтому $M\xi$ называют иногда *средневзвешенным значением* случайной величины.

Если конечнозначная случайная величина имеет равномерное распределение, т. е. $p_1 = p_2 = \dots = p_n = \frac{1}{n}$, то

$$M\xi = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Таким образом, *математическое ожидание конечнозначной равномерно распределенной случайной величины есть среднее арифметическое ее значений.*

Найдем смысл математического ожидания для произвольной конечнозначной случайной величины. Допустим, что произведено N независимых испытаний над случайной величиной ξ , в результате которых значение x_1 появилось k_1 раз, значение x_2 появилось k_2 раз и т. д. Очевидно, что

$$k_1 + k_2 + \dots + k_n = N.$$

Отношения $\frac{k_1}{N}, \frac{k_2}{N}, \dots, \frac{k_n}{N}$ будут частотами появления соответствующих значений. Если число испытаний N велико, то эти частоты, согласно общему определению вероятности, должны лишь незначительно отличаться от соответствующих вероятностей:

$$\frac{k_1}{N} \approx p_1, \quad \frac{k_2}{N} \approx p_2, \quad \dots, \quad \frac{k_n}{N} \approx p_n.$$

Поэтому, заменяя в формуле математического ожидания вероятности p_1, p_2, \dots, p_n на частоты $\frac{k_1}{N}, \frac{k_2}{N}, \dots, \frac{k_n}{N}$, мы допустим лишь незначительную ошибку, т. е.

$$M\xi \approx \frac{k_1 x_1 + k_2 x_2 + \dots + k_n x_n}{N}.$$

В числителе получившейся дроби каждое значение x_i повторяется слагаемым (с помощью умножения на k_i) ровно столько раз, сколько раз это значение возникало в процессе испытаний. Иными словами, получившийся числитель равен сумме всех N результатов испытаний (каждый результат прибавляется к общей сумме независимо от того, встречался он уже раньше или нет). А так как эта сумма затем делится на N , то оказывается, что *математическое ожидание конечнозначной случайной величины приближенно равно среднему арифметическому всех результатов, полученных при большом числе испытаний над этой величиной.*

Формула математического ожидания легко обобщается на дискретные случайные величины. Математическим ожиданием такой величины называется сумма ряда

$$M\xi = \sum_{k=1}^{\infty} p_k x_k;$$

если ряд расходится, то математическое ожидание равно бесконечности.

Найдем, например, математическое ожидание в случае распределения Пуассона (см. предыдущий пункт). Случайная величина ξ , распределенная по закону Пуассона, принимает значения $k=0, 1, 2, \dots$ с вероятностями

$$p_k = \frac{a^k e^{-a}}{k!}.$$

Поэтому

$$M\xi = \sum_{k=1}^{\infty} k \frac{a^k e^{-a}}{k!} = a \sum_{k=1}^{\infty} \frac{a^{k-1} e^{-a}}{(k-1)!}.$$

Но

$$\sum_{k=1}^{\infty} \frac{a^{k-1} e^{-a}}{(k-1)!} = e^{-a} \sum_{k=1}^{\infty} \frac{a^{k-1}}{(k-1)!} = e^{-a} e^a = 1.$$

Следовательно, $M\xi = a$. Мы видим, что математическое ожидание совпадает с основным параметром распределения Пуассона.

Математическое ожидание непрерывной случайной величины определяется с помощью интеграла

$$M\xi = \int_{-\infty}^{\infty} x f(x) dx,$$

где $f(x)$ — плотность распределения.

Рассмотрим в качестве примера непрерывное равномерное распределение на отрезке $[0, 1]$, плотность которого была найдена в предыдущем пункте:

$$f(x) = \begin{cases} 1, & \text{если } 0 \leq x \leq 1, \\ 0 & \text{для прочих } x. \end{cases}$$

Для этого распределения

$$M\xi = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}.$$

Можно показать, что и для любого отрезка математическое ожидание непрерывной равномерно распределенной случайной величины есть середина этого отрезка.

Найденный выше смысл математического ожидания как среднего арифметического результатов при большом числе испытаний сохраняется для дискретных и непрерывных случайных величин. Это позволяет находить математическое ожидание любой случайной величины опытным путем (приближенно).

Математическое ожидание случайной величины обладает рядом свойств, облегчающих его вычисление. Мы приведем

эти свойства без доказательства (их нетрудно проверить, например, для равномерного распределения, что мы и рекомендуем проделать читателю).

1. Если все значения случайной величины, не меняя их вероятностей, уменьшить (увеличить) на некоторое число, то математическое ожидание уменьшится (увеличится) на это же число.

2. Если все значения случайной величины, не меняя их вероятностей, уменьшить (увеличить) в некоторое число раз, то математическое ожидание уменьшится (увеличится) во столько же раз.

Благодаря этим двум свойствам при вычислении математического ожидания (в частности, среднего арифметического) можно все значения сокращать, избавляться от дробей, приводить значения к новому началу отсчета и т. п. Например, разыскивая среднее арифметическое чисел 47,3; 47,1; 47,2; 46,8; 46,9; 47,1, можно отнять от всех чисел 47 и умножить их на 10. В результате получатся новые числа,

$$3, 1, 2, -2, -1, 1,$$

среднее арифметическое которых подсчитывается много легче, чем для прежних, и равно $\frac{2}{3}$. Значит, среднее арифметическое первоначальных чисел равно $\frac{1}{10} \cdot \frac{2}{3} + 47 = 47\frac{2}{30}$.

Следующие два свойства математического ожидания связаны с действиями над случайными величинами — сложением и умножением. Случайные величины складываются и умножаются своими значениями. При этом слагаемые (или сомножители) могут встречаться в различных комбинациях; вероятность каждой такой комбинации подсчитывается по законам теории вероятностей, если только известны вероятности сомножителей (слагаемых). Примером суммы двух случайных величин может служить сумма очков, выпавших на двух игральных кубиках — закон распределения этой суммы был найден в предыдущем пункте (таблица 2.1).

Распределение суммы или произведения двух случайных величин можно найти, если известны распределения этих величин. Однако расчеты при этом, как правило, бывают очень сложными. Тем ценнее возможность находить

математические ожидания суммы и произведения непосредственно по математическим ожиданиям слагаемых или сомножителей.

Продолжим изложение свойств математического ожидания.

3. *Математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых.*

4. *Математическое ожидание произведения независимых*) случайных величин равно произведению математических ожиданий сомножителей.*

В качестве приложения этих свойств найдем математическое ожидание числа появлений события A с вероятностью p в n независимых испытаниях (биномиальное распределение). Это число k равно сумме n случайных величин:

$$k = \mu_1 + \mu_2 + \dots + \mu_n,$$

где каждое μ_i равно 1, если событие A появилось в i -м испытании, и равно 0, если событие A в этом испытании не появилось. Все μ_i имеют одинаковое распределение: они принимают значение 1 с вероятностью p и значение 0 с вероятностью $q=1-p$. Поэтому все они имеют одинаковое математическое ожидание

$$M\mu_i = p \cdot 1 + q \cdot 0 = p.$$

Но тогда в силу третьего свойства

$$Mk = M\mu_1 + M\mu_2 + \dots + M\mu_n = np.$$

Таким образом, математическое ожидание числа появлений события A в n испытаниях оказалось равным np . В п. 1.5 было найдено наивероятнейшее число появлений события A . Легко видеть, что это число отличается от np менее, чем на 1. Поэтому в случае, когда np есть целое число, *наивероятнейшее число появлений события A и математическое ожидание числа появлений совпадают друг с другом*. Это раскрывает еще один смысл математического ожидания.

Если представлять себе, что значения случайной величины рассеяны вдоль числовой оси, то математическое ожидание

*) Случайные величины называют *независимыми*, если каждая из них имеет самостоятельное распределение, не зависящее от возможных значений других величин. Более подробно об этом см. ниже, в § 9.

играет роль центра этого рассеяния. Очевидно, нужна еще одна общая числовая характеристика, показывающая, как сильно рассеяны значения случайной величины вокруг этого центра.

Дисперсией случайной величины называется математическое ожидание квадрата отклонения этой величины от ее математического ожидания:

$$D\xi = M(\xi - M\xi)^2.$$

Для конечнозначной случайной величины это дает формулу

$$D\xi = p_1(x_1 - M\xi)^2 + p_2(x_2 - M\xi)^2 + \dots + p_n(x_n - M\xi)^2.$$

Для дискретной величины появится сумма ряда:

$$D\xi = \sum_{k=1}^{\infty} p_k(x_k - M\xi)^2,$$

для непрерывной величины появится интеграл:

$$D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 f(x) dx.$$

Чтобы выяснить целесообразность понятия дисперсии, рассмотрим конечнозначную случайную величину ξ с равномерным распределением:

$$p_1 = p_2 = \dots = p_n = \frac{1}{n}.$$

Математическое ожидание такой величины есть просто среднее арифметическое ее значений. Чтобы охарактеризовать рассеяние значений, нужно как-то усреднить все отклонения этой величины от $M\xi$. Но если брать просто среднее арифметическое всех отклонений, то отклонения с разными знаками взаимно погасятся, и результат всегда будет равен нулю. Чтобы избежать этого, нужно брать отклонения по абсолютной величине, либо возвести их в квадрат. Абсолютная величина неудобна для аналитических выкладок (ее нельзя, например, дифференцировать), остается возведение в квадрат. Это даст в качестве меры рассеяния число

$$\frac{(x_1 - M\xi)^2 + (x_2 - M\xi)^2 + \dots + (x_n - M\xi)^2}{n},$$

которое в точности совпадает с дисперсией рассматриваемой случайной величины.

Таким образом, *дисперсия есть естественная, простейшая мера рассеяния случайной величины вокруг ее математического ожидания.*

Для практических приложений часто оказывается удобной другая формула дисперсии.

Используя свойства математического ожидания, можно записать, что

$$D\xi = M[\xi^2 - 2\xi M\xi + (M\xi)^2] = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2,$$

т. е., окончательно,

$$D\xi = M\xi^2 - (M\xi)^2.$$

При вычислении дисперсии также полезно опираться на ее свойства.

1. *Дисперсия случайной величины не изменится, если все ее значения уменьшить (увеличить) на одно и то же число, не меняя их вероятностей.*

2. *Если все значения случайной величины, не меняя их вероятностей, умножить на некоторый множитель, то дисперсия умножится на квадрат этого множителя.*

3. *Дисперсия суммы независимых случайных величин равна сумме дисперсий этих величин.*

Найдем, пользуясь третьим свойством, дисперсию числа появлений k события A в n независимых испытаниях (биномиальное распределение). Для этого найдем вначале дисперсии величин μ_i , определенных при вычислении математического ожидания числа k .

Мы получим, что

$$\begin{aligned} D\mu_i &= M(\mu_i - M\mu_i)^2 = M(\mu_i - p)^2 = (1-p)^2 p + (0-p)^2 q = \\ &= q^2 p + p^2 q = pq(q+p) = pq, \end{aligned}$$

т. е. дисперсии всех μ_i совпадают между собой. А тогда

$$Dk = D\mu_1 + D\mu_2 + \dots + D\mu_n = npq.$$

Третье свойство дисперсии особенно важно при изучении воздействия каких-либо факторов на результат наблюдения. Эти факторы редко удается изучать по отдельности, поэтому

по наблюдениям определяется, как правило, лишь общая дисперсия всех изучаемых факторов. Если факторы независимы, общая дисперсия есть просто сумма дисперсий, связанных с каждым фактором по отдельности. Убрав один фактор (а это обычно можно сделать), мы сразу же увидим, насколько изменилась общая дисперсия; разность как раз и даст нам дисперсию убранного фактора, которую в чистом виде выделить было невозможно. Эта идея лежит в основе дисперсионного анализа, изучаемого ниже, в § 8.

По сравнению со значениями случайной величины дисперсия измеряется в квадратных единицах. Для того чтобы иметь меру рассеяния, сопоставимую со значениями, часто рассматривают корень квадратный из дисперсии, который называют *средним квадратичным отклонением* случайной величины.

§ 3. НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

3.1. Пример нормального распределения (задача о рассеянии снарядов). Мы уже неоднократно отмечали, что большинство результатов, получаемых при наблюдениях, являются случайными величинами. Для того чтобы иметь возможность обрабатывать такие результаты, а тем более делать дальнейшие прогнозы, необходимо знать распределения соответствующих случайных величин.

Число всевозможных типов распределений, разумеется, неограничено. Однако на практике далеко не все распределения встречаются одинаково часто. Анализ различных случайных величин, как изучаемых теоретически, так и вычисляемых на основании опытов, показывает существование одного наиболее часто встречающегося распределения, называемого *нормальным*.

Совокупность свойств случайной величины, по которым можно найти ее распределение, назовем *определяющими свойствами*. Эти свойства обычно проверяются намного легче, чем непосредственно само распределение. Теоремы, устанавливающие, какие свойства определяют то или иное распределение, составляют важный раздел теории вероятностей. Оказывается, что многие весьма простые и часто встречающиеся свойства случайных величин определяют именно нормальное распределение.

Приведем пример свойств, определяющих нормальное распределение; по этому примеру можно будет судить, насколько распространенными бывают такие свойства. Пусть координаты случайной точки на плоскости являются независимыми случайными величинами, т. е. имеют совершенно самостоятельные и независимые друг от друга функции распределения. Эту же точку можно рассматривать в какой-либо другой системе координат. Новые координаты снова будут случайными величинами, но они уже могут не быть

независимыми. Оказывается, достаточно, чтобы существовала еще одна система с тем же началом координат, в которой обе координаты случайной точки независимы — и распределение любой из упомянутых координат уже не может быть никаким другим, кроме нормального.

Доказательство сформулированного утверждения слишком сложно. Мы рассмотрим более частный случай, когда случайная точка имеет одинаковое распределение вдоль любой оси, проходящей через заданное начало координат. В этом случае все системы координат равноправны, поэтому в любой из них координаты независимы и имеют одну и ту же функцию распределения. Для такого распределения имеется хорошая физическая интерпретация — рассеяние снарядов при стрельбе из артиллерийского орудия. Действительно, все основные несимметричные факторы, действующие на снаряд (ветер, сила тяжести и т. п.), учитываются при наводке орудия. Остальные же факторы, являющиеся случайными и вызывающие рассеяние снарядов, действуют во всех направлениях с равной вероятностью. К этим факторам относятся, например, колебания воздуха, неодинаковость снарядов. Если точку наводки принять за начало координат, а точки попадания снарядов считать случайными, то распределение таких точек как раз и будет удовлетворять всем необходимым условиям (примерная картина попаданий дана на рис. 10).

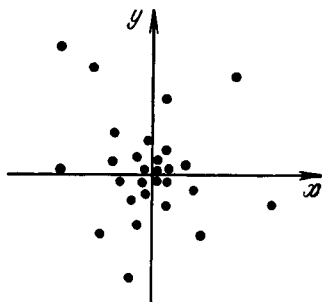


Рис. 10.

Итак, мы воспользуемся примером рассеяния снарядов, чтобы вывести основные формулы нормального распределения. Рассмотрим на плоскости некоторую точку $M(x, y)$ и какую-либо площадку вокруг нее. Если площадка невелика, то вероятность попадания в нее приблизительно (с точностью до бесконечно малых более высокого порядка) пропорциональна площади Δs этой площадки. Следовательно, для различных площадок, окружающих точку M и имеющих одинаковую площадь Δs , вероятности попадания

«почти» равны — они могут отличаться лишь на такую величину α , что

$$\lim_{\Delta s \rightarrow 0} \frac{\alpha}{\Delta s} = 0.$$

Возьмем теперь в качестве площадки прямоугольник с центром в точке M и сторонами, параллельными осям координат (рис. 11). Вероятность попадания в такой прямоугольник обозначим через P . Это попадание можно рассматривать как совместное осуществление двух событий: попадание в вертикальную полосу шириной Δx (см. рис. 11) и попадание в горизонтальную полосу шириной Δy . В силу независимости координат оба эти события не-

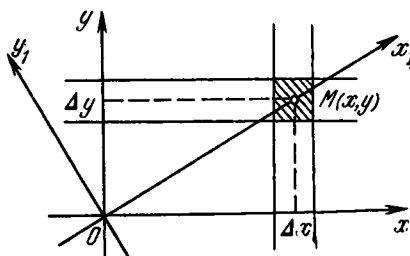


Рис. 11.

зависимы, поэтому, найдя их вероятности и перемножив, мы и получим P .

Попадание снаряда в вертикальную полосу равносильно тому, что абсцисса точки попадания окажется в отрезке $\left[x - \frac{1}{2} \Delta x, x + \frac{1}{2} \Delta x \right]$; вероятность такого события есть

$$P(\Delta x) = \int_{x - \frac{1}{2} \Delta x}^{x + \frac{1}{2} \Delta x} f(x) dx \approx f(x) \Delta x,$$

где $f(x)$ — плотность распределения координаты x . У координаты y должно быть то же самое распределение, поэтому вероятность попадания в горизонтальную полосу равна

$$P(\Delta y) = \int_{y - \frac{1}{2} \Delta y}^{y + \frac{1}{2} \Delta y} f(y) dy \approx f(y) \Delta y.$$

Оба равенства здесь приближенные, но погрешности при этом имеют более высокий порядок малости, чем Δx и Δy .

Таким образом, вероятность попадания в прямоугольник будет

$$P = P(\Delta x) P(\Delta y) \approx f(x) f(y) \Delta s,$$

где $\Delta s = \Delta x \Delta y$ есть площадь прямоугольника.

Все проведенные рассуждения можно повторить для любой другой системы Ox_1y_1 . Вероятность попадания в новый прямоугольник будет равна

$$P_1 \approx f(x_1) f(y_1) \Delta s_1,$$

где x_1, y_1 — новые координаты точки M , а Δs_1 — площадь нового прямоугольника. Если площади Δs и Δs_1 выбрать одинаковыми, то, как уже указывалось выше,

$$f(x) f(y) \Delta s = f(x_1) f(y_1) \Delta s_1 + \alpha$$

(в добавку α мы включили и все прочие погрешности, поэтому равенство стало точным). Деля обе части этого равенства на $\Delta s = \Delta s_1$, получим:

$$f(x) f(y) = f(x_1) f(y_1) + \frac{\alpha}{\Delta s}.$$

Устремляя теперь Δs к нулю и учитывая, что $\frac{\alpha}{\Delta s}$ при этом также стремится к нулю, придем к основному равенству

$$f(x) f(y) = f(x_1) f(y_1).$$

Если выбрать новые оси координат так, чтобы ось Ox_1 проходила через точку M , то новые координаты выразятся через старые формулами:

$$x_1 = \sqrt{x^2 + y^2}, \quad y_1 = 0.$$

В этом случае для функции получится соотношение

$$f(x) f(y) = f(\sqrt{x^2 + y^2}) f(0).$$

Введем обозначения

$$u = \sqrt{x^2 + y^2}, \quad k = f(0),$$

тогда

$$f(x) f(y) = kf(u).$$

От этого равенства можно перейти к некоторому дифференциальному уравнению, позволяющему найти функцию

f в явном виде. Для этого продифференцируем обе его части по x :

$$\frac{df(x)}{dx} f(y) = k \frac{df(u)}{dx} = k \frac{df(u)}{du} \frac{\partial u}{\partial x}.$$

Аналогично, дифференцируя обе части по y , получим

$$f(x) \frac{df(y)}{dy} = k \frac{df(u)}{dy} = k \frac{df(u)}{du} \frac{\partial u}{\partial y}.$$

Из обоих равенств можно найти $k \frac{df(u)}{du}$:

$$k \frac{df(u)}{du} = f(y) \frac{df(x)}{dx} : \frac{\partial u}{\partial x} = f(x) \frac{df(y)}{dy} : \frac{\partial u}{\partial y}.$$

Но

$$\frac{\partial u}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial u}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}.$$

Поэтому

$$f(y) \frac{df(x)}{dx} \frac{\sqrt{x^2 + y^2}}{x} = f(x) \frac{df(y)}{dy} \frac{\sqrt{x^2 + y^2}}{y}.$$

Отсюда уже нетрудно получить равенство

$$\frac{1}{xf(x)} \frac{df(x)}{dx} = \frac{1}{yf(y)} \frac{df(y)}{dy}.$$

В полученном равенстве левая часть не зависит от y , правая часть не зависит от x . Но поскольку они все-таки равны друг другу, то значит, они обе не зависят ни от x , ни от y , т. е. равны какому-то постоянному числу b . В частности,

$$\frac{1}{xf(x)} \frac{df(x)}{dx} = b.$$

Мы получили простое дифференциальное уравнение первого порядка, из которого и найдем $f(x)$. Для этого запишем уравнение в виде

$$\frac{df(x)}{f(x)} = bx dx,$$

откуда после интегрирования получим

$$\ln f(x) = \frac{bx^2}{2} + \ln C, \quad f(x) = Ce^{bx^2/2},$$

Уточним смысл постоянных b и C , исходя из того, что для $f(x)$, как и для всякой плотности,

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

Отсюда сразу же видно, что b не может быть положительным (не будет сходиться интеграл). Поэтому можно положить $b = -h^2$. Это же равенство позволяет выразить C через h . Действительно,

$$\int_{-\infty}^{\infty} C e^{-h^2 x^2 / 2} dx = 1,$$

откуда

$$C = \frac{1}{\int_{-\infty}^{\infty} e^{-h^2 x^2 / 2} dx} = \frac{h}{\sqrt{2\pi}}$$

(стоящий в знаменателе интеграл вычисляется с помощью сложных методов; его значение можно найти в различных специальных таблицах определенных интегралов).

Итак, мы получили, что плотность распределения точек попадания снарядов вдоль любого направления определяется формулой

$$f(x) = \frac{h}{\sqrt{2\pi}} e^{-h^2 x^2 / 2}.$$

Здесь остается неопределенной постоянная h ; она зависит от условий стрельбы и называется *точностью* стрельбы. Нетрудно видеть, что чем больше точность, тем кучнее, ближе к цели ложатся снаряды.

Если бы цель (точка наводки) находилась не в начале координат, а была смещена в направлении оси Ox на расстояние a , то для приведения к уже рассмотренному случаю нужно сделать замену x на $x - a$. Это даст плотность рассеяния снарядов вдоль оси Ox в общем случае:

$$f(x) = \frac{h}{\sqrt{2\pi}} e^{-h^2 (x-a)^2 / 2}.$$

Полученная функция и называется *плотностью нормального распределения*.

3.2. Свойства нормального распределения. Нормальное распределение играет основную роль в дальнейшем изложении. Поэтому изучим предварительно его свойства.

Случайная величина ξ называется распределенной по *нормальному закону*, если ее функция распределения имеет вид

$$F(x) = \frac{h}{\sqrt{2\pi}} \int_{-\infty}^x e^{-h^2(x-a)^2/2} dx.$$

Такая величина, очевидно, является непрерывной.

Найдем математическое ожидание и дисперсию нормально распределенной случайной величины. Для математического ожидания получим

$$M\xi = \frac{h}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-h^2(x-a)^2/2} dx.$$

Введем новую переменную $t = x - a$, тогда $dx = dt$, а пределы интегрирования не изменятся. Мы получим:

$$\begin{aligned} M\xi &= \frac{h}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (t+a) e^{-h^2 t^2/2} dt = \\ &= \frac{h}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{-h^2 t^2/2} dt + \frac{ha}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-h^2 t^2/2} dt. \end{aligned}$$

Первый из получившихся интегралов равен нулю, так как подынтегральная функция нечетна, а пределы интегрирования симметричны относительно нуля. Второй интеграл уже встречался в предыдущем пункте, он равен $\frac{\sqrt{2\pi}}{h}$. Поэтому окончательно

$$M\xi = a.$$

Применительно к примеру предыдущего пункта мы получаем, что точка наводки (цель) есть математическое ожидание случайных попаданий снаряда — результат, еще раз подтверждающий смысл математического ожидания как среднего, наивероятнейшего значения случайной величины.

Перейдем к вычислению дисперсии. Согласно общей формуле

$$D\xi = \frac{h}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-h^2(x-a)^2/2} dx.$$

Введем новую переменную $t=h(x-a)$, тогда $dx = \frac{1}{h} dt$, а пределы интегрирования не изменятся. Мы получим, что

$$D\xi = \frac{h}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{t^2}{h^2} e^{-t^2/2} \frac{1}{h} dt = \frac{1}{h^2 \sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} dt.$$

Полученный интеграл будем вычислять по частям, полагая $u=t$, $dv = te^{-t^2/2} dt$. Тогда $du=dt$,

$$v = \int te^{-t^2/2} dt = \int e^{-t^2/2} d\left(\frac{t^2}{2}\right) = -e^{-t^2/2}.$$

Поэтому

$$D\xi = \frac{1}{h^2 \sqrt{2\pi}} \left(-te^{-t^2/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-t^2/2} dt \right).$$

Выражение $-te^{-t^2/2}$ при подстановке ∞ и $-\infty$ вместо t обращается в нуль (точнее говоря, оно стремится к нулю при t , стремящемся к ∞ или $-\infty$). Оставшийся интеграл принадлежит к типу, уже неоднократно встречавшемуся раньше; из таблиц находим, что он равен $\sqrt{2\pi}$. Это приводит нас к окончательной формуле для дисперсии:

$$D\xi = \frac{1}{h^2}.$$

Для примера с рассеянием снарядов дисперсия обратно пропорциональна квадрату точности. Чем выше точность, тем меньше дисперсия, следовательно, и здесь дисперсия может служить мерой рассеяния.

Среднее квадратичное отклонение нормально распределенной величины называется *стандартом* и обозначается буквой σ . Отсюда $D\xi = \sigma^2$ и, следовательно, $\sigma = \frac{1}{h}$. С помощью стандарта плотность нормального распределения

запишется в виде

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2},$$

а функция распределения

$$F(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^x e^{-(x-a)^2/2\sigma^2} dx.$$

Именно этими формулами мы и будем пользоваться в дальнейшем.

График плотности нормального распределения называется *нормальной кривой* или *кривой Гаусса*. Он зависит от значений a и σ ; на рис. 12 представлена нормальная кривая при $a=2$, $\sigma=0,5$.

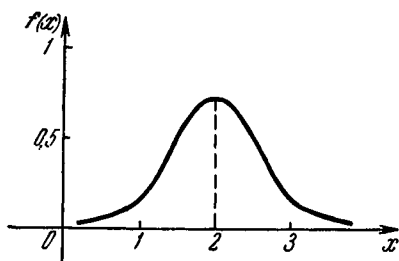


Рис. 12.

Нормальная кривая симметрична относительно прямой $x=a$ и при $x \rightarrow \pm \infty$ неограниченно приближается к оси абсцисс (эта ось является асимптотой кривой). При $x=a$ кривая имеет максимум, равный

$$\frac{1}{\sigma \sqrt{2\pi}}.$$

Если вокруг каждой точки оси абсцисс взять небольшую окрестность заданного радиуса ϵ *), то вероятность попадания в такой интервал пропорциональна значению плотности в рассматриваемой точке (с точностью до бесконечно малых более высокого порядка, чем ϵ). Непосредственно из графика видно, что для нормального распределения наиболее вероятно попадание в окрестность математического ожидания, а в обе стороны от него эта вероятность монотонно убывает. Скорость убывания при этом зависит от крутизны нормальной кривой, которая в свою очередь зависит от σ . На рис. 13 приведены для сравнения три нормальные кривые при $a=0$ и различных σ . Мы видим, что чем меньше σ ,

*) Окрестностью радиуса ϵ для какой-либо точки x называется интервал длины 2ϵ с центром в этой точке.

тем сильней убывает вероятность, т. е. тем плотней вокруг математического ожидания расположены значения случайной величины, обладающие достаточно заметной вероятностью.

В п. 2.2 указывалось, что с помощью плотности можно найти вероятность попадания случайной величины ξ в

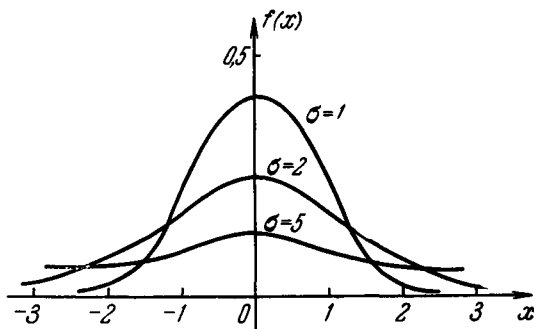


Рис. 13.

любой заданный интервал (x_1, x_2) . Для нормального распределения получается формула

$$P\{x_1 \leq \xi \leq x_2\} = \frac{1}{\sigma \sqrt{2\pi}} \int_{x_1}^{x_2} e^{-(x-a)^2/2\sigma^2} dx.$$

К сожалению, получившийся интеграл нельзя вычислить с помощью простых приемов интегрирования. Существуют более сложные методы, позволяющие вычислить его с достаточной степенью точности. Однако нет смысла каждый раз самостоятельно проводить эти вычисления — лучше воспользоваться готовыми таблицами, уже составленными для нормального распределения. В таблицах указываются обычно значения функции распределения

$$F(x) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^x e^{-(x-a)^2/2\sigma^2} dx.$$

По этим значениям нетрудно вычислить

$$P\{x_1 \leq \xi \leq x_2\} = F(x_2) - F(x_1).$$

Табулирование функции $F(x)$ наталкивается на одну трудность — для каждого конкретного значения a и σ нужно составлять свою таблицу. Этой трудности удастся избежать,

приводя все нормальные распределения к такому распределению, у которого $a=0$, $\sigma=1$.

Нормальное распределение с математическим ожиданием, равным нулю, и стандартом, равным единице, называется

стандартным. Его функция распределения имеет вид

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx.$$

График этой функции представлен на рис. 14. В некоторых статистических таблицах даются значения $F_0(x)$ при изменении x от $-\infty$ до ∞ . Однако такую таблицу можно сократить еще вдвое, используя следующее соображение. Сама функция $F_0(x)$ не обладает четностью или нечетностью, но функция $\Phi(x) = F_0(x) - \frac{1}{2}$ является нечетной, т. е. $\Phi(-x) = -\Phi(x)$. Поэтому таблицу значений функции $\Phi(x)$ достаточно составить лишь для $x \geq 0$. А восстановить $F_0(x)$ по $\Phi(x)$ уже не представляет труда. Более того, мы постараемся все встречающиеся в дальнейшем формулы преобразовывать таким образом, чтобы в них вообще встречалась только функция $\Phi(x)$.

Функция $\Phi(x)$ называется *функцией Лапласа*. Используя равенство

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-x^2/2} dx = F_0(0) = \frac{1}{2},$$

ее можно представить в виде

$$\Phi(x) = F_0(x) - F_0(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-x^2/2} dx - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-x^2/2} dx,$$

или, окончательно,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-x^2/2} dx.$$

Таблица значений этой функции приводится в Приложении к настоящей книге (таблица I).

Случайная величина ξ_0 , имеющая стандартное распределение, называется *нормированной*. Для такой величины

$$P\{x_1 \leq \xi_0 \leq x_2\} = F_0(x_2) - F_0(x_1).$$

Но

$$F_0(x_1) = \Phi(x_1) + \frac{1}{2}, \quad F_0(x_2) = \Phi(x_2) + \frac{1}{2}.$$

Поэтому

$$P\{x_1 \leq \xi_0 \leq x_2\} = \Phi(x_2) - \Phi(x_1). \quad (3.1)$$

Если какое-нибудь из заданных значений x_1, x_2 отрицательно, то для него функцию Лапласа вычисляют с использованием свойства нечетности.

Пример 1. Найти вероятность того, что нормированная случайная величина примет значение между -1 и $\frac{1}{2}$.

Согласно общей формуле (3.1),

$$P\left\{-1 \leq \xi_0 \leq \frac{1}{2}\right\} = \Phi\left(\frac{1}{2}\right) - \Phi(-1).$$

По таблице I Приложения находим, что $\Phi(0,5) = 0,1915$. Значение -1 в таблице отсутствует, но мы знаем, что $\Phi(-1) = -\Phi(1)$. Поэтому $\Phi(-1) = -0,3413$. В результате

$$P\left\{1 \leq \xi_0 \leq \frac{1}{2}\right\} = 0,1915 + 0,3413 = 0,5328.$$

Вычисления еще более упрощаются, если числа x_1 и x_2 симметричны относительно математического ожидания a . С такой ситуацией приходится сталкиваться в задаче об *абсолютном отклонении*. Абсолютным отклонением называется разность между случайной величиной и ее математическим ожиданием, взятая по абсолютной величине:

$$\Delta\xi = |\xi - M\xi|.$$

Задача об абсолютном отклонении звучит так: *найти вероятность того, что абсолютное отклонение случайной величины не превзойдет некоторого заданного числа ε* . Эта задача сводится к уже рассмотренной, так как неравенство $\Delta\xi \leq \varepsilon$ равносильно неравенствам $M\xi - \varepsilon \leq \xi \leq M\xi + \varepsilon$.

Для нормального распределения $M\xi = a$ и, значит,

$$P\{\Delta\xi \leq \varepsilon\} = P\{a - \varepsilon \leq \xi \leq a + \varepsilon\}.$$

В частности, для нормированной случайной величины

$$P\{\Delta\xi_0 \leq \varepsilon\} = P\{-\varepsilon \leq \xi_0 \leq \varepsilon\} = \Phi(\varepsilon) - \Phi(-\varepsilon).$$

Учитывая, что $\Phi(-\varepsilon) = -\Phi(\varepsilon)$, получим очень важную для приложений формулу

$$P\{\Delta\xi_0 \leq \varepsilon\} = 2\Phi(\varepsilon).$$

Пример 2. Точность стрельбы $h=1$. Найти вероятность попадания в мишень радиуса 2 (все данные в метрах).

Как уже было доказано, рассеяние снарядов подчиняется нормальному закону распределения со стандартом $\sigma = \frac{1}{h} = 1$.

Значит, в нашем примере распределение точек попадания стандартное. Обозначим расстояние от точки попадания снаряда до центра мишени через $\Delta\xi_0$. Тогда задача сведется к нахождению вероятности $P\{\Delta\xi_0 \leq 2\}$. Используя таблицы, находим $P\{\Delta\xi_0 \leq 2\} = 2\Phi(2) = 0,9544$.

Перейдем теперь к рассмотрению общего нормального распределения с произвольными параметрами a и σ . Преобразуем заданную случайную величину по формуле $\xi_0 = \frac{\xi - a}{\sigma}$ (такое преобразование называется *нормировкой*). Случайная величина ξ_0 будет нормированной. Действительно, используя свойства математического ожидания и дисперсии (см. п. 2.3), найдем

$$M\xi_0 = M\left(\frac{\xi - a}{\sigma}\right) = \frac{1}{\sigma}(M\xi - a) = \frac{1}{\sigma}(a - a) = 0.$$

Аналогично,

$$D\xi_0 = D\left(\frac{\xi - a}{\sigma}\right) = \frac{1}{\sigma^2} D(\xi - a) = \frac{1}{\sigma^2} D\xi = \frac{\sigma^2}{\sigma^2} = 1,$$

откуда вытекает, что и стандарт для ξ_0 равен 1. Таким образом, для полученной после нормировки величины ξ_0 должны быть справедливы все найденные выше формулы.

Допустим вначале, что нам нужно найти вероятность $P\{x_1 \leq \xi \leq x_2\}$. Если ξ изменяется от x_1 до x_2 , то $\xi_0 = \frac{\xi - a}{\sigma}$ изменяется в это время от $\frac{x_1 - a}{\sigma}$ до $\frac{x_2 - a}{\sigma}$. Следовательно,

$$P\{x_1 \leq \xi \leq x_2\} = P\left\{\frac{x_1 - a}{\sigma} \leq \xi_0 \leq \frac{x_2 - a}{\sigma}\right\}.$$

Отсюда и из (3.1) немедленно вытекает основная формула:

$$P\{x_1 \leq \xi \leq x_2\} = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right). \quad (3.2)$$

Если же решается задача об абсолютном отклонении, то

$$P\{\Delta\xi \leq \varepsilon\} = P\left\{\Delta\xi_0 \leq \frac{\varepsilon}{\sigma}\right\} = 2\Phi\left(\frac{\varepsilon}{\sigma}\right).$$

Пример 3. Случайная величина ξ имеет нормальное распределение с параметрами $a=20$ и $\sigma=10$. Какова вероятность того, что она примет значение, лежащее между 15 и 40?

Используя формулу (3.2) и таблицу I Приложения, находим

$$\begin{aligned} P\{15 \leq \xi \leq 40\} &= \Phi\left(\frac{40-20}{10}\right) - \Phi\left(\frac{15-20}{10}\right) = \\ &= \Phi(2) + \Phi(0,5) = 0,4772 + 0,1915 = 0,6687. \end{aligned}$$

Пример 4. Детали, изготавливаемые на станке, в силу различных случайных причин отличаются по своему диаметру. Удалось установить, что диаметр d есть нормально распределенная случайная величина со стандартом $\sigma=2$ мм. Какова вероятность брака, если бракуются детали, диаметр которых отклоняется от нормы (математического ожидания) более, чем на 3,5 мм?

По условиям задачи требуется найти вероятность неравенства $\Delta d > 3,5$. Мы такого неравенства не рассматривали, поэтому перейдем к противоположному событию, т. е. неравенству $\Delta d \leq 3,5$. Полагая $\xi=d$ и учитывая, что $\varepsilon=3,5$,

найдем

$$P \{ \Delta d \leq 3,5 \} = 2\Phi \left(\frac{3,5}{2} \right) = 2\Phi (1,75) = 0,9198.$$

Значит,

$$P \{ \Delta d > 3,5 \} = 1 - P \{ \Delta d \leq 3,5 \} = 1 - 0,9198 = 0,0802.$$

В задаче об абсолютном отклонении математическое ожидание случайной величины не играет никакой роли — достаточно знать лишь стандарт σ (или дисперсию σ^2). Обозначая $\frac{\varepsilon}{\sigma}$ через k , получим соотношение

$$P \{ \Delta \xi \leq k\sigma \} = 2\Phi (k).$$

Если k одно и то же, то и вероятность $P \{ \Delta \xi \leq k\sigma \}$ не меняется, какую бы случайную величину ξ с нормальным распределением мы ни рассматривали. Мы приходим к следующему важному правилу: *вероятность того, что абсолютное отклонение нормально распределенной случайной величины не превзойдет некоторого предела, зависит только от того, во сколько раз этот предел больше стандарта рассматриваемой случайной величины.* Отсюда, кстати, ясно и происхождение названия «стандарт» — это величина, с которой сравниваются все отклонения.

Для некоторых k полезно соответствующие значения $\Phi(k)$ запомнить — это поможет мгновенно оценивать различные результаты измерений. В частности, нужно помнить следующие три формулы:

$$P \{ \Delta \xi \leq \sigma \} = 2\Phi (1) = 0,6826,$$

$$P \{ \Delta \xi \leq 2\sigma \} = 2\Phi (2) = 0,9544,$$

$$P \{ \Delta \xi \leq 3\sigma \} = 2\Phi (3) = 0,9973.$$

Если брать приближенные значения вероятностей, то получатся удобные для запоминания правила:

- 1) вероятность стандартного отклонения равна $\frac{2}{3}$ (правило сигмы);
- 2) вероятность удвоенного стандарта равна 95% (правило двух сигм);
- 3) вероятность утроенного стандарта равна 1 (правило трех сигм).

Считая, что вероятность 1 могут иметь только достоверные события, последнему правилу можно придать другую

формулировку: *отклонения, большие чем утроенный стандарт, практически невозможны* *).

Пример 5. Какова должна быть точность стрельбы h , чтобы все снаряды попали в мишень диаметром 2,4 м?

Предел абсолютного отклонения здесь равен 1,2 (радиус мишени). Значит, согласно правилу трех сигм, стандарт должен быть не больше, чем $\frac{1,2}{3} = 0,4$. А тогда $h \geq \frac{1}{0,4} = 2,5$, т. е. точность должна быть не меньше 2,5.

В заключение пункта отметим (без доказательства), что нормальное распределение обладает важным свойством *линейности*: если независимые случайные величины ξ_1 и ξ_2 имеют нормальные распределения, то для произвольных чисел α , β величина

$$\eta = \alpha\xi_1 + \beta\xi_2$$

также имеет нормальное распределение. При этом из свойств математического ожидания и дисперсии независимых случайных величин (см. п. 2.3) непосредственно вытекает, что параметры a и σ распределения величины η выражаются через параметры a_1 и σ_1 распределения величины ξ_1 и параметры a_2 и σ_2 распределения величины ξ_2 формулами

$$a = \alpha a_1 + \beta a_2, \quad \sigma = \sqrt{\alpha^2 \sigma_1^2 + \beta^2 \sigma_2^2}.$$

3.3. Использование нормального распределения для изучения других случайных величин. Нормальное распределение является наиболее изученным распределением. Поэтому его стараются использовать и при изучении случайных величин, распределения которых отличны от нормального. Здесь могут быть два основных пути.

Первый путь заключается в переходе от заданной величины к другой, имеющей нормальное распределение, по определенной формуле, которую впоследствии можно будет учесть. Например, при изучении свойств случайной величины может оказаться, что нормальное распределение имеет ее логарифм (распределение самой случайной величины ξ называется в этом случае *логарифмически нормальным*). Тогда вместо ξ следует рассматривать случайную величину

*) Суждения подобного типа более подробно изучаются ниже, в п. 5.3.

$\eta = \lg \xi$, пересчитав все исходные данные применительно к новой величине. Получив с помощью формул нормального распределения все необходимые результаты для η , мы затем можем снова вернуться к исходной случайной величине ξ .

Другой путь заключается в том, чтобы распределение заданной случайной величины заменить приближенно нормальным (если это возможно). Второй путь особенно часто применяется при обработке экспериментальных данных, где обычно нет возможности установить распределение случайной величины с абсолютной точностью. Возможность приближенной замены одного распределения другим основана на следующем соображении. Вычисление вероятности на практике никогда не бывает самоцелью. Вероятность того или иного события нужна нам обычно лишь для того, чтобы в соответствии с ней выбрать дальнейший план действий: принять или не принять гипотезу, удовлетвориться ли результатом, учитывать ли тот или иной фактор и т. п. В такого рода «психологическом» выборе небольшие различия между вероятностями не будут, конечно, играть роли. Следовательно, и функцию распределения незачем знать с особо высокой точностью.

Для того чтобы одну случайную величину можно было приближенно заменить другой, нужно, чтобы достаточно мало отличались их функции распределения. Если случайные величины непрерывны, то можно сравнивать их плотности. Разумеется, в каждом конкретном случае нужно уметь проверять, что погрешности перехода действительно малы.

Приближенная замена одних распределений другими особенно эффективна при замене конечнозначных и дискретных величин непрерывными, так как аппарат интегрирования намного проще в работе, чем аппарат суммирования. В этом случае, однако, можно сравнивать лишь функции распределения случайных величин, ибо у дискретных и конечнозначных распределений нет плотности.

Последнее обстоятельство несколько огорчительно, так как сравнение плотностей удобнее непосредственного сравнения самих функций распределения. Оказывается роль плотности конечнозначной (или дискретной) величины, в некотором смысле, играет ряд распределения этой величи-

ны, т. е. совокупность ее значений и соответствующих вероятностей. Покажем это на примере конечнозначной случайной величины ξ , принимающей только целочисленные значения $k=0, 1, 2, \dots, n$; вероятности этих значений обозначим через p_k .

Рассмотрим совокупность интервалов $\left[k - \frac{1}{2}, k + \frac{1}{2}\right]$ для всех k . Эти интервалы, не перекрываясь, заполняют промежуток от $-\frac{1}{2}$ до $n + \frac{1}{2}$, и числа $k=0, 1, 2, \dots, n$ будут

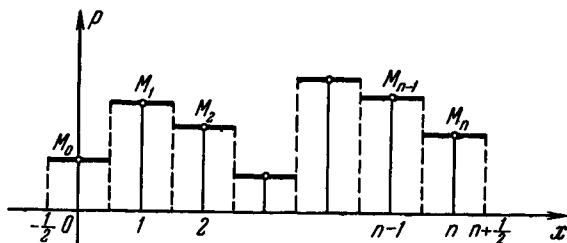


Рис. 15.

серединами интервалов. Случайная величина ξ может принимать только одно значение из каждого такого интервала (как раз середину) с вероятностью p_k для k -го интервала. Поэтому можно сказать, что для случайной величины ξ вероятность попадания в интервал $\left[k - \frac{1}{2}, k + \frac{1}{2}\right]$ равна p_k .

Для того чтобы «превратить» конечнозначную случайную величину ξ в непрерывную величину η , нужно как бы «размазать» каждое значение величины ξ по всему окружающему его интервалу; при этом желательно, чтобы вероятности попаданий в эти интервалы не изменились, или изменились незначительно. Подобное «размазывание» можно осуществлять самыми различными способами, из которых наиболее употребительны способы гистограммы и полигона.

Отложим значения $k=0, 1, 2, \dots, n$ на оси абсцисс и в каждой точке восставим перпендикуляр соответствующей длины p_k (см. рис. 15). Мы получим точки M_0, M_1, \dots, M_n ; совокупность этих точек называется *огивой* случайной величины ξ . Распространим теперь каждое значение p_k на

весь интервал $\left[k - \frac{1}{2}, k + \frac{1}{2}\right]$. Геометрически это равносильно тому, что через каждую точку M_k проводится горизонтальный отрезок длины 1. Получится ступенчатый график на промежутке от $-\frac{1}{2}$ до $n + \frac{1}{2}$; этот график называется *гистограммой* случайной величины ξ . Наконец, на рис. 16 изображен *полигон* случайной величины ξ ; он получится,

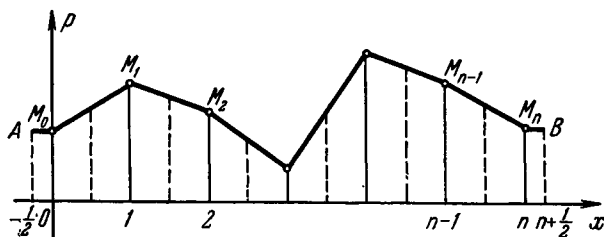


Рис. 16.

если соединить отрезком прямой каждые две соседние точки огины. Точки M_0 и M_n соединяются при этом с точками $A\left(-\frac{1}{2}, p_0\right)$ и $B\left(n + \frac{1}{2}, p_n\right)$.

Если точкам огины соответствуют лишь целочисленные значения $k=0, 1, 2, \dots, n$ на оси абсцисс, то значения, соответствующие точкам гистограммы или полигона, уже сплошь заполняют весь промежуток от $-\frac{1}{2}$ до $n + \frac{1}{2}$. Следовательно, гистограмму и полигон можно рассматривать как графики некоторых функций $f_1(x)$ и $f_2(x)$, определенных на интервале $\left[-\frac{1}{2}, n + \frac{1}{2}\right]$; вне этого интервала их можно считать равными нулю. Тогда

$$\left. \begin{aligned} \int_{-\infty}^{\infty} f_1(x) dx &= \int_{-1/2}^{n+1/2} f_1(x) dx, \\ \int_{-\infty}^{\infty} f_2(x) dx &= \int_{-1/2}^{n+1/2} f_2(x) dx. \end{aligned} \right\} \quad (3.3)$$

Но интеграл от функции по некоторому промежутку равен площади фигуры, заключенной между графиком функции и

осью абсцисс, а справа и слева ограниченной пределами интегрирования. Поэтому, чтобы вычислить интегралы (3.3) от функций $f_1(x)$ и $f_2(x)$, нужно вычислить площади фигур, заключенных между гистограммой (соответственно полигоном) и осью абсцисс.

В случае гистограммы мы имеем совокупность прямоугольников с основаниями 1 и высотами p_k , значит, и площади их равны p_k . А тогда

$$\int_{-1/2}^{n+1/2} f_1(x) dx = p_0 + p_1 + \dots + p_n = 1.$$

В случае полигона мы имеем совокупность трапеций с основаниями p_k и p_{k+1} и высотами 1. Площадь каждой такой трапеции равна $\frac{p_k + p_{k+1}}{2}$. Кроме того, здесь участвуют еще два прямоугольника с основаниями $\frac{1}{2}$ и высотами p_0 и p_n . Окончательно получим, что

$$\begin{aligned} \int_{-1/2}^{n+1/2} f_2(x) dx &= \frac{p_0}{2} + \frac{p_0 + p_1}{2} + \frac{p_1 + p_2}{2} + \dots + \frac{p_{n-1} + p_n}{2} + \frac{p_n}{2} = \\ &= p_0 + p_1 + \dots + p_n = 1. \end{aligned}$$

Итак, обе функции $f_1(x)$ и $f_2(x)$ удовлетворяют равенствам

$$\int_{-\infty}^{\infty} f_1(x) dx = 1, \quad \int_{-\infty}^{\infty} f_2(x) dx = 1$$

и, кроме того, они неотрицательны. Значит, их можно считать плотностями некоторых непрерывных случайных величин η_1 и η_2 . Нетрудно проверить, что функции распределения случайных величин η_1 и η_2 мало отличаются от функции распределения величины ξ и отличие это тем слабее, чем меньше сами вероятности p_k . Поэтому в практических расчетах величину ξ можно приближенно заменять непрерывными величинами η_1 и η_2 . При этом необходимо учитывать, что каждое значение k величины ξ заменится целым интервалом $\left[k - \frac{1}{2}, k + \frac{1}{2}\right]$. Например, неравенство $k_1 \leq \xi \leq k_2$ равносильно тому, что величина η_1 (или η_2) попадет в один

из интервалов $\left[k_1 - \frac{1}{2}, k_1 + \frac{1}{2}\right]$, $\left[k_1 + \frac{1}{2}, k_1 + \frac{3}{2}\right]$, ..., ..., $\left[k_2 - \frac{1}{2}, k_2 + \frac{1}{2}\right]$, т. е. равносильно неравенству $k_1 - \frac{1}{2} \leq \eta_1 \leq k_2 + \frac{1}{2}$ (или $k_1 - \frac{1}{2} \leq \eta_2 \leq k_2 + \frac{1}{2}$). Следовательно, в приближенных расчетах можно пользоваться формулами

$$P\{k_1 \leq \xi \leq k_2\} = \int_{k_1 - 1/2}^{k_2 + 1/2} f_1(x) dx = \int_{k_1 - 1/2}^{k_2 + 1/2} f_2(x) dx.$$

Плотности $f_1(x)$ и $f_2(x)$ обладают одним важным свойством: если x принимает целочисленное значение $k=0, 1, 2, \dots, n$, то

$$f_1(k) = f_2(k) = p_k.$$

Именно этот факт мы имели в виду, когда говорили, что ряд распределения конечнозначной величины ξ играет роль плотности — ведь у ξ нет других значений, кроме целочисленных.

Величины η_1 и η_2 , с точки зрения своей структуры, являются простейшими непрерывными случайными величинами, связанными с конечнозначной величиной ξ . Однако аналитические выражения для их плотностей $f_1(x)$ и $f_2(x)$ довольно громоздки, неудобны в работе. Для того чтобы их записать, нужно знать все вероятности p_k , что тоже не всегда возможно. Поэтому η_1 и η_2 стараются приближенно заменить какой-нибудь третьей непрерывной случайной величиной η с более «удобной» плотностью $f(x)$. Такую замену можно делать, если плотность $f(x)$ мало отличается от плотностей $f_1(x)$ и $f_2(x)$. В частности, для $k=0, 1, 2, \dots, n$ должны выполняться соотношения $f(k) \approx p_k$. Функцией $f(x)$ можно будет тогда пользоваться при оценке вероятностей неравенств:

$$P\{k_1 \leq \xi \leq k_2\} \approx \int_{k_1 - 1/2}^{k_2 + 1/2} f(x) dx. \quad (3.4)$$

Описанный способ перехода от конечнозначной случайной величины к непрерывной применим к любым конечнозначным и дискретным величинам, при этом будут лишь необходимы некоторые поправки, учитывающие, что раз-

ность между соседними значениями случайной величины может отличаться от единицы.

Особенно эффектный результат получается, если изложенные выше соображения применить к случайной величине, имеющей биномиальное распределение. Напомним, что биномиальным называлось распределение числа появлений события A , имеющего вероятность p , в n независимых испытаниях. Это число может быть только целым $k=0, 1, 2, \dots, n$; вероятность каждого значения уже была вычислена в п. 1.5:

$$p_k = \frac{n!}{k!(n-k)!} p^k q^{n-k} \quad (q = 1 - p).$$

При больших n и k вычисление факториалов вызывает серьезные затруднения. Соответственно усложняются и все другие задачи, связанные с биномиальным распределением, так что переход к непрерывной случайной величине был бы здесь весьма желательным.

Опуская довольно сложные и громоздкие выкладки, сформулируем получающийся здесь окончательный результат: *биномиальное распределение можно приближенно заменять нормальным распределением с математическим ожиданием $a=np$ и дисперсией $\sigma^2=npq$; погрешность при этом будет тем меньше, чем больше дисперсия npq* . В практических вычислениях погрешностью перехода можно пренебрегать уже при $npq \geq 9$. При заданном p это позволяет находить минимально допустимое число испытаний n , позволяющее биномиальное распределение заменять нормальным без всяких поправок. Например, для $p=0,1$ имеем

$$n \cdot 0,1(1-0,1) \geq 9, \quad n \geq 100.$$

С помощью нормального распределения нетрудно исследовать различные неравенства для числа осуществлений события A в n независимых испытаниях. Из общего равенства (3.4) вытекает формула

$$P\{k_1 \leq \xi \leq k_2\} = \frac{1}{\sqrt{2\pi npq}} \int_{k_1 - 1/2}^{k_2 + 1/2} e^{-(x - np)^2 / 2npq} dx.$$

Как и для всякого нормального распределения, здесь можно использовать функцию Лапласа, что приведет нас к

следующей, обычно и употребляемой на практике формуле:

$$P\{k_1 \leq \xi \leq k_2\} = \Phi\left(\frac{k_2 + \frac{1}{2} - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - \frac{1}{2} - np}{\sqrt{npq}}\right).$$

Пример. Найти вероятность того, что при 50 бросаниях монеты герб выпадет не менее 20 и не более 28 раз.

Здесь $n=50$, $p=\frac{1}{2}$, значит, и $q=\frac{1}{2}$. Находим $np=25$ и $npq=12,5$. Мы видим, что $npq \geq 9$, следовательно, для расчетов можно пользоваться нормальным распределением.

Непосредственно по формуле и таблице I Приложения находим

$$\begin{aligned} P\{20 \leq \xi \leq 28\} &= \Phi\left(\frac{28 + \frac{1}{2} - 25}{\sqrt{12,5}}\right) - \Phi\left(\frac{20 - \frac{1}{2} - 25}{\sqrt{12,5}}\right) = \\ &= \Phi(1,00) - \Phi(-1,57) = 0,3413 + 0,4418 = 0,7831. \end{aligned}$$

Неравенство $npq \geq 9$ показывает, что когда p или q близки к нулю, число n должно быть очень велико. Поэтому биномиальное распределение редко удается заменять нормальным при очень малых или очень больших (близких к единице) значениях p . Тем не менее, и в этом случае не обязательно вести расчеты по точным формулам. Можно показать, что для $p \leq 0,1$ и достаточно большого n биномиальное распределение мало отличается от распределения Пуассона (см. п. 2.2) с параметром $a=np$. В связи с этим распределение Пуассона называют иногда «законом редких событий». Если $p \geq 0,9$, то нужно перейти к рассмотрению противоположного события \bar{A} .

Напомним, что отношение числа появлений события A к числу всех испытаний называется частотой события A . Если число появлений события A как случайную величину обозначить через ξ , то частота ω также есть случайная величина, связанная с ξ соотношением $\omega = \frac{\xi}{n}$. В п. 1.5 была сформулирована теорема Бернулли, которая в принятых

нами сейчас обозначениях может быть записана так:

$$\lim_{n \rightarrow \infty} P \{ |\omega - p| > \varepsilon \} = 0,$$

где ε — любое фиксированное положительное число.

Используя связь между биномиальным и нормальным распределениями, мы можем теперь доказать эту теорему (здесь $n \rightarrow \infty$, поэтому ни о каких погрешностях перехода можно не думать). Действительно,

$$P \{ |\omega - p| > \varepsilon \} = 1 - P \{ |\omega - p| \leq \varepsilon \} = 1 - P \left\{ \left| \frac{\xi}{n} - p \right| \leq \varepsilon \right\}.$$

Неравенство $\left| \frac{\xi}{n} - p \right| \leq \varepsilon$ равносильно неравенству $|\xi - np| \leq n\varepsilon$ и, так как np есть математическое ожидание ξ , мы получаем задачу об абсолютном отклонении. Стандарт для ξ равен \sqrt{npq} , поэтому

$$\begin{aligned} P \{ |\xi - np| \leq n\varepsilon \} &= P \{ \Delta \xi \leq n\varepsilon \} = 2\Phi \left(\frac{n\varepsilon}{\sqrt{npq}} \right) = \\ &= 2\Phi \left(\frac{\varepsilon}{\sqrt{pq}} \sqrt{n} \right). \end{aligned}$$

Если $n \rightarrow \infty$, то и $\frac{\varepsilon}{\sqrt{pq}} \sqrt{n} \rightarrow \infty$. Но тогда

$$\Phi \left(\frac{\varepsilon}{\sqrt{pq}} \sqrt{n} \right) \rightarrow \frac{1}{2}$$

(в этом можно убедиться хотя бы из рис. 14). Поэтому

$$\lim_{n \rightarrow \infty} P \{ |\xi - np| \leq n\varepsilon \} = 1,$$

откуда и вытекает утверждение теоремы Бернулли.

§ 4. НАБЛЮДЕНИЯ

4.1. Наблюдение как этап исследования. В процессе изложения предыдущих параграфов мы накопили уже достаточный теоретический материал для непосредственного перехода к основному содержанию настоящей книги. Мы попытаемся применить методы теории вероятностей к практическому изучению различных явлений природы и к обработке получаемых при этом результатов.

Прежде всего уточним, какой смысл нужно придавать понятию испытания. В практических исследованиях *испытанием* называют чаще всего эксперимент, опыт, т. е. такое изучение явлений, при котором изучаемые факторы вызываются искусственно — создаются специальные приборы, установки и т. п. Однако искусственная установка — не единственный способ изучения природных процессов и явлений. Поэтому нам удобнее сохранить за понятием испытания как можно более широкий смысл, считая испытанием любое осуществление комплекса основных факторов, искусственное или протекающее в природных условиях.

К основным факторам мы будем относить все изучаемые факторы, а также факторы, служащие для стабилизации явления. Всякие побочные, посторонние факторы желательно по возможности устранять. Однако, как уже неоднократно упоминалось, устранить все такие факторы невозможно. Те из них, которые хотя и нельзя устранить, но зато можно с достаточной полнотой учесть, мы также будем относить к основным факторам. Основные факторы должны быть одинаковыми для всех испытаний, посвященных исследованию какого-либо одного свойства или признака (такую совокупность испытаний мы будем называть *серией испытаний*). Испытания с неизменным комплексом основных факторов называются *однородными*; однородность испытаний яв-

ляется одним из важнейших условий правильного применения статистических методов обработки наблюдений. Чтобы обеспечить однородность испытаний, нужно каждую их серию проводить в как можно более стабильных условиях: на одних и тех же приборах и установках, одними и теми же исследователями и, наконец, в предельно короткий срок, так как многие факторы заметно меняются во времени. Если стабильность испытаний обеспечить не удастся, то их изменение желательно учитывать и как особый фактор включать в число основных факторов.

Все прочие неустраняемые факторы, не поддающиеся или трудно поддающиеся учету, относят обычно к случайным факторам. Вообще говоря, вопрос о том, какие факторы нужно во что бы то ни стало учитывать, а какие можно считать случайными, приходится решать каждый раз применительно к конкретным условиям. Например, используя для измерений температуры в различных испытаниях один и тот же термометр, его погрешность учитывают как основной фактор. Если же пользуются в разных испытаниях разными термометрами, то соответствующие погрешности присоединяют к остальным случайным факторам.

Еще раз подчеркнем, однако, что деление факторов на основные и случайные весьма условно и может быть совершенно произвольным — лишь бы основные факторы не менялись от испытания к испытанию в пределах одной серии. Например, испытывая на прочность стальные бруски, мы можем их поперечники считать случайными (т. е. не учитывать), хотя такой учет может быть совсем не трудным.

Выделение тех или иных факторов в качестве основных влияет на вероятности возникающих случайных событий, на распределения соответствующих случайных величин. Изменится комплекс основных факторов — изменятся и эти распределения. Но не все распределения одинаково удобны для изучения. И если мы хотим в результате испытаний получить удобное, хорошо изученное распределение, то мы уже не можем выделять основные факторы совершенно произвольно. Именно получение случайных величин с «хорошими» распределениями и является главным критерием при выделении основных факторов.

При каждом испытании возникает бесчисленное множество событий, изучая частоту которых можно составить

представление об их вероятностях и в дальнейшем применять к их изучению вероятностные методы. Если изучается одно какое-нибудь случайное событие, то в результате отдельного испытания удастся выяснить только, осуществилось это событие или нет. Однако практическое проведение любого испытания является достаточно трудоемким делом. Поэтому одно и то же испытание используется для одновременного изучения нескольких случайных событий. Обычно это выглядит следующим образом: перед проведением испытаний выделяют целый класс случайных событий, а после каждого испытания регистрируют, какие случайные события из этого класса осуществились. Такая регистрация и называется *наблюдением*.

В качестве наблюдаемого класса случайных событий выбирают чаще всего совокупность свойств какого-нибудь явления или совокупность числовых значений какой-нибудь случайной величины. Например, при каждом астрономическом наблюдении за небесным светилом одновременно отмечают его координаты, яркость, цвет, время наблюдений и многое другое.

При наблюдениях приходится следить не только за результатами испытания, но и за правильностью его проведения. Поэтому признаки, отмечаемые при наблюдении, делятся на *изучаемые* и *контрольные*. Контрольные признаки служат для проверки однородности испытаний. Например, изучая сопротивление электролита при фиксированной температуре, мы при каждом наблюдении должны отмечать, что температура действительно не изменилась. Эти же контрольные признаки помогают учитывать изменение условий испытания, если таковое произойдет.

По своему характеру результаты, регистрируемые при наблюдении, делятся на *качественные* и *количественные*. К качественным результатам относятся появление какого-либо события (например, зажигание контрольной лампочки, выпадение осадка в растворе и т. п.), а также цвет, вкус, форма изучаемого объекта, короче, все результаты, не имеющие числового характера. Последнее обстоятельство, кстати, и не позволяет непосредственно применить к обработке качественных результатов математические методы. В связи с этим в экспериментальных исследованиях (особенно в последнее время) все более заметно стремление переходить от

качественных результатов к количественным, подбирая единицы измерения для таких качеств, как цвет, вкус, яркость и т. п. Развитие теории информации позволило получать количественные оценки даже в таких «далеких» от математики науках, как лингвистика или медицина.

Количественные результаты наиболее удобны для математической обработки. Источниками таких результатов служат в основном наблюдения двух видов: *подсчет* и *измерение*. С наблюдениями первого вида мы сталкиваемся, подсчитывая число вспышек на фотографии ядерного процесса, число дефектов на готовой детали, число зерен на шлифе; к подсчетам относятся почти все демографические наблюдения, экономические данные и т. п.

Измерения возникают тогда, когда наблюдаемое свойство сравнивается в количественном отношении с некоторым эталоном (единицей измерения). Для измерений служат всевозможные измерительные приборы: весы, мензурки, метрические линейки, электроизмерительные приборы и т. д. Результаты, полученные непосредственным измерением, в дальнейшем нередко приходится пересчитывать по различным формулам; после подсчета получаются так называемые *результаты косвенных измерений*. К измерениям (главным образом, косвенным) относится большинство методов анализа вещества, а также многочисленные исследования производственных процессов.

Каждое наблюдение производится для изучения некоторой причинно-следственной связи, изучаемый фактор при этом всегда включается в число основных факторов испытания. Кроме того, как уже говорилось выше, на результат наблюдения оказывают влияние и многочисленные примесные факторы. Если такие факторы включены в число основных, то их действие учитывается и, следовательно, не искажает результат наблюдения. Результат, который появился бы при воздействии одних только основных факторов испытания, называется *истинным результатом* (а при измерении — *истинным значением* измеряемой величины). Отыскание истинного результата и есть идеальная цель каждого исследования.

По-видимому, нет необходимости вновь убеждать читателя в том, что выделение основных факторов (т. е. факторов, которые можно учесть) в чистом виде невозможно,

На каждый результат оказывают воздействие всевозможные неучитываемые нами (сознательно и бессознательно) случайные факторы. Следовательно, *реальный результат наблюдения всегда является случайной величиной*. И если результат реального подсчета может иногда совпадать с истинным результатом, то при измерениях получить истинное значение измеряемой величины, как правило, невозможно. Связано это с тем, что результатам подсчета соответствует конечнозначная или дискретная величина, у которой каждое значение имеет ненулевую вероятность; результатам же измерения соответствует непрерывная случайная величина, у которой каждое отдельное значение (в том числе и истинное) имеет вероятность нуль.

Итак, каждый реальный результат отклоняется от истинного. Это отклонение называется *ошибкой наблюдения*. Ошибка наблюдения также есть случайная величина — фактически она является результатом действия только случайных (неучитываемых) факторов.

Получается довольно неотрадная картина: мы знаем, что каждое наблюдение содержит ошибку, но не знаем ее величины. Иными словами, мы должны делать «верное» заключение по «неверным» данным. Многие буржуазные философы даже видят в этом «противоречии» доказательство непознаваемости мира, в противовес одному из основных положений диалектического материализма; они считают, что выводы, получаемые на основе обработки наблюдений, не являются объективными, и зависят от личных качеств экспериментатора.

Разумеется, никакого противоречия с диалектическим материализмом здесь нет. Из того, что каждое наблюдение содержит некоторую случайную ошибку, следует лишь, что выводы, сделанные на основе наблюдений, не являются абсолютно достоверными. Однако эти выводы будут вполне объективными — существуют методы, позволяющие оценивать вероятности этих выводов *). Разработкой указанных методов и занимается математическая статистика.

*) Кстати сказать, положение о том, что вероятность есть некоторая объективная характеристика события, не зависящая от того, кто проводит испытание, также разделяется не всеми современными буржуазными учеными.

4.2. Наблюдение как случайная величина. В дальнейшем мы не будем рассматривать наблюдения с качественными результатами, так как их обработка не требует математических методов. Кроме того, вместо слов «результат наблюдения» мы будем нередко говорить просто «наблюдение», придавая, таким образом, слову «наблюдение» двойной смысл: и регистрация результата, и сам результат.

Предположим теперь, что проводится серия однородных испытаний. Наблюдения по этим испытаниям будут различаться из-за наличия случайных факторов, следовательно, каждый истинный результат в реальных условиях превращается в случайную величину. Для того чтобы найти этот истинный результат или хотя бы дать для него достаточно «хорошую» оценку, необходимо иметь определенный запас сведений о соответствующей случайной величине.

Здесь возникают два основных вопроса. Во-первых, как связано распределение случайной величины с истинным результатом, во-вторых, как найти (или оценить) это распределение по данным наблюдений?

Решение первого вопроса связано с характером ошибок, возникающих при наблюдениях. Различают ошибки трех видов.

1. **Систематическая ошибка**, т. е. ошибка, повторяющаяся и одинаковая во всей серии наблюдений. Эта ошибка связана обычно с неправильным ведением эксперимента: неисправными измерительными приборами, ошибкой экспериментатора, снимающего показания, наличием неучтенных, но постоянных факторов (изменившаяся температура, повышенная влажность, наличие сильного магнитного поля и т. п.).

2. **Грубая ошибка**, т. е. ошибка, связанная с резким нарушением условий испытания при отдельном наблюдении. Сюда относятся ошибки, связанные с толчком или поломкой прибора, грубым просчетом экспериментатора, непредвиденным посторонним вмешательством и т. д. Если систематическая ошибка характеризуется в первую очередь своей неизменностью во всей серии испытаний, то грубая ошибка присутствует обычно не более, чем в одном-двух испытаниях и характерна именно своим отличием по величине от прочих рядовых ошибок.

3. С л у ч а й н а я о ш и б к а, включающая все остальные виды ошибок. В таком определении понятие случайности оказывается суженным по сравнению с предыдущим изложением. А именно, к случайным факторам (т. е. факторам, порождающим случайную ошибку) не относятся факторы с постоянным и факторы с однократным, но очень сильным действием. Такое сужение понятия случайности не нужно для теории, однако весьма полезно на практике; это, так сказать, первый наш шаг на пути получения случайных величин с «хорошими» распределениями.

С общей точки зрения ошибки всех трех видов являются случайными величинами, коль скоро порождающие их факторы не учитываются и не включаются в число основных. При этом распределение случайных ошибок обладает одной важной особенностью — оно симметрично относительно нуля. Это значит, что ошибки, противоположные по знаку, но одинаковые по абсолютной величине, встречаются одинаково часто (в среднем). Действительно, если такой симметрии нет, то из рассматриваемой ошибки можно выделить систематическую (или соответствующую грубой ошибке) составляющую, так что остаток (а он-то и соответствует собственно случайной ошибке) уже будет иметь симметричное относительно нуля распределение.

Из симметричности распределения случайных ошибок вытекает важный вывод: *при отсутствии систематических и грубых ошибок истинный результат наблюдения есть математическое ожидание соответствующей случайной величины*. В связи с этим особую важность приобретает проблема освобождения результатов наблюдений от всех систематических и грубых ошибок.

При наличии систематической ошибки истинный результат не совпадает с математическим ожиданием наблюдаемой случайной величины, отличаясь от него как раз на величину этой ошибки. Поэтому, если от систематической ошибки не удастся избавиться, ее, как правило, можно учесть — для этого достаточно найти ее величину. Чтобы найти систематическую ошибку (или убедиться в ее отсутствии) используют следующий прием: заменяют изучаемый объект другим, достаточно изученным (эталоном) и проводят над ним ту же серию испытаний. Если невозможно провести такую замену в целом, изучают систематические

ошибки каждого прибора в отдельности. Устранению систематических ошибок помогает и более тщательный учет всех действующих факторов (температуры, давления, магнитного поля и т. п.).

Грубые ошибки учитывать заранее невозможно, поэтому с ними нужно бороться в процессе самих испытаний, проводя их достаточно тщательно. Если же все-таки появляется сомнение в каком-либо из наблюдений, то соответствующее значение ни в коем случае нельзя исправлять, подгоняя под остальные,— лучше совсем его отбросить. Однако и здесь нужно быть осторожным и не отбросить вместо грубой случайную ошибку, которая тоже может при некоторых условиях быть весьма значительной. Действительно, отбрасывая хотя бы одну случайную ошибку, мы можем исказить всю картину распределения таких ошибок (особенно при малом числе испытаний), что приведет нас в конечном счете к неверным общим выводам. Лучше всего при появлении сомнений в отдельных результатах переделать всю серию наблюдений. В некоторых случаях при известном общем характере распределения ошибок можно использовать специальные критерии, позволяющие совершенно объективно выделять в каждой серии наблюдений грубые ошибки, если таковые имеются (см., например, п. 6.5).

В дальнейших рассуждениях мы будем исходить из предположения, что все систематические и грубые ошибки учтены (в виде добавочных основных факторов) или отброшены. Поэтому в качестве истинного результата всегда будет рассматриваться математическое ожидание соответствующей случайной величины. Но для того чтобы найти это математическое ожидание или хотя бы оценить его точность, нужно знать распределение случайной величины. Существуют различные методы, позволяющие приближенно находить функцию распределения случайной величины по результатам наблюдений. Частично об этом говорится в следующем пункте, частично — в п. 7.3. Во многих случаях без всяких вычислений удастся определить тип распределения, используя определяющие свойства (см. п. 3.1) наблюдаемой случайной величины.

В подавляющем большинстве реальных испытаний наблюдения имеют нормальное (или достаточно близкое к нему) распределение. Поэтому при обработке наблюдений

первая же выдвигаемая гипотеза — нормальность соответствующего распределения; мы будем называть эту гипотезу *основной*.

Проверять основную гипотезу можно двумя способами. По первому способу исследуются условия испытаний, из которых выводятся нужные определяющие свойства. Иногда здесь удастся провести те же рассуждения, что и при исследовании рассеяния снарядов (п. 3.1), иногда видна связь с последовательностью независимых испытаний. Но чаще всего помогает следующее общее утверждение (т е о р е м а А. М. Л я п у н о в а): *если случайная величина ξ представляет собой сумму очень большого числа взаимно независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n$, влияние каждой из которых на всю сумму ничтожно мало, то величина ξ имеет распределение, близкое к нормальному*.

В этой теореме (доказательство которой слишком сложно для настоящей книги) не нужно знать ни величин ξ_i , ни их распределений, лишь бы этих величин было много, а сами они взаимно независимы и очень малы. Это и обеспечивает теореме Ляпунова широкое применение на практике. Действительно, случайная ошибка является обычно суммарным результатом действия большого числа взаимно независимых случайных факторов. А это значит, что ее можно рассматривать как сумму большого числа отдельных взаимно независимых «частных ошибок», соответствующих упомянутым факторам. Если эти «частные ошибки» очень малы (т. е. среди случайных факторов нет доминирующих), то в общей сумме они и дадут «почти» нормальное распределение.

Нормальное распределение, благодаря своей детальной изученности, наиболее удобно для практической работы. Теорема Ляпунова заставляет нас по-иному, чем раньше, взглянуть на то, какие факторы можно считать случайными (не учитывать), если мы хотим получить нормальное распределение. А именно, к случайным желательно относить только те факторы, влияние которых в отдельности очень мало; исключение можно делать лишь для тех факторов, которые сами по себе (из каких-либо других соображений) дают нормальное распределение результатов.

Проверка основной гипотезы по описанному первому способу обладает важным достоинством — она может быть

проведена еще до наблюдений и поэтому не зависит от того, сколько проведено наблюдений и какие при этом получены результаты. К сожалению, такая проверка возможна не всегда. Иногда мы просто не в состоянии проверить, что все неучитываемые факторы действительно малы; в других случаях приходится обрабатывать статистический материал, полученный другими исследователями в различных лабораториях, и условия соответствующих испытаний вообще неизвестны. В подобных ситуациях основную гипотезу приходится проверять вторым способом — непосредственно по результатам наблюдений. Соответствующие правила проверки носят название *критериев согласия*. Мы будем их рассматривать ниже, в п. 7.1; сейчас отметим только, что для получения достаточно надежных выводов критерии согласия требуют очень большого числа наблюдений.

4.3. Основная схема производства наблюдений (выборочный метод). Применение методов математической статистики к обработке наблюдений оказывается возможным благодаря тому, что производство наблюдений полностью соответствует основной схеме статистических испытаний, называемой *выборочным методом*.

Выборочный метод в самой общей форме выглядит следующим образом. Имеется некоторая большая совокупность объектов, называемая *генеральной совокупностью*. Из этой совокупности извлекаются n объектов, которые образуют *выборку*; число n называется *объемом* выборки. Эти n объектов подвергаются детальному исследованию, по результатам которого требуется описать всю генеральную совокупность или какие-нибудь ее свойства, характеристики.

Приведем простой пример применения выборочного метода. Завод, выпускающий электролампы, должен контролировать свою продукцию, в частности, проверять долговечность ламп. Чтобы проверить срок службы лампы, нужно держать ее на испытательном стенде включенной до тех пор, пока она не перегорит. Если бы завод проверял все свои лампы, то его продукция не пошла бы дальше стенда. Из создавшегося положения находят простой выход: отбирают, скажем, одну лампу на тысячу и проверяют только отобранные лампы. В этом случае по долговечности

ламп из выборки судят о долговечности всей генеральной совокупности выпускаемых заводом ламп.

Выборочный метод применяют при исследовании семян на всхожесть, при различных демографических и экономических исследованиях, при контроле за производством. На первый взгляд этот метод мало чем отличается от обычного метода малых проб. Например, при анализе вещества все исследования проводят над малыми количествами (пробами) этого вещества. Однако разница тут есть и весьма существенная: при анализе вещества мы заведомо знаем, что интересующий нас признак (количество тех или иных ионов) распределен по всей массе вещества равномерно и, следовательно, любая малая проба является точной копией всей совокупности вещества. При выборочном же методе исследуемый признак распределен по генеральной совокупности неравномерно, причем даже характер этой неравномерности неизвестен. Поэтому далеко не всякая выборка хорошо отражает структуру всей генеральной совокупности. Представьте себе, что вы хотите исследовать средний рост жителей некоторого города, а вам в качестве выборки предлагают сборную баскетбольную команду. Нетрудно понять, насколько будет искажен результат.

Не имея никаких сведений о генеральной совокупности, мы, делая выборку, можем полагаться только на случай — все прочие способы отбора будут необъективными, носящими следы влияния посторонних факторов (см. об этом ниже, в п. 10.1). И лампы для проверки долговечности, и семена для проверки всхожести, и жителей для выяснения среднего роста — все нужно отбирать совершенно случайным образом. Иное дело, если мы заранее знаем, что генеральная совокупность состоит из нескольких классов, различных по своим характеристикам. При этих условиях случайную выборку лучше делать из каждого класса в отдельности. Например, изучая рост жителей, делают отдельную выборку мужчин, отдельную — женщин; иногда при этом учитывают возраст, профессию, место жительства.

Из случайного характера выборок немедленно вытекает, что *любое суждение о генеральной совокупности по выборке само является случайным* *).

*) Имеется в виду суждение, затрагивающее хотя бы один элемент генеральной совокупности, не попавший в выборку.

Перейдем к изучению связи между наблюдениями и общим выборочным методом. Будем считать, что при каждом наблюдении, помимо контрольных признаков, отмечается один количественный результат (подсчет или измерение). Результаты любой серии наблюдений будут случайным образом колебаться вокруг истинного результата. Как уже указывалось в предыдущем пункте, это означает, что с истинным результатом связана некоторая случайная величина и каждое реальное наблюдение дает одно из значений этой величины.

Получается следующая абстрактная схема производства наблюдений: имеется случайная величина ξ и в результате n независимых испытаний получают n ее допустимых значений. Если все допустимые значения случайной величины ξ считать генеральной совокупностью, то полученные при наблюдениях n значений образуют выборку. По этой выборке мы и должны определить распределение случайной величины ξ (в дальнейшем оно называется *распределением генеральной совокупности*).

Итак, *производство наблюдений является частным случаем выборочного метода, когда в качестве генеральной совокупности берутся все допустимые значения некоторой случайной величины и исследуется распределение этой величины.*

Чтобы найти неизвестное распределение генеральной совокупности, используют следующие рассуждения. Полученные при наблюдениях числа x_1, x_2, \dots, x_n (называемые *элементами выборки*) можно считать полной совокупностью значений некоторой конечнозначной случайной величины ξ_n . При этом все полученные при наблюдениях числа нужно считать различными элементами, независимо от того, повторяются они или нет. В этих предположениях каждый элемент выборки появляется лишь в результате одного наблюдения, и значит, опыт не позволяет приписать одним элементам (как значениям случайной величины ξ_n) большую вероятность, чем другим. Иными словами, каждому элементу x_i нужно приписать вероятность $\frac{1}{n}$.

Полученное равномерное распределение величины ξ_n называется *эмпирическим*, или *выборочным*, распределением.

Если объем выборки n достаточно велик, то распределение случайной величины ξ_n должно быть в каком-то смы-

сле близко к распределению изучаемой случайной величины ξ . Это предположение оправдывается при сравнении функций распределения этих величин. Приведем без доказательства соответствующее утверждение (теорема Гливенко): с вероятностью 1 при $n \rightarrow \infty$ максимальная разность между функциями распределения случайных величин ξ_n и ξ стремится к нулю. Практически это означает, что

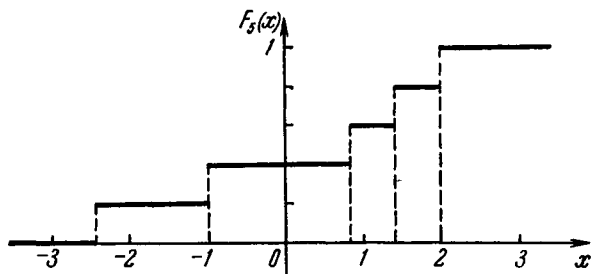


Рис. 17.

при достаточно большом объеме выборки функцию распределения генеральной совокупности можно приближенно заменять выборочной функцией распределения.

Напомним, как вычисляется функция распределения $F_n(x)$ конечнозначной случайной величины ξ_n (см. п. 2.2): для любого x она равна сумме вероятностей значений величины ξ_n , не превосходящих x . В нашем случае все элементы выборки имеют одинаковую вероятность $\frac{1}{n}$. Следовательно, функция распределения выборки в каждой точке равна числу элементов выборки, меньших, чем x , деленному на объем выборки n .

Рассмотрим, например, выборку, состоящую из элементов

—2,5 —1 0,8 1,3 2.

Объем этой выборки $n=5$. Для любого значения x функция распределения выборки $F_5(x)$ вычисляется непосредственным сравнением этого x с элементами выборки. Например, $F_5(1) = \frac{3}{5}$, так как в выборке есть три элемента (—2,5; —1 и 0,8), меньшие, чем 1. На рис. 17 приведен окончательный график функции $F_5(x)$. Мы видим, что все элементы выборки оказываются точками разрыва этой функции.

Наличие в выборке повторяющихся элементов не приводит к дополнительным трудностям. Например, для выборки

3,6 3,1 3,0 3,1 3,1 3,3 3,0 3,3

значение функции распределения при $x=3,2$ равно $\frac{5}{8}$.

4.4. Среднее и дисперсия выборки. Допустим, что, проводя наблюдения, мы сумели учесть все систематические ошибки и избежать грубых ошибок. Это значит, что истинный результат совпадает с математическим ожиданием $M\xi$ соответствующей случайной величины. Это число нам неизвестно, но, благодаря теореме Гливенко, вместо математического ожидания величины ξ при больших n можно рассматривать математическое ожидание величины ξ_n . Действительно, оба математических ожидания одинаковой формулой выражаются через свои функции распределения (эта формула содержит интеграл Стильтьеса и поэтому мы ее не приводим), откуда и вытекает, что $M\xi_n \rightarrow M\xi$ при $F_n(x) \rightarrow F(x)$.

Из этой же теоремы следует, что погрешность замены в среднем должна быть тем меньше, чем больше объем выборки.

Случайная величина ξ_n конечнозначна и имеет равномерное распределение. Поэтому (см. п. 2.3) ее математическое ожидание есть просто среднее арифметическое элементов выборки:

$$M\xi_n = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

В дальнейшем мы будем его называть *средним* выборки и обозначать \bar{x} .

Возможность приближенной замены $M\xi$ на $M\xi_n$ означает, таким образом, что *в качестве истинного результата можно брать среднее выборки наблюдений, причем точность замены тем выше, чем больше объем выборки*. Полученное утверждение составляет основу всей математической обработки наблюдений. Правда, сам по себе «принцип среднего» известен очень давно и широко используется экспериментаторами без всяких теоретических обоснований. Но математическая статистика позволяет пойти дальше и оценить,

насколько неточен переход от среднего выборки к истинному результату.

Для того чтобы найти погрешность среднего, нужно уметь оценивать, насколько точны сами наблюдения. Если бы все наблюдения давали один и тот же результат, то никакой погрешности вообще бы не было. Именно разброс результатов, т. е. превращение истинного результата в случайную величину, и порождает эту погрешность. Следовательно, точность найденного истинного результата связана в первую очередь с мерой рассеяния — дисперсией наблюдаемой случайной величины.

Дисперсия случайной величины, так же как и математическое ожидание, полностью определяется функцией распределения этой величины. Следовательно, и здесь можно использовать теорему Гливенко, в силу которой дисперсия $D\xi$ приближенно равна дисперсии $D\xi_n$ и это приближение тем лучше, чем больше объем выборки n .

Дисперсия конечнозначной равномерно распределенной случайной величины рассматривалась в п. 2.3, откуда

$$D\xi_n = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}.$$

По теореме Гливенко $D\xi_n \approx D\xi$; это равенство было бы еще более надежным, если бы в формуле для $D\xi_n$ вместо \bar{x} стоял непосредственно истинный результат $M\xi$. Из-за того, что в реальных наблюдениях всегда $\bar{x} \neq M\xi$, дисперсия $D\xi_n$ дает, как правило, заниженную оценку рассеяния значений генеральной совокупности и нуждается в некотором видоизменении. Путь этого изменения станет ясен, если несколько уточнить суть выборочного метода.

В предыдущих рассуждениях рассматривалась одна фиксированная выборка объема n , с которой связывалась конечнозначная случайная величина ξ_n . Выборку заданного объема из бесконечной генеральной совокупности можно, однако, осуществлять бесконечным числом способов. Ясно, что любая выборка при этом сама станет случайным событием, в связи с чем любая закономерность выборок может носить лишь вероятностный характер (отсюда, кстати, понятно, почему утверждение теоремы Гливенко не абсолютно достоверное, а лишь имеющее вероятность 1).

Из случайности выборок вытекает, что *все числовые характеристики выборки (в частности, среднее и дисперсия) при неизменном объеме n будут случайными величинами со своими распределениями*. Эти распределения можно находить, зная распределение основной случайной величины ξ .

Рассмотрим вначале среднее выборки, равное сумме результатов отдельных наблюдений, деленной на n . Если рассматривать всевозможные выборки, то результат первого наблюдения окажется совершенно случайным; то же самое верно для второго наблюдения, третьего и т. д. Следовательно, каждому по счету наблюдению соответствует своя случайная величина: первому — μ_1 , второму — μ_2 и т. д. Все эти величины имеют то же распределение, что и основная величина ξ — ведь именно ее мы наблюдаем всякий раз. Поэтому они имеют одинаковые математические ожидания $M\mu_1 = M\mu_2 = \dots = M\mu_n = M\xi$, одинаковые дисперсии $D\mu_1 = D\mu_2 = \dots = D\mu_n = D\xi$. Среднее выборки выражается через результаты отдельных наблюдений по формуле

$$\bar{x} = \frac{\mu_1 + \mu_2 + \dots + \mu_n}{n}.$$

Следовательно, в силу свойств математического ожидания и дисперсии любой случайной величины

$$M\bar{x} = \frac{M\mu_1 + M\mu_2 + \dots + M\mu_n}{n} = \frac{nM\xi}{n} = M\xi$$

и

$$D\bar{x} = \frac{D\mu_1 + D\mu_2 + \dots + D\mu_n}{n^2} = \frac{nD\xi}{n^2} = \frac{1}{n} D\xi.$$

Сравним теперь случайные величины ξ и \bar{x} . Обе эти величины имеют одинаковые математические ожидания, т. е. наблюдения над ними соответствуют одному и тому же истинному результату. Благодаря этому, вместо величины ξ можно изучать величину \bar{x} . Правда, каждое наблюдение над величиной \bar{x} в n раз труднее, чем наблюдение над ξ (нужно сначала получить всю выборку и лишь потом сумму элементов поделить на n), но зато у величины \bar{x} в n раз меньше дисперсия.

Перейдем к рассмотрению дисперсии $D\xi_n$. Считая, так же как и выше, каждое наблюдение случайной величиной,

получим, что

$$D\xi_n = \frac{(\mu_1 - \bar{x})^2 + (\mu_2 - \bar{x})^2 + \dots + (\mu_n - \bar{x})^2}{n},$$

откуда

$$M(D\xi_n) = \frac{M(\mu_1 - \bar{x})^2 + M(\mu_2 - \bar{x})^2 + \dots + M(\mu_n - \bar{x})^2}{n}.$$

Производя преобразования (которые мы опускаем ввиду их громоздкости), приходим к равенству:

$$M(D\xi_n) = \frac{n-1}{n} D\xi.$$

Иными словами, истинный результат наблюдений над дисперсией $D\xi_n$ как случайной величиной не совпадает с дисперсией $D\xi$, а оказывается несколько меньше последней. В связи с этим $D\xi_n$ называется *смещенной* оценкой дисперсии $D\xi$.

Из полученных формул непосредственно видно, как нужно изменить $D\xi_n$, чтобы получить *несмещенную* оценку дисперсии $D\xi$. А именно, в качестве дисперсии выборки нужно рассмотреть величину

$$s^2 = \frac{n}{n-1} D\xi_n.$$

Согласно свойствам математического ожидания

$$Ms^2 = D\xi$$

и, значит, s^2 действительно является несмещенной оценкой $D\xi^*$). Переход к несмещенной оценке s^2 важен в основном для малых выборок, ибо разница между s^2 и $D\xi_n$ при больших n незаметна. Однако во избежание разногласий мы в дальнейшем под выборочной дисперсией будем понимать только s^2 .

Используя знак Σ для обозначения суммы по всем элементам выборки, получим удобные сокращенные формулы среднего и дисперсии выборки

$$\bar{x} = \frac{1}{n} \sum x_i, \quad s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2. \quad (4.1)$$

*) Более подробно о несмещенных оценках см. ниже, в п. 5.1.

В практических вычислениях для дисперсии s^2 часто удобна формула

$$s^2 = \frac{1}{n-1} \left[\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right],$$

легко вытекающая из (4.1).

Величина s (корень квадратный из выборочной дисперсии) называется *средним квадратичным отклонением выборки* или *выборочным стандартом*.

Тот факт, что для получения несмещенной оценки дисперсии $D\xi$ в знаменателе выборочной дисперсии пришлось n заменить на $n-1$, непосредственно связан с тем, что величина \bar{x} , относительно которой берутся отклонения, сама зависит от элементов выборки. Если бы в формуле выборочной дисперсии были две такие величины, то n нужно было бы заменить на $n-2$ и т. д. (со значительным уменьшением знаменателя выборочной дисперсии нам еще придется столкнуться в регрессионном анализе, п. 9.2).

Каждая величина, зависящая от элементов выборки и участвующая в формуле выборочной дисперсии, называется *связью*. Оказывается (это можно строго доказать), *знаменатель выборочной дисперсии всегда равен разности между объемом выборки и числом связей, наложенных на эту выборку*. Эта разность фактически показывает, какое количество элементов выборки можно произвольно изменять, не нарушая связей, поэтому она называется *числом степеней свободы* выборки. Число степеней свободы участвует не только в формуле выборочной дисперсии, но и в формулах всех случайных величин, так или иначе связанных с этой дисперсией.

Математическое ожидание $M\xi$ и дисперсию $D\xi$ самой случайной величины ξ называют обычно *генеральным средним* и *генеральной дисперсией*. Применяя различные методики испытаний, мы будем получать различные случайные величины, даже исследуя один и тот же объект; соответственно будут меняться $M\xi$ и $D\xi$. Следовательно, генеральное среднее и генеральную дисперсию можно использовать для характеристики методик испытаний. При этом нужно хорошо представлять себе, что генеральное среднее и генеральная дисперсия, взятые по отдельности, слабо характеризуют методику испытаний. Так, методика с большой

дисперсией (малой точностью) может из-за отсутствия систематических ошибок дать лучшее приближение к истинному результату, чем методика с малой дисперсией, но с систематической ошибкой.

Итак, среднее и дисперсия характеризуют две различные важные стороны применяемой методики испытаний: *среднее характеризует результат, даваемый методикой, а дисперсия — точность этого результата, точность методики.* На этом «разделении ролей» среднего и дисперсии основана обработка так называемых «текущих измерений», к изложению которой мы и перейдем.

Как уже указывалось, генеральное среднее и генеральная дисперсия оцениваются средним и дисперсией выборки тем точнее, чем больше объем выборки. В практической работе, однако, не всегда есть возможность провести достаточно большое число наблюдений; если же такая возможность есть, то на проведение большой серии наблюдений потребуется много времени, в течение которого результат или точность методики могут измениться. В то же время в руках исследователя часто имеются большие совокупности наблюдений, в которых неизменна только дисперсия или только среднее. Например, состав неизвестного вещества одновременно исследуется в нескольких лабораториях, в силу чего точность методики (дисперсия) по всем наблюдениям не будет одинакова, но среднее всех наблюдений (при отсутствии систематических и грубых ошибок) одно и то же. Часто встречается и такая ситуация, когда для различных измерений с различными средними применяется одна и та же методика, одни и те же приборы, и значит, дисперсия по всем наблюдениям не изменится.

Оказывается, изменение одного из чисел (среднее, дисперсия) не мешает использовать все наблюдения для нахождения второго числа, если оно остается неизменным. Проще всего обстоят дела с вычислением среднего — здесь изменением дисперсии можно просто пренебрегать. При вычислении дисперсии нужно уже учитывать изменение среднего, что позволит найти общую дисперсию «текущих измерений».

Для вычисления дисперсии все наблюдения разбивают на отдельные выборки, в каждой из которых среднее можно считать неизменным. Пусть эти частные выборки имеют

объемы n_1, n_2, \dots, n_k . Вычислим частные дисперсии $s_1^2, s_2^2, \dots, s_k^2$ для каждой такой выборки в отдельности. Общая дисперсия всех наблюдений будет теперь равна средневзвешенному значению частных дисперсий (в качестве весов берутся степени свободы):

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{n_1 + n_2 + \dots + n_k - k}.$$

Отметим, что в знаменателе, как всегда, стоит общее число степеней свободы для всей объединенной выборки. Действительно, каждая частная выборка имеет свою связь, значит, всего k связей.

Рассмотрим пример вычисления дисперсии по «текущим измерениям». Спектральный метод определения фосфора в чугуне по различным образцам дал значения, приведенные в таблице 4.1 (в % содержания фосфора). Используем все эти данные для вычисления дисперсии указанного метода; соответствующая схема рекомендуется для обработки любых «текущих измерений».

Таблица 4.1

Номер наблюдения j	Номер образца i				
	1	2	3	4	5
1	0,42	0,26	0,09	0,60	0,47
2	0,38	0,24	0,08	0,64	0,44
3	0,39	0,21	0,08	0,62	0,46
4	0,36	0,23	0,09	0,62	0,47
5	0,41	—	0,12	0,64	0,49
6	0,39	—	0,08	0,59	0,45
7	0,40	—	—	0,61	0,48
8	0,41	—	—	0,63	—
Σx_j	3,16	0,94	0,54	4,95	3,26
Σx_j^2	1,2508	0,2222	0,0498	3,0651	1,5200
n_i	8	4	6	8	7

В нижних трех строчках таблицы 4.1 указаны результаты подсчета сумм данных каждой колонки, сумм квадратов этих данных, а также объемы наблюдений по каждому образцу (число данных в колонке). Эти числа нужны для вычисления дисперсий s_i^2 по формуле

$$s_i^2 = \frac{1}{n_i - 1} \left[\sum x_j^2 - \frac{(\sum x_j)^2}{n_i} \right].$$

Для вычисления общей дисперсии s^2 нам понадобятся, однако, не сами s_i^2 , а произведения

$$(n_i - 1) s_i^2 = \sum x_j^2 - \frac{(\sum x_j)^2}{n_i}.$$

По результатам первой колонки легко находим

$$(n_1 - 1) s_1^2 = 1,2508 - \frac{3,16^2}{8} = 1,2508 - 1,2482 = 0,0026.$$

Аналогичные расчеты дают значения

$$\begin{aligned} (n_2 - 1) s_2^2 &= 0,0013, & (n_3 - 1) s_3^2 &= 0,0012, \\ (n_4 - 1) s_4^2 &= 0,0023, & (n_5 - 1) s_5^2 &= 0,0018. \end{aligned}$$

Общее число степеней свободы здесь равно

$$8 + 4 + 6 + 8 + 7 - 5 = 28.$$

Поэтому

$$s^2 = \frac{0,0026 + 0,0013 + 0,0012 + 0,0023 + 0,0018}{28} = 0,00033.$$

Извлекая квадратный корень, можем найти

$$s = 0,018.$$

При обработке наблюдений чаще всего приходится сталкиваться с нормальным распределением. Для такого распределения математическое ожидание и дисперсия обозначаются через a и σ^2 . Во всех случаях, где это не может вызвать недоразумений, мы будем использовать эти обозначения и для любых других распределений: a — генеральное среднее, σ^2 — генеральная дисперсия (соответственно σ — генеральный стандарт).

Генеральный стандарт σ играет очень важную роль в большинстве вопросов обработки наблюдений. Как мы

увидим ниже, знание генерального стандарта дает всегда более точные оценки и в то же время заметно облегчает их получение. К сожалению, данные наблюдений не позволяют находить точное значение генерального стандарта, и мы вынуждены использовать лишь выборочный стандарт. Возникающая при этом погрешность тем меньше, чем большее число наблюдений участвовало в вычислениях выборочного стандарта, точнее, чем больше число степеней свободы у выборочной дисперсии.

Число степеней свободы у средневзвешенной дисперсии s^2 гораздо больше, чем у каждой дисперсии s_i^2 в отдельности. Поэтому s намного точнее отражает генеральный стандарт σ . В приведенном выше примере вычисления дисперсии по «текущим измерениям» можно теперь считать, что $\sigma = 0,018$, используя это значение стандарта при дальнейших применениях спектрального метода определения фосфора в чугунах. Подобное соображение особенно ценно в тех случаях, когда одна и та же методика повторяется много раз (например, при контроле за производством).

Рассмотрим теперь некоторые вопросы, связанные с обработкой косвенных измерений, т. е. случайных величин, полученных не непосредственно из наблюдений, а путем некоторого функционального перехода. Почти все исследования бывают связаны с косвенными измерениями, ибо величины, найденные из опыта, редко используются в дальнейшем сами по себе — гораздо чаще их приходится пересчитывать по тем или иным формулам.

Пусть случайная величина z зависит от наблюдений x_1, x_2, \dots, x_n по известному закону

$$z = \psi(x_1, x_2, \dots, x_n).$$

Тогда истинное значение величины z может не совпадать с математическим ожиданием Mz , и его естественнее определить этим же законом

$$a_z = \psi(a_1, a_2, \dots, a_n),$$

где a_i — генеральные средние величин x_i . Число a_z называется обычно *средним косвенного измерения*.

Дисперсия косвенного измерения σ_z^2 определяется так же, как обычная дисперсия, только отклонения берутся не от Mz , а от среднего косвенного измерения a_z . Эта дисперсия

обладает всеми свойствами обычной дисперсии; ее можно найти, если известны дисперсии σ_i^2 отдельных наблюдений x_i . На практике чаще приходится иметь дело с выборочными дисперсиями s_i^2 , по которым определяется некоторое число $s_z^2 \neq \sigma_z^2$. При достаточно больших числах степеней свободы у дисперсий s_i^2 найденное s_z^2 оказывается близким к дисперсии косвенного измерения σ_z^2 . Из соображений аналогии s_z^2 называют обычно *выборочной дисперсией косвенного измерения z* .

Чтобы найти s_z^2 , разложим z в ряд Тейлора, ограничиваясь членами первого порядка и предполагая, что отдельные наблюдения x_i мало отличаются от своих истинных значений a_i :

$$z \approx \psi(a_1, a_2, \dots, a_n) + \frac{\partial \psi(a_1, a_2, \dots, a_n)}{\partial x_1} (x_1 - a_1) + \\ + \frac{\partial \psi(a_1, a_2, \dots, a_n)}{\partial x_2} (x_2 - a_2) + \dots + \frac{\partial \psi(a_1, a_2, \dots, a_n)}{\partial x_n} (x_n - a_n).$$

Воспользовавшись тем, что дисперсия суммы независимых величин равна сумме дисперсий и что дисперсия постоянной величины равна нулю, найдем

$$s_z^2 = \left(\frac{\partial \psi}{\partial x_1} \right)^2 s_1^2 + \left(\frac{\partial \psi}{\partial x_2} \right)^2 s_2^2 + \dots + \left(\frac{\partial \psi}{\partial x_n} \right)^2 s_n^2.$$

В частности, если z зависит только от одного наблюдения x по закону $z = \psi(x)$, то

$$a_z = \psi(a_x), \quad s_z^2 = [\psi'(x)]^2 s_x^2. \quad (4.2)$$

В заключение пункта отметим, что при неизвестном распределении наблюдаемой случайной величины генеральное среднее a и генеральная дисперсия σ^2 представляют в основном лишь самостоятельную ценность (см. п. 7.2). Если же известно, что изучаемое распределение нормально, то числа a и σ^2 полностью определяют его, и следовательно, знание этих чисел является исчерпывающим знанием о величине ξ .

§ 5. ОСНОВНЫЕ ЗАДАЧИ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

5.1. Параметры распределения. Важнейшая задача математической статистики, решение которой позволило бы, в принципе, решить и все остальные задачи — это нахождение функции распределения наблюдаемой случайной величины. Для решения этой задачи можно пользоваться теоремой Гливенко (см. п. 4.3), которая позволяет неизвестное распределение приближенно заменять эмпирическим распределением случайной величины ξ_n . Теорема Гливенко не использует никаких специфических свойств генеральной совокупности, целиком опираясь на случайность выборки и соответствующие вероятностные закономерности; она применима к *любым* случайным величинам. Естественно, что от теоремы с такими общими условиями трудно ждать тонких результатов. И действительно, теорема Гливенко может быть практически использована лишь при очень больших объемах выборки. Например, для того чтобы быть достаточно уверенным, что эмпирическая функция распределения $F_n(x)$ отличается от неизвестной функции распределения $F(x)$ не более, чем на 0,1, нужно брать выборку объемом не меньше 185 элементов *).

К счастью, при обработке наблюдений редко приходится прибегать к построению эмпирической функции распределения. Даже простейший анализ условий испытаний позволяет с достаточной степенью уверенности определять тип неизвестной функции распределения — распределение Пуассона, биномиальное, нормальное распределение и т. д. В подобном случае окончательное уточнение неизвестной

*) Подсчет произведен на основании теоремы Колмогорова (см. ниже, п. 7.1).

функции распределения сводится к определению некоторых числовых параметров распределения. Эти параметры определяются по выборке, разумеется, приближенно, однако нужная точность здесь достигается при гораздо меньших объемах выборки, чем при непосредственном использовании теоремы Гливенко.

Большинство параметров определяется и для наблюдаемой случайной величины ξ (в этом случае они называются *генеральными*), и для эмпирической случайной величины ξ_n (в этом случае они называются *выборочными*). С двумя важнейшими параметрами мы уже встречались — это среднее и дисперсия. В этом пункте мы укажем еще несколько важных параметров; случайная величина ξ при этом все время предполагается непрерывной.

Формальное определение большинства параметров распределения дается с помощью функции распределения*). Если в таком определении заменить генеральную функцию распределения $F(x)$ эмпирической функцией $F_n(x)$, то из генеральных получатся соответствующие выборочные параметры. Поэтому тот факт, что выборочные параметры стремятся к генеральным при увеличении объема выборки, обеспечивается уже теоремой Гливенко, хотя при этом, как правило, получается весьма грубая оценка погрешности. В результате возникает основная задача, связанная с параметрами: используя специфические свойства каждого параметра в отдельности, найти для него более удобную оценку.

Пусть изучается генеральный параметр α и пусть по выборке объема n определена некоторая величина α_n (не обязательно соответствующий выборочный параметр). Говорят, что α_n является *состоятельной* оценкой параметра α , если с вероятностью единица $\alpha_n \rightarrow \alpha$ при $n \rightarrow \infty$. Используя теорему Гливенко, мы можем теперь сказать, что выборочные параметры являются состоятельными оценками своих генеральных параметров.

Оценка α_n называется *несмещенной*, если при каждом фиксированном n математическое ожидание $M\alpha_n = \alpha$. В част-

*) Такие определения используют, как правило, сложное понятие интеграла Стильтеса. Поэтому в дальнейшем даются отдельные формулы параметров: для генеральной совокупности через плотность распределения, для выборки — через ее элементы.

ности, выборочное среднее \bar{x} является несмещенной оценкой генерального среднего a . А вот выборочная дисперсия $D\xi_n$ (см. п. 4.4) оказалась смещенной оценкой генеральной дисперсии σ^2 и поэтому ее пришлось заменить на s^2 . Таким образом, несмещенные оценки, в отличие от состоятельных, дают уже не все выборочные параметры: каждый такой параметр нужно проверять отдельно и при необходимости исправлять наподобие дисперсии.

Еще одной важной характеристикой оценок генеральных параметров является их *эффективность*, которая для различных несмещенных оценок одного и того же параметра при фиксированном объеме выборок обратно пропорциональна дисперсиям этих оценок — чем меньше дисперсия, тем выше эффективность оценки. Там, где это возможно, стараются использовать *максимально эффективные оценки*, которым приписывается эффективность 1 (именно такими являются, например, \bar{x} и s^2). Однако получение максимальной эффективной оценки может быть сопряжено с большими трудностями и выгоднее бывает использовать менее эффективные, но зато и менее трудоемкие оценки, например, для оценки генерального стандарта σ , наряду с выборочным стандартом s , нередко используется выборочный размах W (см. п. 6.2).

Напомним, что каждый выборочный параметр является случайной величиной из-за случайности самой выборки. Поэтому лучший способ исследования получающихся оценок — вывести законы распределения соответствующих выборочных параметров. Именно на этом принципе основана обработка самого распространенного и изученного распределения — нормального.

Нормальное распределение полностью определяется двумя параметрами (среднее и дисперсия). Если же нет уверенности в том, что генеральное распределение нормально, или же если такая уверенность есть, но нужно проконтролировать среднее или дисперсию, прибегают к дополнительным параметрам.

Первая группа параметров, непосредственно обобщающая понятие дисперсии — это *моменты**) Моментом

*) Мы будем рассматривать только так называемые центральные моменты.

непрерывной случайной величины ξ с плотностью распределения $f(x)$ называется величина

$$m_k = \int_{-\infty}^{+\infty} (x - M\xi)^k f(x) dx.$$

Число k называется *порядком* момента; оно может принимать любое целое положительное значение. Нетрудно проверить, что момент первого порядка равен нулю, момент второго порядка есть дисперсия. Если дисперсия дает лишь общую оценку рассеяния значений случайной величины, то моменты дают уже более детальные сведения — они характеризуют крутизну, степень симметричности графика плотности распределения и т. п.

Для выборки с элементами x_1, x_2, \dots, x_n моменты определяются формулой

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

Наиболее важную роль играют выборочные моменты третьего и четвертого порядков (см. п. 7.1).

Моменты являются общими (интегральными) характеристиками распределения. Вторая группа параметров характеризует отдельные значения функции распределения. К ним в первую очередь относятся квантили.

Квантилем ξ_p распределения случайной величины ξ с функцией распределения $F(x)$ называется решение уравнения

$$F(\xi_p) = p.$$

Иными словами, квантиль ξ_p есть такое значение случайной величины ξ , что $P\{\xi \leq \xi_p\} = p$. Вероятность p , задаваемая в процентах, дает название соответствующему квантилю; например, $\xi_{0,3}$ называется 30%-ным квантилем.

Квантили стандартного нормального распределения (т. е. распределения с параметрами $\alpha=0$, $\sigma=1$) обозначаются через u_p ; их легко найти непосредственно из таблицы I Приложения. Если $p < \frac{1}{2}$, то, подбирая такое x , для которого $\Phi(x) = \frac{1}{2} - p$, мы найдем, что $u_p = -x$. Если же $p > \frac{1}{2}$,

то подбирают такое x , для которого $\Phi(x) = p - \frac{1}{2}$, и тогда $u_p = x$, например, 40%-ный квантиль $u_{0,4} = -0,25$; 85%-ный квантиль $u_{0,85} = 1,04$. Для удобства пользования некоторые часто употребляемые квантили стандартного нормального распределения приводятся в отдельной таблице II Приложения.

Квантиль v_p общего нормального распределения с параметрами a и σ выражается через квантиль u_p стандартного распределения по формуле

$$v_p = a + \sigma u_p. \quad (5.1)$$

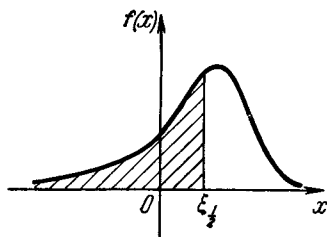


Рис. 18.

Например, 40%-ный квантиль для нормального распределения с параметрами $a=4$, $\sigma=2$ равен

$$v_{0,4} = 4 + 2(-0,25) = 4 - 0,5 = 3,5.$$

Понятие квантиля используется не только для нормального, но и для большинства встречающихся в дальнейшем распределений. Если известны два квантиля, ξ_p и ξ_q , то

$$P\{\xi_p \leq \xi \leq \xi_q\} = q - p;$$

на этом равенстве и основывается использование квантилей.

Некоторые часто встречающиеся квантили носят специальные названия. Так, квантили $\xi_{\frac{1}{4}}$ и $\xi_{\frac{3}{4}}$ называют *квартíлями*, квантили $\xi_{0,1}$, $\xi_{0,2}, \dots, \xi_{0,9}$ — *деци́лями*, квантили $\xi_{0,01}$, $\xi_{0,02}, \dots, \xi_{0,99}$ — *процентíлями*.

Наиболее важное значение имеет квантиль $\xi_{\frac{1}{2}}$, называемый *медианой* распределения. Если изобразить график плотности распределения (рис. 18), то вертикальная прямая, проходящая через медиану, расщепит площадь между графиком плотности и осью абсцисс (одна из таких половин на рис. 18 заштрихована); площадь каждой половины равна $\frac{1}{2}$.

Отметим, что медиана может не совпадать с математическим ожиданием распределения. Если же распределение симметрично, то $\xi_{\frac{1}{2}} = M\xi$. Выше уже упоминалось, что рас-

пределение случайных ошибок симметрично. Поэтому при отсутствии систематических и грубых ошибок в качестве истинного результата можно брать как математическое ожидание, так и медиану распределения.

Квантили ξ_p и ξ_{1-p} называются *симметричными*. Для симметричного относительно нуля распределения всегда $\xi_p = -\xi_{1-p}$.

Для конечнозначных и дискретных величин (а значит, и для выборок) понятие квантиля используется редко. Чаше других используется *выборочная медиана*, т. е. такое значение, для которого эмпирическая функция распределения $F_n(x) = \frac{1}{2}$. В силу ступенчатости графика $F_n(x)$ подобных значений бесчисленное множество, поэтому понятие выборочной медианы несколько уточняют. А именно, нужно все элементы выборки расположить в возрастающем порядке и в качестве медианы взять средний из них (т. е. такой, слева и справа от которого расположено одинаковое число элементов). Если выборка имеет четный объем, то у нее два средних элемента и нужно брать их полусумму.

Выборочная медиана является состоятельной и несмещенной оценкой генерального среднего, поэтому ее, так же как и выборочное среднее, можно брать в качестве приближения к истинному результату. Помимо простоты вычисления, у медианы есть еще одно преимущество перед выборочным средним: при достаточно большом объеме выборки ее распределение как случайной величины близко к нормальному, независимо от того, какое распределение имеет генеральная совокупность. Эти преимущества, правда, портит малая по сравнению с выборочным средним эффективность медианы — ее дисперсия в полтора с лишним раз больше дисперсии среднего. Поэтому медиану редко используют при обработке нормально распределенной совокупности.

Третья группа параметров определяется только для выборок и служит главным образом для проверки однородно-

сти испытаний (см. п. 6.5). Сюда относятся такие параметры, как *размах* (или амплитуда) выборки — разность между наибольшим и наименьшим элементами выборки; *наибольшее абсолютное отклонение*, т. е. наибольшая по абсолютной величине разность между элементами и средним выборки. Если в выборке многие элементы повторяются, то используют понятие *моды* — элемента с наибольшим числом повторений.

Определенные в этом пункте параметры используются в различных разделах статистического анализа. Отметим еще одну их сторону, не менее важную, по крайней мере для обработки наблюдений. Каждая серия наблюдений над одним объектом связана с очень большим цифровым материалом, куда относятся данные всех параллельных, контрольных наблюдений и т. д.; иными словами, эта серия несет в себе большое количество информации. И вся эта информация существенна для оценки полученных результатов. Особенно важна такая информация, если сопоставляются и анализируются результаты, полученные в различных лабораториях, различными исследователями. Эти результаты найдены на разных установках, с различной тщательностью — не зная всех деталей, провести надежный анализ невозможно.

С результатами других исследователей экспериментатор знакомится главным образом по публикациям. Каких-нибудь 50-100 лет назад экспериментаторов было немного, и они могли в публикациях подробно описывать все приборы, методику, качество работы. Однако со временем число публикаций настолько возросло (подсчитано, что этот рост имеет экспоненциальный характер), что их размеры пришлось резко сокращать; этот процесс продолжается и сейчас. Нет возможности опубликовывать не только описания, но даже все цифровые данные, полученные при наблюдениях. Все настоятельней стала необходимость свертывать информацию, причем для свертывания нужны такие показатели, которые самым наглядным и компактным образом характеризовали бы не только результаты, но и качество исследования.

Наилучшим в этом отношении показателем является дисперсия (или среднее квадратичное отклонение), которую и нужно обязательно указывать наряду с результатами

исследования. Простое сравнение дисперсий позволяет находить лучший метод исследования, выделять в изучаемом процессе наиболее сильно действующие факторы, выяснять неслучайность тех или иных событий и многое другое.

Если есть возможность, желательно указывать и некоторые другие параметры, что еще лучше охарактеризует полученный результат. Особенно важны такие сведения, если распределение генеральной совокупности не является нормальным.

5.2. Доверительные интервалы и доверительные вероятности. Выше неоднократно отмечалось, что выборочные параметры могут служить приближенными оценками соответствующих генеральных параметров. При этом ограничивались простым утверждением, что погрешность такой оценки тем меньше, чем больше объем выборки. Теперь настало время выяснить, каким образом оценивается подобная погрешность.

Все выборочные параметры являются случайными величинами, следовательно, и их отклонения от генеральных параметров (погрешности) также будут случайными. Таким образом, вопрос об оценке этих отклонений носит вероятностный характер: а именно, можно лишь указать вероятность той или иной погрешности. Фактически мы решаем при этом задачу, рассматривавшуюся в п. 2.2, — найти вероятность того, что некоторая случайная величина Δv (в нашем случае — отклонение выборочного параметра v от исследуемого генерального) не превосходит по абсолютной величине некоторого заданного числа ε , т. е. находится в пределах от $-\varepsilon$ до ε . Эта задача легко решается, если известна функция распределения $F(x)$ или плотность распределения $f(x)$ величины Δv :

$$P\{|\Delta v| \leq \varepsilon\} = F(\varepsilon) - F(-\varepsilon) = \int_{-\varepsilon}^{\varepsilon} f(x) dx. \quad (5.2)$$

Распределение отклонения Δv есть смещенное распределение самого параметра v . У этих распределений одинаковые дисперсии и все соответствующие моменты, различаются у них лишь математические ожидания. Иногда рас-

пределение Δv удастся довольно точно (в пределах точности всех производимых вычислений) определить по элементам выборки, иногда это распределение вообще зависит только от объема выборки n и его можно вывести теоретически. Во всяком случае, знание этого распределения является обязательным условием для проведения соответствующего анализа.

Итак, допустим, что распределение Δv известно; в частности, известно $M(\Delta v)$. Если бы при этом было известно математическое ожидание самого параметра v , то величина $l = Mv - M(\Delta v)$ дала бы точное значение генерального параметра. Однако Mv , как правило, неизвестно. Поэтому задачу о генеральном параметре решают следующим образом: находят из опыта (по выборке) одно значение v_0 выборочного параметра v и принимают его за приближенное значение генерального параметра l . Полученное выше неравенство (5.2) позволяет оценить это приближение.

Действительно, задаваясь некоторым положительным числом ε , мы можем найти вероятность P того, что $|\Delta v| = |v - l| \leq \varepsilon$. Поскольку v_0 есть одно из допустимых значений случайной величины v , то вероятность неравенства $|v_0 - l| \leq \varepsilon$ также равна P .

Мы получаем формулу

$$P \{ |v_0 - l| \leq \varepsilon \} = F(\varepsilon) - F(-\varepsilon) = \int_{-\varepsilon}^{\varepsilon} f(x) dx,$$

позволяющую сравнивать найденное значение выборочного параметра с неизвестным генеральным параметром.

Неравенство $|v_0 - l| \leq \varepsilon$ может быть переписано в виде $v_0 - \varepsilon \leq l \leq v_0 + \varepsilon$, что дает более наглядную оценку неизвестному генеральному параметру l ; вероятность нового неравенства по-прежнему равна P . Мы сталкиваемся здесь с неравенством иного типа, чем раньше, при изучении случайных величин, а именно, неизвестная (но не случайная) величина l оценивается случайными границами, ибо определенное по выборке значение v_0 является, вообще говоря, случайным. Подобная ситуация постоянно встречается в математической статистике, где для оценки любого параметра генеральной совокупности используются выборочные, а значит, случайные величины.

Итак, любая статистическая оценка есть оценка вида $v' \leq l \leq v''$, где v' и v'' — некоторые случайные величины. Придавая v' и v'' конкретные значения, мы сможем вычислять вероятность соответствующей оценки. Наиболее удобно в качестве границ v' и v'' брать квантили одной какой-либо случайной величины v . Вероятность оценки $v_p \leq l \leq v_q$

находится тогда очень легко и равна $q - p$.

Можно решать и обратную задачу: по заданной вероятности определить границы. Эта задача имеет бесчисленное множество решений — например, вероятности p соответствует любая оценка вида

$$v_\alpha \leq l \leq v_{p+\alpha},$$

где $0 < \alpha < 1 - p$. При обработке наблюдений для оценок берут, как правило, симметричные квантили. В этом случае вероятности p соответствует оценка

$$v_{(1-p)/2} \leq l \leq v_{(1+p)/2}.$$

Связь между квантильными границами и соответствующей вероятностью хорошо видна, если воспользоваться графиком плотности распределения величины v (рис. 19). Квантили $v_{(1-p)/2}$ и $v_{(1+p)/2}$ находятся на одинаковом расстоянии от начала координат соответствующие им ординаты отсекают площадь, равную p (на рис. 19 заштрихована).

Как же пользоваться полученными оценками на практике? Ведь все практические рекомендации должны носить категорический характер. Например, исследуя грунт, мы должны дать ответ, можно или нельзя здесь строить плотину; вряд ли строителей удовлетворит ответ: «Стройте с вероятностью 0,9». Более того, на основании наблюдений человечество сумело за всю историю науки сделать немало важных выводов, правильность которых подтверждена всем дальнейшим прогрессом. А ведь мы только что показали, что все эти выводы держатся лишь на случайных оценках!

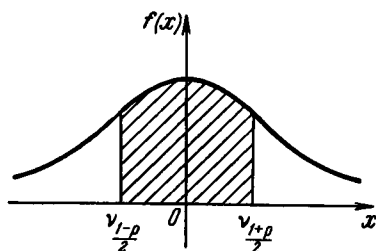


Рис. 19.

И еще одно соображение: производя различные измерения, мы привыкли всегда указывать определенный, достоверный результат измерения. Если же мы и указываем погрешность, то это опять-таки вполне определенная, достоверная величина, причем небольшая, в то время как квантильные границы оценок могут быть сколь угодно большими.

Возникшее противоречие между теорией и практикой оказывается легко устранимым. Во всех перечисленных случаях результаты действительно являются достоверными. Однако при этом речь идет о так называемой *практической достоверности*, в отличие от *абсолютной* (или теоретической).

Событие называется *абсолютно достоверным*, если оно появляется при любом осуществлении комплекса основных факторов (именно так определялась достоверность в п. 1.1). Абсолютную достоверность нельзя установить никакой самой длительной проверкой, ее можно вывести лишь теоретически, путем логических умозаключений. Сюда относятся в основном математические истины и некоторые выводы других точных наук.

Большинство привычных достоверных событий при ближайшем рассмотрении не оказываются достоверными абсолютно. Нельзя, например, считать абсолютно достоверным тот факт, что подброшенная монета упадет или гербом, или цифрой — ведь у монеты есть и другие состояния равновесия (скажем, на ребре). Даже строго математически доказанные теоремы не всегда можно считать абсолютно достоверными, так как сюда примешивается возможность ошибки доказавшего теорему математика; могут ошибаться и те, кто проверял доказательство.

Таким образом, безупречное с научной точки зрения понятие абсолютной достоверности оказывается совершенно неприемлемым с практической точки зрения. Однако отбрасывать это понятие нельзя. Вспомним выводы, связанные с последовательностью независимых испытаний (п. 1.5). Вероятность того, что событие A с вероятностью p осуществится во всех n испытаниях, равна p^n . Но при $p < 1$ (как бы ни было p близко к 1) обязательно $p^n \rightarrow 0$ при $n \rightarrow \infty$. А это значит, что проводя достаточно большое число испытаний, мы обязательно получим такое испытание, в котором

событие A не произойдет. Если же вспомнить геометрическое определение вероятности, то мы столкнемся с событиями, вероятность которых даже равна 1 и которые, тем не менее, в отдельном испытании могут не произойти.

Итак, вопрос о том, какие результаты исследования можно считать практически достоверными, оказывается далеко не простым. Все в конечном счете зависит от того, сколь велико число дальнейших применений этого результата, а также сколь велика опасность единичной ошибки. Так, например, вероятность ясной погоды 0,9 достаточна для того, чтобы выйти из дому без зонтика; однако если 0,9 — это вероятность того, что у некоторого вещества не будет самопроизвольного взрыва, вряд ли вы станете небрежно хранить это вещество.

Из всего сказанного ясно, что событие A с вероятностью $p \approx 1$ может считаться практически достоверным, если число n всех реально проводившихся ранее и проводимых в будущем испытаний над этим событием невелико, т. е. вероятность p^n мало отличается от 1. Определение, как мы видим, весьма расплывчатое и к тому же несет в себе психологический элемент (оптимизм или скептицизм самого исследователя). Поэтому лучше всего регулярно указывать вероятность (*уровень достоверности*) каждого получаемого результата.

Отметим еще одно обстоятельство. Вероятность большинства реальных событий заранее неизвестна и вычисляется опять-таки с помощью испытаний. Поэтому каждое реально проведенное испытание, при котором появилось событие A , с одной стороны, приближает то «роковое» испытание, в котором событие A может не появиться, а с другой стороны, увеличивает вероятность события A и, значит, отодвигает «роковое» испытание. Именно это обстоятельство и позволяет нам быть уверенными в появлении абстрактно не достоверных событий, осуществлявшихся уже в большом числе предыдущих испытаний.

Вернемся к вопросу об оценках генеральных параметров. Использование принципа практической достоверности позволяет не доводить окончательную вероятность оценки до 1 (что дало бы бесконечный интервал в качестве границ), а считать окончательной менее вероятную оценку. Принимаемый при этом уровень достоверности называется до-

верительной вероятностью. В зависимости от конкретных обстоятельств в качестве доверительной вероятности берут 0,95; 0,98; 0,99; реже 0,90 или 0,999.

Соответствующие доверительной вероятности квантильные границы (только такие границы и будут рассматриваться в дальнейшем) называются *доверительными границами*, образуемый ими интервал — *доверительным интервалом* (или *доверительной оценкой*).

Найдем, например, доверительную оценку генерального среднего a по одному наблюдению $x_0=3$, если известно, что генеральная совокупность имеет нормальное распределение со стандартом $\sigma=0,9$. В качестве доверительной вероятности возьмем 0,95.

Как показано в начале пункта, соответствующая оценка имеет вид $x_0 - \varepsilon \leq a \leq x_0 + \varepsilon$, где ε есть оценка абсолютного отклонения. Иными словами, в качестве доверительных границ можно взять симметричные квантили $u_{0,025}$ и $u_{0,975}$ нормального распределения со средним $x_0=3$ и стандартом $\sigma=0,9$. Используя формулу (5.1) предыдущего пункта, найдем, что

$$u_{0,025} = 3 + 0,9 (-1,96) = 1,236,$$

$$u_{0,975} = 3 + 0,9 \cdot 1,96 = 4,764$$

(здесь число $u_{0,975}=1,96$ найдено из таблицы II Приложения). Окончательно получим, что

$$1,236 \leq a \leq 4,764.$$

Возможно, полученный доверительный интервал нас не устроит. Однако любое его сужение повлечет снижение доверительной вероятности, что нежелательно. Поэтому единственный путь улучшения оценок — снижение соответствующей дисперсии (путем улучшения методики, уточнения действующих факторов и т. д.).

5.3. Проверка статистических гипотез. Из принципа практической достоверности, изложенного в предыдущем пункте, немедленно вытекает принцип практической невозможности: *события с очень малыми вероятностями можно в практических приложениях считать невозможными*.

В качестве примера использования этого принципа напомним сформулированное в п. 3.2 правило трех сдгм.

Вероятность того, что абсолютное отклонение нормально распределенной случайной величины превзойдет ее стандарт σ не более, чем в три раза, равна 0,9973. Значит, вероятность отклонения, большего 3σ , равна 0,0027. С практической точки зрения столь малой вероятностью можно пренебречь, что и приводит к правилу трех сигм.

Принципы практической достоверности и практической невозможности представляют собой, в сущности, одно и то же утверждение, примененное к противоположным событиям. Поэтому, вместо того чтобы говорить о практической достоверности некоторой доверительной оценки, можно говорить о практической невозможности отклонений, превышающих эту оценку. Если для нормально распределенной случайной величины в качестве доверительной вероятности взять 95%, то соответствующей доверительной оценкой абсолютного отклонения будет неравенство

$$\Delta\xi \leq 1,96\sigma.$$

Отклонения, большие чем $1,96\sigma$, нужно теперь считать практически невозможными. Разумеется, такая оценка более «рискована», чем правило трех сигм, поэтому ею можно пользоваться лишь при малом числе предстоящих испытаний (чаще всего при одном испытании).

Допустим теперь, что отклонения $\Delta\xi$, большие, чем некоторое доверительное число, признаны нами практически невозможными. Этот вывод был, очевидно, сделан на основании некоторого теоретического распределения, которое, по тем или иным соображениям, мы считаем распределением величины ξ . Производя наблюдение, мы получим реальное значение отклонения $\Delta\xi$. И может оказаться, что это реальное значение превосходит доверительную оценку. Какой вывод нужно из этого сделать?

Первое, что может прийти на ум — это влияние случая. Ведь доверительная оценка не является абсолютно достоверной; возможно, здесь в первом же испытании сыграла роль та ничтожно малая (но существующая!) вероятность, которой мы пренебрегли в доверительной оценке. Однако при таком допущении пришлось бы отказаться от принципа практической достоверности, который уже в течение многих столетий надежно проверен человеческой практикой.

Поэтому более естественным будет второе предположение: несоответствие принятого нами теоретического распределения реальному распределению величины ξ . Это несоответствие может быть коренным (не тот тип распределения) или может быть связано с неправильным определением параметров распределения. Таким образом, принцип практической невозможности удастся использовать в этом примере как один из критериев проверки гипотезы о распределении величины ξ .

Принцип практической невозможности может быть использован в самых различных задачах, где возникает необходимость проверять, случайно или неслучайно появилось то или иное событие. При этом всякий раз практическая невозможность события полностью отвергает случайность его появления, заставляя пересмотреть исходные предпосылки вычисления вероятности.

Эти рассуждения ясно вырисовывают разницу использования теории вероятностей для изучения предстоящих или уже осуществившихся событий. По отношению к предстоящим событиям главное — это надежное предвидение, поэтому здесь мы интересуемся лишь событиями с большими вероятностями. Из осуществившихся же событий нас интересуют в первую очередь события с малой вероятностью. Чем меньше расчетная вероятность уже осуществившегося события, тем больше его «неслучайность» и тем важнее эту «неслучайность» раскрыть. Событие как бы сильнее привлекает к себе внимание наблюдателя, становится для него более *значимым*.

Использование принципа практической невозможности для доказательства неслучайного появления события с малой вероятностью называется *принципом значимости*. Наибольшее значение вероятности, несовместимой со случайностью события, называется *уровнем значимости*. Иными словами, уровень значимости есть максимум таких вероятностей, при которых события можно считать практически невозможными. Но событие, противоположное практически невозможному, является практически достоверным. Поэтому *принятые нами уровень значимости и уровень достоверности должны в сумме давать единицу*.

Теперь можно дать более строгое определение значимости событий. Событие A называется *значимым*, если его

вероятность $P(A)$ меньше, чем принятый уровень значимости. Чем выше уровень значимости, тем он «жестче», ибо тем большее число событий нельзя рассматривать как случайные. Уровень значимости — это как бы величина ячеек «сита», сквозь которое отсеиваются неслучайные события. Наиболее употребительны уровни значимости 0,05; 0,02; 0,01; реже 0,10 или 0,001. Чтобы лучше «почувствовать» уровень

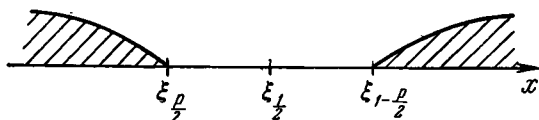


Рис. 20.

значимости, можно пользоваться аналогией между вероятностью и частотой, считая, что *уровень значимости, выраженный в процентах, показывает, сколько раз в ста испытаниях мы рискуем ошибиться, объявив изучаемое событие неслучайным*. Так, наиболее употребительный в данной книге 5%-ный уровень значимости допускает ошибку в пяти случаях из ста. Поскольку каждую проверку гипотезы можно считать одним испытанием, то такой уровень вполне допустим при единичных проверках.

Чаще всего принцип значимости применяется для проверки так называемых *статистических гипотез*. Эти гипотезы имеют самые различные формулировки, но, в конечном счете, являются гипотезами о распределении той или иной случайной величины. Проверка каждой такой гипотезы осуществляется следующим образом.

Выбирается уровень значимости p , ему соответствует доверительная вероятность $1-p$. По этой вероятности, используя гипотезу о распределении величины ξ , находят квантильные доверительные границы, как правило симметричные, т. е. $\xi_{p/2}$ и $\xi_{1-p/2}$. Числа $\xi_{p/2}$ и $\xi_{1-p/2}$ называются *критическими значениями* гипотезы; значения x , меньшие, чем $\xi_{p/2}$, и большие, чем $\xi_{1-p/2}$, образуют *критическую область* гипотезы (рис. 20). Для многих широко используемых на практике распределений составлены таблицы критических значений при различных уровнях значимости.

Следующим этапом находят реальное значение ξ_0 изучаемой случайной величины (обычно его вычисляют по вы-

борке). Если найденное значение ξ_0 попадает в критическую область, то, по нашей гипотезе оно является практически невозможным. Но так как оно все-таки появилось, то должна быть отвергнута гипотеза. Если же ξ_0 попадает между $\xi_{p/2}$ и $\xi_{1-p/2}$, то гипотеза вполне допускает такое значение в качестве случайного (на данном уровне значимости); поэтому нет никаких оснований ее отвергать.

Мы видим, что первое суждение (гипотеза неверна) гораздо более категорично, чем второе (гипотеза не отвергается, но и не утверждается, что она верна). И это не удивительно — в обоих случаях для проверки гипотезы используется лишь одно значение ξ_0 . Но если для опровержения гипотезы достаточно одного противоречащего примера, то доказать правильность гипотезы нельзя даже с помощью тысячи подтверждающих примеров — опровергающим может оказаться еще не найденный нами тысяча первый. Конечно, и подтверждающие примеры не бесполезны для научного познания, ибо каждый такой пример есть испытание, увеличивающее вероятность того, что гипотеза правильна*).

Принимая решение по результатам проверки, мы можем допустить ошибку. Возможные ошибки различаются по своему характеру. *Ошибка первого рода* состоит в том, что отвергается гипотеза, которая на самом деле верна. Вероятность такой ошибки не выше уровня значимости, следовательно, достаточно мала. *Ошибка второго рода* состоит в том, что гипотеза принимается, а на самом деле она не верна. Вероятность ошибки второго рода зависит от характера проверяемой гипотезы, от способа проверки и от многих других причин, что сильно усложняет ее оценку. Ясно только, что эта вероятность тем меньше, чем «жестче» принятый уровень значимости, ибо при этом увеличивается число отвергаемых гипотез.

Одну и ту же статистическую гипотезу можно исследовать с помощью различных случайных величин. Каждый такой способ исследования называется *критерием значимости*. Для проверки гипотезы стараются из всех

*) Понятие «вероятность» употреблено здесь не совсем точно — это, скорее, житейская уверенность (ибо правильность гипотезы не есть случайное событие).

возможных критериев выбрать тот, у которого при заданном уровне значимости меньше вероятность ошибки второго рода. Теория такого выбора в общем виде слишком сложна и выходит за рамки настоящей книги; отдельные сведения о контроле за ошибками второго рода имеются в следующем пункте и п. 10.2.

5.4. Односторонние и двусторонние критерии. Наиболее часто встречаются статистические гипотезы, связанные со сравнением различных выборок.

Рассмотрим следующий пример. Изучаются два типа резцов, применяемых при обработке деталей на токарном станке. С помощью резцов получают некоторое число деталей. Диаметры этих деталей образуют две выборки, соответствующие каждому типу резца; дисперсии этих выборок несколько различаются. Такое различие может, конечно, оказаться результатом случайных причин, а может быть и следствием разницы резцов. Зная, что распределения результатов по каждому резцу являются нормальными, мы должны фактически проверить гипотезу, одинаковы ли генеральные дисперсии этих распределений. Если такая гипотеза будет отвергнута, то одному из резцов нужно будет отдать предпочтение.

Другим примером может служить сравнительное испытание на всхожесть двух сортов пшеницы. Вычисляя количество проросших семян каждого сорта на нескольких участках, мы, как и выше, получим две выборки, у которых теперь нужно сравнивать средние. Если генеральные средние обоих соответствующих распределений окажутся одинаковыми, то различие между сортами пшеницы будет только случайным; если же они окажутся разными, то различны по всхожести и сами сорта.

Сравнение двух или нескольких выборок приходится проводить, сравнивая различные методики анализа, различные условия производства; с такой же задачей приходится сталкиваться при обработке «текущих измерений» (см. п. 4.4). Весьма важно следить за неизменностью основных параметров при исследованиях, требующих длительного времени.

Приведенным примерам соответствует следующая общая схема. Найдены два значения α_1 и α_2 некоторого выбо-

точного параметра. Эти значения можно рассматривать как оценки генеральных параметров A_1 и A_2 . Высказывается гипотеза, что различие между α_1 и α_2 чисто случайное и что на самом деле $A_1 = A_2$, т. е. между генеральными параметрами нет различий. Такая гипотеза называется *нулевой*. Для проверки этой гипотезы нужно выяснить, значительно ли расхождение между α_1 и α_2 в условиях нулевой гипотезы. С этой целью обычно исследуют случайную величину $\Delta\alpha = \alpha_1 - \alpha_2$ и проверяют, значительно ли ее отличие от нуля.

Иногда удобнее рассматривать величину $\frac{\alpha_1}{\alpha_2}$, сравнивая ее с единицей. Конкретные методы таких исследований приводятся в следующем параграфе.

Гипотеза $A_1 \neq A_2$ называется *альтернативной*. Отвергая нулевую гипотезу, мы тем самым принимаем альтернативную гипотезу. Альтернативная гипотеза в свою очередь распадается на две: $A_1 > A_2$ и $A_1 < A_2$. Если одно из этих неравенств заведомо невозможно, то альтернативная гипотеза называется *односторонней* и для ее проверки применяются *односторонние критерии значимости* (в отличие от обычных, *двусторонних*).

Как будет показано ниже, односторонний критерий значимости имеет намного меньшую вероятность ошибки второго рода, чем соответствующий двусторонний. Уже из этого видно, насколько полезно предварительно выяснить, какой из сравниваемых параметров A_1 и A_2 не может быть меньше другого. Односторонний характер альтернативной гипотезы зачастую вытекает из самой постановки задачи. Например, изучая эффективность некоторого усовершенствования производственного процесса, мы заранее можем считать, что это усовершенствование способно лишь уменьшить дисперсию процесса. Точно так же при исследовании удобрения можно считать, что его применение увеличивает среднюю урожайность (т. е. генеральное среднее).

Односторонний критерий значимости легко получать из двустороннего. Обратимся к рис. 20. Мы видим, что критическая область гипотезы (заштрихованная на рис. 20) состоит из двух частей. Каждая часть соответствует своему неравенству: $A_1 > A_2$ или $A_1 < A_2$. Если мы заранее знаем, что возможно лишь одно из этих неравенств, то и рассматривать мы должны лишь одну из половин критической

области. Вероятность попадания в критическую область уменьшится, тем самым, ровно вдвое и станет равна $\frac{p}{2}$.

Таким образом, при одностороннем критерии значимости можно использовать те же критические значения, что и при двустороннем, однако этим значениям будет соответствовать вдвое меньший уровень значимости. Например, уровню значимости 0,05 при двустороннем критерии соответствуют критические значения $\xi_{0,025}$ и $\xi_{0,975}$, т. е. значимыми (неслучайными) считаются значения ξ_0 , удовлетворяющие неравенствам $\xi_0 < \xi_{0,025}$ и $\xi_0 > \xi_{0,975}$. Если же перейти к одностороннему критерию, то одно из этих неравенств (например, $\xi_0 < \xi_{0,025}$) заведомо невозможно и значимыми будут лишь значения ξ_0 , удовлетворяющие другому неравенству ($\xi_0 > \xi_{0,975}$). Вероятность последнего неравенства равна 0,025, таков и будет уровень значимости одностороннего критерия.

Обычно для одностороннего критерия берут тот же уровень значимости, что и для двустороннего, так как ошибка первого рода в обоих случаях нежелательна совершенно одинаково. Для этого нужно вывести односторонний критерий из двустороннего, соответствующего вдвое большему уровню значимости, чем тот, что нами принят. Так, в предыдущем примере, желая сохранить уровень значимости 0,05 для одностороннего критерия, мы для двустороннего должны были бы взять уровень 0,10, что дало бы критические значения $\xi_{0,05}$ и $\xi_{0,95}$. Из этих значений для одностороннего критерия сохраняется одно (скажем, $\xi_{0,95}$), которое и будет окончательным критическим значением, соответствующим одностороннему критерию при уровне значимости 0,05.

Итак, при одном и том же уровне значимости 0,05 одному и тому же неравенству $A_1 > A_2$ в случае двустороннего критерия соответствует критическое значение $\xi_{0,975}$, а одностороннего — $\xi_{0,95}$. Но $\xi_{0,95} < \xi_{0,975}$, значит, при одностороннем критерии большее число значений ξ_0 придется считать не случайными (значимыми), большее число гипотез будет отвергнуто. Тем самым уменьшится вероятность принять неверную гипотезу, допустить ошибку второго рода. А вероятность ошибки первого рода как для одностороннего, так и для двустороннего критерия остается одинаковой, ибо она равна уровню значимости.

Чтобы нагляднее подчеркнуть преимущества одностороннего критерия значимости перед двусторонним, приведем следующий пример. Сталеплавильный завод изготавливает специальную сталь, которая должна содержать 40% ванадия. Контроль ведется на уровне значимости 0,05; методика контроля дает нормальное распределение результатов со стандартом $\sigma=2\%$. Контрольный анализ партии стали дал для содержания ванадия значение 36,4%. Достаточно ли этого результата, чтобы забраковать партию?

Обозначим через ξ результат произвольного анализа над доброкачественной сталью. Согласно условиям задачи величина ξ имеет нормальное распределение с параметрами $a=40$ и $\sigma=2$. Правило вычисления квантилей такого распределения было указано в п. 5.1. Используя таблицы II Приложения, найдем

$$\begin{aligned}v_{0,025} &= 40 + 2u_{0,025} = 40 - 2 \cdot 1,96 = 36,08, \\v_{0,975} &= 40 + 2u_{0,975} = 40 + 2 \cdot 1,96 = 43,92.\end{aligned}$$

В качестве нулевой гипотезы здесь нужно взять гипотезу о том, что исследуемая сталь доброкачественна и, следовательно, значение $\xi_0=36,4$ появилось в результате случайностей анализа. Критическими значениями такой гипотезы при двустороннем критерии будут числа $v_{0,025}=36,08$ и $v_{0,975}=43,92$; критическая область образуется неравенствами $\xi < 36,08$ и $\xi > 43,92$. Значение $\xi_0=36,4$ не попадает в эту критическую область, следовательно, двусторонний критерий не позволяет отвергнуть нулевую гипотезу и считать сталь недоброкачественной.

Условия задачи позволяют применить односторонний критерий значимости. Действительно, найденное значение $\xi_0=36,4$ меньше медианы $v_{0,50}=40$, поэтому его можно сравнивать только с теми критическими значениями, которые меньше 40. Критическим значением проверяемой нулевой гипотезы при одностороннем критерии является квантиль $v_{0,05}=40+2u_{0,05}=40-2 \cdot 1,64=36,72$.

Мы видим, что $\xi_0 < 36,72$, т. е. ξ_0 попадает в критическую область. Таким образом, односторонний критерий, как более точный, сумел при тех же исходных данных выявить недоброкачественность стали.

§ 6. ОЦЕНКА РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ НАД НОРМАЛЬНО РАСПРЕДЕЛЕННОЙ СЛУЧАЙНОЙ ВЕЛИЧИНОЙ

6.1. Оценка генерального среднего. В настоящем параграфе рассматриваются некоторые методы статистических оценок. При этом постоянно предполагается, что наблюдаемая случайная величина (или, как мы условились говорить, генеральная совокупность) имеет нормальное распределение. Через p все время обозначается принятый уровень значимости; доверительная вероятность соответственно будет равна $1-p$.

Основным оцениваемым параметром является генеральное среднее. Особенно важную роль играет среднее в обработке наблюдений — ведь здесь оно совпадает с истинным результатом наблюдений (см. п. 4.2).

Легче всего дать оценку для генерального среднего в тех случаях, когда с достаточно высокой степенью точности известна генеральная дисперсия σ^2 . Генеральную дисперсию можно найти только приближенно по выборочной дисперсии; погрешность такого приближения в зависимости от объема выборки n изучается в следующем пункте. На практике эту погрешность обычно не учитывают уже при $n \geq 50$. Разумеется, такое большое количество наблюдений над одним объектом проводится редко. Однако здесь можно пользоваться сериями наблюдений и над другими объектами, если только у этих серий та же самая генеральная дисперсия (сравнение двух или нескольких дисперсий также возможно методами математической статистики, оно изучается ниже, в п. 6.3). Дисперсия будет тогда вычисляться по «текущим измерениям», как указывалось в конце п. 4.4. Используя наблюдения над большим количеством объектов, мы сможем сделать общее число наблюдений до-

статочно большим и, значит, найти генеральную дисперсию с высокой степенью точности.

Знание генеральной дисперсии позволяет оценивать генеральное среднее даже по одному наблюдению (пример такой оценки дан в конце п. 5.2). А именно, если при наблюдении над случайной величиной ξ получено значение x_0 , то для генерального среднего a имеет место следующая доверительная оценка (с доверительной вероятностью $1-p$):

$$x_0 - \sigma u_{1-p/2} \leq a \leq x_0 + \sigma u_{1-p/2},$$

где $u_{1-p/2}$ — квантиль стандартного нормального распределения, который можно найти из таблицы II Приложения. Например, при стандарте $\sigma=4,5$ и доверительной вероятности $1-p=0,98$ значение $x_0=142$ даст оценку:

$$142 - 4,5 \times 2,33 \leq a \leq 142 + 4,5 \cdot 2,33,$$

откуда $131,5 \leq a \leq 152,5$.

Если над случайной величиной ξ проведено несколько наблюдений, то для оценки генерального среднего можно использовать выборочное среднее \bar{x} . Как следует из п. 3.2, это среднее также является случайной величиной с нормальным распределением. В п. 4.4 было показано, что математическое ожидание у величины \bar{x} то же самое, что и у ξ , а дисперсия уменьшается в n раз (n — число наблюдений) и равна

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}.$$

Каждая выборка есть одно наблюдение над величиной \bar{x} . Поэтому для генерального среднего a получается оценка

$$\bar{x} - \sigma_{\bar{x}} u_{1-p/2} \leq a \leq \bar{x} + \sigma_{\bar{x}} u_{1-p/2}.$$

Как и выше, генеральную дисперсию σ^2 считаем известной, откуда $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Окончательно получаем оценку

$$\bar{x} - \frac{\sigma}{\sqrt{n}} u_{1-p/2} \leq a \leq \bar{x} + \frac{\sigma}{\sqrt{n}} u_{1-p/2}. \quad (6.1)$$

Из этой оценки видно, в частности, что *уменьшение доверительного интервала обратно пропорционально корню*

квадратному из числа наблюдений, т. е. если мы хотим уменьшить возможную ошибку в два раза, мы должны число наблюдений увеличить в четыре раза.

В качестве примера оценим генеральное среднее a по генеральной дисперсии $\sigma^2=0,16$ и по трем наблюдениям $x_1=7,2$; $x_2=7,8$; $x_3=7,6$. Здесь $n=3$,

$$\bar{x} = \frac{7,2+7,8+7,6}{3} = 7,53.$$

В качестве доверительной вероятности возьмем $1-p=0,95$, тогда

$$u_{1-p/2} = 1,96.$$

Поэтому

$$7,53 - \frac{\sqrt{0,16}}{\sqrt{3}} 1,96 \leq a \leq 7,53 + \frac{\sqrt{0,16}}{\sqrt{3}} 1,96.$$

После всех вычислений получим окончательную оценку $7,07 \leq a \leq 7,99$.

В проведенных рассуждениях мы пользовались тем, что \bar{x} , как случайная величина, имеет нормальное распределение с параметрами a и $\frac{\sigma}{\sqrt{n}}$. Благодаря этому величина

$$u = \frac{\bar{x} - a}{\sigma} \sqrt{n} \quad (6.2)$$

имеет стандартное нормальное распределение и с вероятностью $1-p$ удовлетворяет неравенству

$$-u_{1-p/2} \leq u \leq u_{1-p/2}. \quad (6.3)$$

Подставляя в (6.3) значение u из формулы (6.2), мы вновь приходим к оценке (6.1) генерального среднего a .

Как уже указывалось, генеральную дисперсию σ^2 нельзя найти из наблюдений, поэтому вместо нее обычно берут выборочную дисперсию s^2 . Это значит, что вместо величины u на самом деле рассматривается величина

$$t = \frac{\bar{x} - a}{s} \sqrt{n}. \quad (6.4)$$

При больших n дисперсия s^2 мало отличается от σ^2 и

значит, величина t мало отличается от величины u . При малых же объемах выборок различие между t и u оказывается весьма существенным; более того, распределение величины t уже не является нормальным.

Общие законы теории вероятностей позволяют вывести формулы, описывающие распределение величины t . Это распределение называется *t-распределением* или *распределением Стьюдента* *); оно зависит только от числа f степеней свободы, по которым подсчитана дисперсия s^2 . Если дисперсия s^2 и среднее \bar{x} подсчитывались по одним и тем же наблюдениям, то $f=n-1$, где n — объем выборки.

Мы не будем приводить формулу плотности t -распределения, слишком громоздкую и содержащую специальные

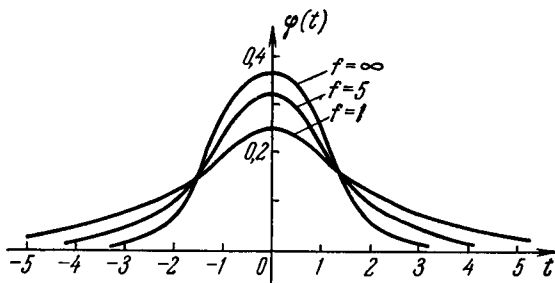


Рис. 21.

функции. Свойства этой плотности хорошо видны на графике; на рис. 21 приведены графики плотности t -распределения при различных значениях f . Они напоминают по форме плотность нормального распределения, но при $t \rightarrow \pm\infty$ значительно медленнее сближаются с осью абсцисс. При $f \rightarrow \infty$ дисперсия $s^2 \rightarrow \sigma^2$, поэтому распределение Стьюдента сближается с нормальным; случай $f=\infty$ вообще соответствует нормальному распределению. При малых же f распределение Стьюдента сильно отличается от нормального, в силу чего его роль особенно велика в так называемой *микростатистике* или статистике малых выборок.

*) Стьюдент — псевдоним английского статистика Госсета, открывшего t -распределение в 1908 г.

В дальнейшем, согласно общему правилу, через t_p обозначаются квантили t -распределения. Это распределение симметрично относительно нуля, поэтому $t_p = -t_{1-p}$. При доверительной вероятности $1-p$ для величины t получается доверительная оценка

$$-t_{1-p/2} \leq t \leq t_{-p/2}.$$

Подставляя сюда выражение для t из формулы (6.4), получим неравенство

$$-t_{1-p/2} \leq \frac{\bar{x}-a}{s} \sqrt{n} \leq t_{1-p/2},$$

откуда, после преобразований, найдем

$$\bar{x} - \frac{s}{\sqrt{n}} t_{1-p/2} \leq a \leq \bar{x} + \frac{s}{\sqrt{n}} t_{1-p/2}.$$

Полученная оценка очень похожа на оценку (6.1), полученную ранее, только здесь σ заменено на s , в связи с чем вместо $u_{1-p/2}$ приходится рассматривать $t_{1-p/2}$. Значения $t_{1-p/2}$ для различных чисел степеней свободы f и уровней значимости p приведены в таблице III Приложения.

Распределение Стьюдента позволяет оценивать генеральное среднее, когда генеральная дисперсия неизвестна. При этом число наблюдений может быть очень малым, даже равным двум. Конечно, скудость информации сказывается на результатах — доверительные границы получаются довольно широкими. Поэтому везде, где только можно, нужно стараться увеличивать число степеней свободы у выборочной дисперсии, привлекая, в частности, «текущие измерения».

Сравним, например, две оценки генерального среднего по выборочному среднему $\bar{x}=18,6$, найденному по трем наблюдениям; в качестве доверительной вероятности возьмем $1-p=0,95$. В обеих оценках будем использовать одну и ту же дисперсию 0,25, только вначале будем считать ее генеральной, а потом — выборочной, найденной по тем же трем наблюдениям. Если 0,25 — это генеральная дисперсия, то $\sigma=0,5$, и используя нормальное распределение, получим доверительную оценку

$$18,6 - \frac{0,5}{\sqrt{3}} 1,96 \leq a \leq 18,6 + \frac{0,5}{\sqrt{3}} 1,96$$

или, после всех вычислений, $18,03 \leq a \leq 19,17$.

Если же 0,25 — это выборочная дисперсия, то $s=0,5$ и нужно воспользоваться распределением Стьюдента, в силу, которого справедлива доверительная оценка

$$18,6 - \frac{0,5}{\sqrt{3}} 4,30 \leq a \leq 18,6 + \frac{0,5}{\sqrt{3}} 4,30$$

(здесь квантиль $t_{1-p/2}=4,30$ найден из таблицы III Приложения соответственно $f=2$ степеням свободы). После вычислений получаем доверительную оценку $17,36 \leq a \leq 19,84$, которая значительно уступает оценке, полученной в предположении, что известная нам дисперсия является генеральной. Этот пример еще раз подчеркивает важность определения именно генеральной дисперсии для получения наиболее узких доверительных интервалов.

В некоторых задачах требуется найти одностороннюю доверительную оценку генерального среднего, т. е. оценку только сверху или только снизу. Такие оценки непосредственно вытекают из общего определения квантилей. А именно, при доверительной вероятности $1-p$ оценка для t сверху имеет вид $t \leq t_{1-p}$, оценка снизу имеет вид $t \geq -t_{1-p}$. Используя выражение для t из формулы (6.4), получим односторонние доверительные оценки генерального среднего:

$$a \leq \bar{x} + \frac{s}{\sqrt{n}} t_{1-p} \quad (\text{сверху}), \quad a \geq \bar{x} - \frac{s}{\sqrt{n}} t_{1-p} \quad (\text{снизу}).$$

Напомним, что в таблице III Приложения указаны квантили $t_{1-p/2}$ для соответствующих уровней значимости p . Поэтому число t_{1-p} нужно искать в столбце с вероятностью $2p$.

Рассмотрим следующий пример. При выплавке чугуна в качестве вредной примеси появляется сера. Шестикратный анализ показал, что в одной тонне выплавленного чугуна содержится $\bar{x}=4$ кг серы. Выборочный стандарт наблюдений s (в подобных случаях он носит название *ошибки воспроизводимости* анализа) найден по тем же шести результатам анализа и равен 0,3 кг. Необходимо найти возможный верхний предел содержания серы при доверительной вероятности $1-p=0,95$.

Обратимся к таблице III Приложения; число степеней свободы в нашем случае равно 5. В столбце с вероятностью $2p=0,10$ находим $t_{0,95}=2,02$. Отсюда

$$a \leq 4 + \frac{0,3}{\sqrt{6}} 2,02 = 4,25,$$

что и дает требуемую оценку.

В приведенных примерах вычислений с помощью распределения Стьюдента предполагалось, что стандарт s вычислен по той же выборке, что и среднее \bar{x} . Подобная ситуация встречается на практике чаще всего, однако для распределения Стьюдента она не обязательна, т. е. числа \bar{x} и s в формуле (6.4) могут быть найдены и по различным выборкам. Это позволяет, в частности, вычислять s по «текущим измерениям» даже и в том случае, когда число этих измерений не настолько велико, чтобы найденное по ним s можно было приравнять (в пределах точности вычислений) генеральному стандарту σ .

6.2. Оценка генеральной дисперсии. Роль дисперсии неоднократно подчеркивалась в предыдущем изложении. Не говоря о том, что знание генеральной дисперсии позволяет получать более удобные оценки генерального среднего (см. предыдущий пункт), дисперсия имеет и самостоятельную ценность, как информация о точности применяемой методики испытаний.

Для оценки генеральной дисперсии используется выборочная дисперсия s^2 . Эта дисперсия в силу случайности выборки сама является случайной величиной; в п. 4.4 было показано, что математическим ожиданием для s^2 служит генеральная дисперсия σ^2 . Отсюда следует, что σ^2 можно оценить по s^2 , если известно распределение величины s^2 .

Распределение величины s^2 можно получить с помощью так называемого *распределения Пирсона* (или χ^2 -распределения), открытого и разработанного Пирсоном в 1900 г. Для выборки с элементами x_1, x_2, \dots, x_n через χ^2 обозначается сумма

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

В этой сумме есть связь \bar{x} , поэтому число степеней свободы $f=n-1$. Плотность χ^2 -распределения зависит только от f , графики плотности при некоторых значениях f приведены на рис. 22. Поскольку $\chi^2 \geq 0$, то и плотность рассматривается лишь на промежутке $[0, \infty)$. Кривые асимметричны, хотя степень асимметрии уменьшается при увеличении f . В связи с этим отдельные квантили величины χ^2 не выражаются друг через друга.

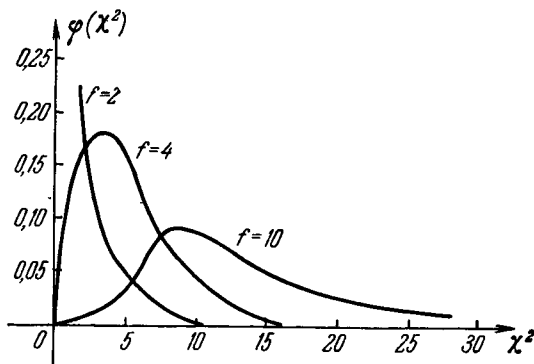


Рис. 22.

При доверительной вероятности $1-p$ двусторонняя доверительная оценка величины χ^2 имеет вид

$$\chi_{p/2}^2 \leq \chi^2 \leq \chi_{1-p/2}^2,$$

односторонние оценки имеют вид

$$\chi^2 \leq \chi_{1-p}^2, \quad \chi^2 \geq \chi_p^2.$$

Квантили χ_{1-p}^2 при различных p и f приведены в таблице IV Приложения.

Нетрудно усмотреть связь между величинами χ^2 и s^2 :

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

откуда

$$\chi^2 = \frac{fs^2}{\sigma^2}.$$

Поэтому с вероятностью $1-p$ справедливо неравенство

$$\chi_{p/2}^2 \leq \frac{fs^2}{\sigma^2} \leq \chi_{1-p/2}^2.$$

Простейшие преобразования приведут нас к соотношению

$$\frac{fs^2}{\chi_{1-p/2}^2} \leq \sigma^2 \leq \frac{fs^2}{\chi_{p/2}^2},$$

которое и является двусторонней доверительной оценкой для генеральной дисперсии σ^2 . Аналогично получаются одно-сторонние доверительные оценки

$$\sigma^2 \leq \frac{fs^2}{\chi_p^2}, \quad \sigma^2 \geq \frac{fs^2}{\chi_{1-p}^2}.$$

Оценим в качестве примера генеральную дисперсию σ^2 для серии из 16 наблюдений с выборочной дисперсией $s^2=2,4$; доверительную вероятность $1-p$ положим равной 0,90. По таблице IV Приложения при числе степеней свободы $f=15$ находим $\chi_{0,95}^2=7,3$ и $\chi_{0,05}^2=25,0$. Это даст двустороннюю доверительную оценку

$$\frac{15}{25,0} 2,4 \leq \sigma^2 \leq \frac{15}{7,3} 2,4$$

или, после вычислений, $1,44 \leq \sigma^2 \leq 4,93$.

Полученные оценки дисперсии можно превратить в оценки стандарта σ , извлекая из всех частей неравенств квадратный корень. Например, двусторонняя доверительная оценка генерального стандарта при доверительной вероятности $1-p$ имеет вид

$$s \sqrt{\frac{f}{\chi_{1-p/2}^2}} \leq \sigma \leq s \sqrt{\frac{f}{\chi_{p/2}^2}}.$$

Вводя в рассмотрение случайную величину

$$v = \sqrt{\frac{f}{\chi^2}},$$

эту двустороннюю оценку можно записать в виде

$$sv_{p/2} \leq \sigma \leq sv_{1-p/2},$$

а соответствующие односторонние — в виде $\sigma \leq sv_{1-p}$, $\sigma \geq sv_p$ (доверительную вероятность по-прежнему считаем равной $1-p$). Для удобства пользования перечисленными оценками в таблице V Приложения приведены квантили

$$v_p = \sqrt{\frac{f}{\chi^2_{1-p}}}.$$

Этими же квантилями можно пользоваться и для оценок генеральной дисперсии, например, с вероятностью $1-p$ справедливо неравенство

$$s^2 v_{p/2}^2 \leq \sigma^2 \leq s^2 v_{1-p/2}^2.$$

Можно показать, что при больших f распределение величины s близко к нормальному с математическим ожиданием σ и дисперсией $\frac{\sigma}{\sqrt{2f}}$. На практике распределение величины s считают нормальным уже при $f \geq 30$.

Отметим, что для определения s^2 и дальнейшей оценки σ^2 можно использовать «текущие измерения» (п. 4.4). При этом нужно помнить, что число степеней свободы f равно общему числу наблюдений, минус число групп. Именно это f и определяет распределение величины χ^2 .

Иногда для оценки генерального стандарта σ удобно использовать размах выборки, т. е. разность между ее наибольшим и наименьшим элементами. Такая оценка менее эффективна, чем оценка с помощью выборочного стандарта s , однако она требует намного меньше вычислений, что нередко оправдывает ее применение.

Обозначим размах выборки через W . Отношение

$$w_n = \frac{W}{\sigma}$$

является случайной величиной, распределение которой зависит только от объема выборки n . Поэтому в качестве приближенного значения σ можно брать величину

$$\frac{W}{Mw_n} = W\alpha_n, \quad (6.5)$$

где Mw_n — математическое ожидание величины w_n . Числа $\alpha_n = \frac{1}{Mw_n}$ приведены в таблице VI Приложения.

Оценка (6.5) мало эффективна и применяется редко. Значительно полезней применение размаха при наличии нескольких выборок из генеральных совокупностей с одинаковым стандартом σ (напомним, что именно такая ситуация возникает при обработке «текущих измерений»). Укажем способ оценки генерального стандарта σ в этом случае.

Допустим, что обработке подлежит некоторая достаточно большая совокупность наблюдений и что все применявшиеся при наблюдениях методики имеют одинаковые генеральные дисперсии σ^2 . Разобьем эту совокупность наблюдений на k отдельных частичных выборок так, чтобы в пределах каждой такой выборки генеральное среднее было неизменным *). Так, например, при обработке «текущих измерений» частичные выборки образуются из измерений одного объекта. Число элементов в выборке с номером i обозначим через n_i ; желательно, чтобы все n_i были меньше 10.

Легче всего производить расчеты, когда все $n_i = n$. Обозначая размах i -й выборки через W_i , найдем величину

$$l = \frac{\alpha_n}{k} \sum_{i=1}^k W_i.$$

При $k \geq 8$ эта величина имеет приближенно нормальное распределение с математическим ожиданием σ и дисперсией $Dl = \frac{\sigma^2}{k} \beta_n^2$. Числа β_n зависят только от n и приведены в таблице VI Приложения.

Таким образом, величину σ можно оценивать через l с помощью квантилей $u_{1-p/2}$ стандартного нормального распределения. При доверительной вероятности $1-p$ получаем оценку

$$l - u_{1-p/2} \frac{\sigma \beta_n}{\sqrt{k}} \leq \sigma \leq l + u_{1-p/2} \frac{\sigma \beta_n}{\sqrt{k}}. \quad (6.6)$$

В этой оценке неизвестный стандарт σ входит во все части неравенств. Решая неравенство (6.6) относительно σ , по-

*) Если генеральное среднее неизменно по всей совокупности, то разбиение нужно проводить случайным образом с помощью таблицы случайных чисел (см. п. 10.1).

лучаем окончательную доверительную оценку генерального стандарта:

$$\frac{l \sqrt{k}}{\sqrt{k} + \beta_{n^{u_{1-p/2}}}} \leq \sigma \leq \frac{l \sqrt{k}}{\sqrt{k} - \beta_{n^{u_{1-p/2}}}}. \quad (6.7)$$

Если частичные выборки имеют различные объемы n_i , то можно снова пользоваться оценкой (6.7), но величина l определится теперь равенством

$$l = \frac{\sum_{i=1}^k W_i \gamma_{n_i}}{\frac{1}{2} + \sum_{i=1}^k \gamma_{n_i}},$$

где числа $\gamma_{n_i} = \frac{1}{\beta_{n_i}^2}$ также приведены в таблице VI Приложения.

6.3. Сравнение дисперсий. Одной из важнейших задач статистической обработки наблюдений является сравнение двух или нескольких выборочных дисперсий. Основной выясняемый вопрос при этом — можно ли считать сравниваемые выборочные дисперсии оценками одной и той же генеральной дисперсии. С такой задачей, в частности, приходится сталкиваться при вычислении дисперсии по «текущим измерениям».

Начнем со сравнения двух дисперсий s_1^2 и s_2^2 , имеющих соответственно f_1 и f_2 степеней свободы. Будем считать, что первая выборка сделана из генеральной совокупности с дисперсией σ_1^2 , а вторая — из генеральной совокупности с дисперсией σ_2^2 . Выдвигается нулевая гипотеза (п. 5.4) — гипотеза о равенстве $\sigma_1^2 = \sigma_2^2$. Для того чтобы отвергнуть эту гипотезу, нужно доказать значимость расхождения между s_1^2 и s_2^2 при выбранном уровне значимости p . В качестве критерия значимости обычно используется так называемое *распределение Фишера*.

Распределением Фишера (или *F-распределением*) называется распределение случайной величины:

$$F = \frac{s_1^2}{\sigma_1^2} : \frac{s_2^2}{\sigma_2^2}.$$

Это распределение зависит только от f_1 и f_2 , при этом $F(f_1, f_2) = \frac{1}{F(f_2, f_1)}$. На рис. 23 приведены графики плотности F -распределения при сочетаниях $(f_1, f_2) = (10, 4)$ и $(10, 50)$.

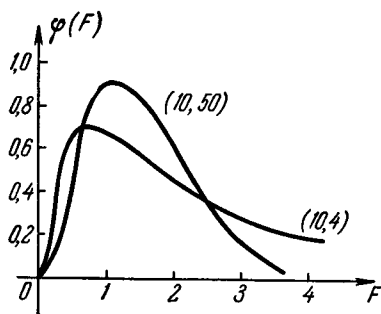


Рис. 23.

Как и в случае χ^2 -распределения, плотность рассматривается лишь на положительной полуоси, т. е. при $0 \leq F < \infty$. Кривые имеют асимметричную форму. В таблице VII Приложения даны квантили F_{1-p} для некоторых наиболее употребительных уровней значимости и различных комбинаций f_1 и f_2 . При нахождении квантилей F_p для значений p , не вошедших в таблицу, используется очевидное соотношение

$$F_p(f_1, f_2) = \frac{1}{F_{1-p}(f_2, f_1)}.$$

Например, $F_{0,95}(4, 3) = 9,1$; $F_{0,05}(3, 4) = \frac{1}{9,1} = 0,11$.

С помощью F -распределения можно находить доверительные оценки для отношения $\frac{\sigma_1^2}{\sigma_2^2}$. Если доверительная вероятность равна $1-p$, то имеем двустороннюю оценку

$$\frac{s_1^2}{s_2^2} \frac{1}{F_{1-p/2}(f_1, f_2)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-p/2}(f_2, f_1)$$

или односторонние оценки

$$\frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-p}(f_2, f_1), \quad \frac{\sigma_1^2}{\sigma_2^2} \geq \frac{s_1^2}{s_2^2} \frac{1}{F_{1-p}(f_1, f_2)}.$$

Вернемся к рассмотрению нулевой гипотезы, согласно которой $\frac{\sigma_1^2}{\sigma_2^2} = 1$. В этом случае $F = \frac{s_1^2}{s_2^2}$ и, следовательно, F -распределение может быть использовано непосредственно

для оценки отношения $\frac{s_1^2}{s_2^2}$. С вероятностью $1-p$ должно выполняться двустороннее неравенство

$$\frac{1}{F_{1-p/2}(f_2, f_1)} \leq \frac{s_1^2}{s_2^2} \leq F_{1-p/2}(f_1, f_2) \quad (6.8)$$

или одно из односторонних неравенств

$$\frac{s_1^2}{s_2^2} \leq F_{1-p}(f_1, f_2), \quad \frac{s_1^2}{s_2^2} \geq \frac{1}{F_{1-p}(f_2, f_1)}.$$

Вероятность любого противоположного неравенства равна уровню значимости p . Значит, согласно принципу значимости (п. 5.3) противоположные неравенства невозможны, если справедлива нулевая гипотеза. Иными словами, противоположные неравенства образуют критическую область проверяемой гипотезы, и если отношение $\frac{s_1^2}{s_2^2}$

на самом деле попадает в эту область, то различие между дисперсиями следует признать значимым.

Сформулируем получившиеся критерии значимости. При этом для удобства изложения будем обозначать через s_1^2 большую из сравниваемых дисперсий. Если большей выборочной дисперсии заведомо не может соответствовать меньшая генеральная, т. е. если неравенство $\sigma_1^2 < \sigma_2^2$ заведомо невозможно, то нужно применять односторонний критерий, сравнивая отношение $\frac{s_1^2}{s_2^2}$ с полученными выше

односторонними доверительными оценками. Нулевая гипотеза отвергается, если $\frac{s_1^2}{s_2^2} > F_{1-p}(f_1, f_2)$, где значение $F_{1-p}(f_1, f_2)$ берется из таблицы VII Приложения.

При неизвестном соотношении между σ_1^2 и σ_2^2 нужно применять двусторонний критерий, т. е. проверять, не нарушается ли установленное выше двустороннее неравенство (6.8). При этом левая часть неравенства (6.8) всегда выполнена, так как для небольших p (а принятый уровень значимости не может быть большим) обязательно

$\frac{1}{F_{1-p/2}(f_2, f_1)} < 1$, в то время как $\frac{s_1^2}{s_2^2} > 1$ по условию.

Поэтому нарушиться может только правая часть неравенства (6.8). Следовательно, при двустороннем критерии значимости отношение $F = \frac{s_1^2}{s_2^2}$ сравнивается только со значением $F_{1-p/2}(f_1, f_2)$ (также взятым из таблицы VII Приложения), и нулевая гипотеза отвергается, если $\frac{s_1^2}{s_2^2} >$

$> F_{1-p/2}(f_1, f_2)$.

Пример. При изучении стабильности температуры в термостате получены данные 21,2; 21,8; 21,3; 21,0; 21,4; 21,3. К стабилизатору температуры применено некоторое усовершенствование, после чего (на другом режиме) получены данные: 37,7; 37,6; 37,6; 37,4. Можно ли при уровне значимости $p=0,05$ считать усовершенствование эффективным?

Эффективность стабилизаторов температуры, очевидно, зависит от даваемой ими дисперсии температур. Таким образом, задача состоит в том, чтобы сравнить генеральные дисперсии данных выборок температур. Вычисляем выборочные дисперсии, уменьшив для удобства вычислений все данные на 21 в первом случае и на 37,5 — во втором:

$$s_1^2 = \frac{1}{5} \left[0,2^2 + 0,8^2 + 0,3^2 + 0,4^2 + 0,3^2 - \frac{(0,2+0,8+0,3+0,4+0,3)^2}{6} \right] = \frac{1}{5} \left(1,02 - \frac{4}{6} \right) = 0,07,$$

$$s_2^2 = \frac{1}{3} \left[0,2^2 + 0,1^2 + 0,1^2 + 0,1^2 - \frac{(0,2+0,1+0,1-0,1)^2}{4} \right] =$$

$$= \frac{1}{3} \left(0,07 - \frac{0,09}{4} \right) = 0,016.$$

Отсюда

$$F = \frac{0,07}{0,016} = 4,4.$$

Числа степеней свободы $f_1=5$, $f_2=3$. Усовершенствование может лишь уменьшить дисперсию, поэтому применяем односторонний критерий значимости. По таблице VII Приложение

ния находим $F_{0,95}(5,3)=9,0$. Мы видим, что $\frac{s_1^2}{s_2^2}=4,4 < 9,0$.

Следовательно, данные наблюдений не позволяют отвергнуть нулевую гипотезу и считать усовершенствование эффективным.

Критерий Фишера можно использовать для сравнения дисперсий и в том случае, когда одна из дисперсий является генеральной. В этом случае ее число степеней свободы считается равным ∞ . Например, сравнивая выборочную дисперсию $s_1^2=0,46$, имеющую $f_1=4$ степеней свободы, с генеральной дисперсией $\sigma_2^2=0,28$, мы в качестве критического значения должны взять из таблицы VII Приложения величину $F_{0,95}(4, \infty)=2,4$. Мы видим, что

$$\frac{s_1^2}{\sigma_2^2} = \frac{0,46}{0,28} = 1,64 < 2,4.$$

Следовательно, различие между выборочной и генеральной дисперсиями не является значимым.

Распределение F с бесконечным числом степеней свободы можно использовать для получения доверительных оценок генеральной дисперсии по выборке объема n . Так, при доверительной вероятности $1-\rho$ справедлива двусторонняя оценка:

$$\frac{s^2}{F_{1-\rho/2}(f_1, \infty)} \leq \sigma^2 \leq \frac{s^2}{F_{\rho/2}(f_1, \infty)} \quad (f_1 = n - 1).$$

К сожалению, полученная оценка не дает ничего нового по сравнению с оценками п. 6.2, ибо справедливо тождество

$$F(f_1, \infty) = \frac{\chi^2(f_1)}{f_1}.$$

Рассмотрим теперь вопрос о сравнении нескольких дисперсий $s_1^2, s_2^2, \dots, s_k^2$, имеющих числа степеней свободы f_1, f_2, \dots, f_k . Требуется выяснить, являются ли числа s_i^2 оценками одной и той же генеральной дисперсии. Сразу же напрашивается мысль использовать F -распределение — сравнить, например, сначала s_1^2 и s_2^2 , затем s_2^2 и s_3^2 и т. д. Однако такой способ сравнения может привести к ошибочному выводу — сравнивая в один прием лишь две дисперсии, мы лишаем себя всей информации остальных дисперсий.

А ведь то, что невозможно на двух случайных объектах (выборках), может стать вполне возможным на большем их числе (чем больше проводится испытаний, тем более редкие события могут произойти).

Кроме того, незначимые различия, накапливаясь от пары к паре, могут стать вполне значимыми, хотя мы этого не заметим. Разумеется, такой ошибки можно избежать, сравнивая сразу самую большую и самую маленькую дисперсию — если уж они различаются незначимо, то и между промежуточными дисперсиями различий нет. Но и этот вывод справедлив, только если все дисперсии определены по выборкам одинаковых объемов.

Таким образом, с помощью F -критерия удастся сравнивать несколько дисперсий только в случае одинаковых чисел степеней свободы у сравниваемых дисперсий. При этом выявить можно только незначимость различий; если же этот критерий показывает, что наибольшая и наименьшая дисперсия различаются значимо, то по отношению ко всем остальным дисперсиям в совокупности вывод о значимом различии может быть неверен. Мы снова сталкиваемся с необходимостью использовать при сравнении полную информацию о всех заданных дисперсиях. Такое квалифицированное сравнение проводится с помощью *критерия Бартлетта*.

Определим вначале средневзвешенную дисперсию

$$s^2 = \frac{\sum_{i=1}^k f_i s_i^2}{f}, \quad \bar{f} = \sum_{i=1}^k f_i.$$

Вычислим величины

$$B = 2,303 \left(f \lg s^2 - \sum_{i=1}^k f_i \lg s_i^2 \right),$$

$$C = 1 + \frac{1}{3(k-1)} \left(\sum_{i=1}^k \frac{1}{f_i} - \frac{1}{\bar{f}} \right).$$

Бартлет показал, что в случае, когда все s_i^2 соответствуют одной генеральной дисперсии, отношение $\frac{B}{C}$ распределено приближенно, как χ^2 с $k-1$ степенями свободы, независимо

от f_i , лишь бы все $f_i \geq 5$. Это значит, что гипотеза о равенстве генеральных дисперсий принимается, если $\frac{B}{C} \leq \chi^2_{1-p}$ при заданном уровне значимости p ; в противном случае различие между дисперсиями $s_1^2, s_2^2, \dots, s_k^2$ нужно считать значимым.

Нетрудно видеть, что всегда $C > 1$. Поэтому вначале вычисляют только B и сравнивают его с χ^2_{1-p} . Если окажется, что $B \leq \chi^2_{1-p}$, то нулевую гипотезу (гипотезу о равенстве дисперсий) нужно принять, ибо и подавно $\frac{B}{C} < B \leq \chi^2_{1-p}$.

Если же $B > \chi^2_{1-p}$, то C придется вычислить, применяя затем критерий Бартлетта полностью.

Формулы для вычисления B и C несколько упрощаются, если все f_i равны между собой. Однако в этом случае существует более удобный (и более точный) критерий Кохрана. Кохран предложил рассматривать отношение максимальной дисперсии к сумме всех остальных,

$$g = \frac{\max s_i^2}{s_1^2 + s_2^2 + \dots + s_k^2},$$

и нашел распределение величины g . Оказалось, что это распределение зависит только от k и от числа f степеней свободы, по которым определена каждая дисперсия s_i^2 .

В таблице VIII Приложения приведены квантили g_{1-p} для $p=0,05$ и $0,01$. Если найденное по заданным дисперсиям значение g окажется больше, чем g_{1-p} (для выбранного уровня значимости), то нулевую гипотезу нужно отбросить и расхождение между дисперсиями считать значимым.

Рассмотрим следующий пример, связанный с обработкой «текущих измерений» (наиболее типичный случай использования критерия Бартлетта). С помощью мостика измеряются электрические сопротивления нескольких проводников. У одних проводников сопротивление мало, у других велико; естественно опасаться, что на разных уровнях сопротивлений мостик обладает разной точностью. Поэтому дисперсии, полученные на разных проводниках, необходимо сравнить.

Данные измерений приведены в таблице 6.1. В этой же таблице приведена часть расчетов для величин B и C .

Таблица 6.1

Номер проводника, i	Кол-во степеней свободы, f_i	Дисперсия, s_i^2	$f_i s_i^2$	$\lg s_i^2$	$f_i \lg s_i^2$	$\frac{1}{f_i}$
1	8	0,17	13,36	$\bar{1},2304$	$\bar{7},8432$	0,1250
2	12	0,40	4,80	$\bar{1},6021$	$\bar{5},2252$	0,0833
3	16	0,38	6,08	$\bar{1},5798$	$\bar{7},2768$	0,0625
4	16	0,62	9,92	$\bar{1},7924$	$\bar{4},6784$	0,0625
5	10	0,54	5,40	$\bar{1},7324$	$\bar{3},3240$	0,1000
Суммы	62		27,56		$\bar{24},3478$	0,4333

Приведем остальные расчеты:

$$s^2 = \frac{27,56}{62} = 0,444, \quad \lg s^2 = \bar{1},6474,$$

$$f \lg s^2 = 62 \cdot \bar{1},6474 = \bar{22},1388, \quad \frac{1}{f} = 0,0161,$$

$$B = 2,303 (\bar{22},1368 - \bar{24},3478) = 2,303 \cdot 1,791 = 4,125,$$

$$C = 1 + \frac{0,4333 - 0,0161}{3(5-1)} = 1,035.$$

По таблице IV Приложения находим, что при четырех степенях свободы $\chi_{0,95}^2 = 9,5$. Мы видим, что $B < \chi_{0,95}^2$ и, значит, на уровне значимости $p=0,05$ гипотезу о равенстве дисперсий нужно принять. Величина C нам не понадобилась, ее вычисление приведено лишь для иллюстрации.

Критерий Бартлета позволяет в разобранным примере утверждать, что генеральная дисперсия измерений на мостике не зависит от измеряемого сопротивления. Следовательно, для оценки этой дисперсии могут быть использованы все 67 измерений. В качестве оценки σ^2 нужно взять средневзвешенную дисперсию s^2 , которая у нас уже вычислена и равна 0,444. Этой дисперсии соответствуют 62 степени свободы, в силу чего можно считать, что $\sigma^2 = s^2$; ошибка такой замены при 62 степенях свободы весьма невелика.

6.4. Сравнение средних. Не менее важной, чем сравнение дисперсий, является задача о сравнении средних. Например, проводя наблюдения одного и того же объекта в

течение длительного времени, мы должны проверять, остается ли при этом неизменным истинное значение измеряемой величины. Другие примеры сравнения средних приводились в п. 5.4.

Пусть заданы две случайные выборки: x_1, x_2, \dots, x_{n_1} и y_1, y_2, \dots, y_{n_2} , взятые из нормально распределенных генеральных совокупностей. Пусть генеральное среднее и генеральная дисперсия первой совокупности равны a_1 и σ_1^2 , второй — a_2 и σ_2^2 . Тогда среднее \bar{x} первой выборки есть нормально распределенная случайная величина с параметрами a_1 и $\frac{\sigma_1^2}{n_1}$, среднее \bar{y} второй выборки есть нормально распределенная случайная величина с параметрами a_2 и $\frac{\sigma_2^2}{n_2}$ (п. 4.4). Рассмотрим случайную величину $\delta = \bar{x} - \bar{y}$.

На основании свойств математического ожидания и дисперсии (п. 2.3) находим, что

$$M\delta = M\bar{x} - M\bar{y} = a_1 - a_2, \quad D\delta = D\bar{x} + D\bar{y} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Более того, распределение величины δ также является нормальным в силу линейности нормального распределения (п. 3.3), поэтому $M\delta$ и $D\delta$ будут параметрами этого распределения. Это позволяет дать оценку разности $a_1 - a_2$ с помощью квантилей стандартного нормального распределения так, как это было сделано в п. 6.1. При доверительной вероятности $1 - p$ мы получим двустороннюю доверительную оценку:

$$\begin{aligned} \bar{x} - \bar{y} - u_{1-p/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} &\leq a_1 - a_2 \leq \\ &\leq \bar{x} - \bar{y} + u_{1-p/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \end{aligned}$$

или односторонние оценки:

$$\begin{aligned} a_1 - a_2 &\leq \bar{x} - \bar{y} + u_{1-p} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \\ a_1 - a_2 &\geq \bar{x} - \bar{y} - u_{1-p} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}. \end{aligned}$$

Если высказывается гипотеза $a_1 = a_2$ (нулевая гипотеза), то полученные оценки могут служить критериями проверки этой гипотезы. Действительно, если нулевая гипотеза верна, то должны выполняться (с вероятностью $1 - p$) все записанные выше неравенства, где вместо $a_1 - a_2$ стоит нуль. Нарушение каждого такого неравенства имеет вероятность p , в силу чего оно является значимым (неслучайным) событием. Таким образом, гипотеза $a_1 = a_2$ отвергается, если выполнено неравенство

$$|\bar{x} - \bar{y}| > u_{1-p/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

при двустороннем критерии или неравенство

$$|\bar{x} - \bar{y}| > u_{1-p} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

при одностороннем критерии. Напомним, что односторонний критерий применяется тогда, когда заранее известно, что большему выборочному среднему (\bar{x} или \bar{y}) не может соответствовать меньшее генеральное среднее.

Указанные критерии просты и надежны, но они требуют знания σ_1^2 и σ_2^2 , что не всегда возможно, особенно при обработке малых выборок. Если генеральные дисперсии σ_1^2 и σ_2^2 неизвестны и можно оперировать только выборочными дисперсиями s_1^2 и s_2^2 , то приходится обращаться к распределению Стьюдента.

Предположим вначале, что $\sigma_1 = \sigma_2 = \sigma$. Это равенство можно проверять по заданным s_1 и s_2 с помощью изложенного в предыдущем пункте критерия Фишера. В этом случае генеральную дисперсию σ^2 можно оценивать средне-взвешенной дисперсией

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

Число степеней свободы у s^2 равно $f = n_1 + n_2 - 2$.

Дисперсия величины δ имеет теперь простой вид

$$D\delta = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right).$$

На основании общих правил (п. 3.2) нормированная случайная величина

$$\frac{\delta - M\delta}{\sqrt{D\delta}} = \frac{(\bar{x} - \bar{y}) - (a_1 - a_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

имеет стандартное нормальное распределение. Если же σ заменить на s , то получится величина

$$t = \frac{(\bar{x} - \bar{y}) - (a_1 - a_2)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

имеющая распределение Стьюдента с f степенями свободы.

Используя квантили $t_{1-p/2}$ из таблицы III Приложения, мы можем теперь записать доверительные оценки для разности $a_1 - a_2$, не содержащие неизвестной генеральной дисперсии σ^2 . При доверительной вероятности $1-p$ имеем двустороннюю оценку

$$\begin{aligned} \bar{x} - \bar{y} - t_{1-p/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} &\leq a_1 - a_2 \leq \\ &\leq \bar{x} - \bar{y} + t_{1-p/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \end{aligned}$$

или односторонние оценки

$$\begin{aligned} a_1 - a_2 &\leq \bar{x} - \bar{y} + t_{1-p} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \\ a_1 - a_2 &\geq \bar{x} - \bar{y} - t_{1-p} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}. \end{aligned}$$

Вернемся вновь к рассмотрению нулевой гипотезы $a_1 = a_2$. Те же рассуждения, что и раньше, дадут нам критерий проверки этой гипотезы с помощью распределения Стьюдента: нулевая гипотеза отвергается, если

$$|\bar{x} - \bar{y}| \geq t_{1-p/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

при двустороннем критерии или

$$|\bar{x} - \bar{y}| \geq t_{1-p} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

при одностороннем критерии. Напомним, что в оценках используются квантили величины t , соответствующие $f = n_1 + n_2 - 2$ степеням свободы.

Сравним с помощью полученного критерия производительность некоторого завода (с установившимся производственным процессом) в дневную и ночную смены. Среднемесячная ($n=30$) выработка дневной смены $\bar{x}=62,7$ тыс. руб., ночной смены $\bar{y}=62,4$ тыс. руб., дисперсии соответственно равны $s_1^2=0,66$ и $s_2^2=0,80$.

Сначала сравним дисперсии по критерию Фишера:

$$F = \frac{s_2^2}{s_1^2} = \frac{0,80}{0,66} = 1,21 < F_{0,95}(29,29) = 1,8.$$

Различие между дисперсиями оказывается незначимым, поэтому можно применять t -критерий. Найдем средневзвешенную дисперсию

$$s^2 = \frac{29s_1^2 + 29s_2^2}{58} = \frac{0,66 + 0,80}{2} = 0,73,$$

откуда $s = \sqrt{0,73} = 0,85$. Число степеней свободы $f=58$. Из таблицы III Приложения находим $t_{0,95}=1,67$, поэтому

$$t_{1-p} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 1,67 \cdot 0,73 \sqrt{\frac{2}{30}} = 0,15.$$

Мы видим, что $\bar{x} - \bar{y} = 0,3 > 0,15$. Значит, при 5%-ном уровне значимости разницу между производительностями дневной и ночной смены нужно считать значимой.

Заметим, что мы здесь рассматривали односторонний критерий, так как средние данные позволяют предполагать, что в ночную смену производительность ниже. Если бы мы не делали этого предположения и применили двусторонний критерий, то пришлось бы взять $t_{0,975}=2,00$. Читателю нетрудно убедиться, что и при двустороннем критерии расхождение средних производительностей является значимым.

Рассмотрим теперь общий случай, когда генеральные дисперсии σ_1^2 и σ_2^2 не равны между собой. Полученные выше точные критерии оказываются в этом случае непригодными. Существует несколько приближенных критериев, позволяющих проверять гипотезу о равенстве генеральных средних

с той или иной точностью. Приведем один, наиболее распространенный и удобный в работе.

Пусть, как и прежде, s_1^2 и s_2^2 есть дисперсии первой и второй выборок и пусть первой дисперсии соответствуют $f_1 = n_1 - 1$, а второй $f_2 = n_2 - 1$ степеней свободы. Вычислим отношения

$$v_1 = \frac{s_1^2}{n_1}, \quad v_2 = \frac{s_2^2}{n_2}.$$

По таблице III Приложения найдем квантили $t_{1-p/2}(f_1)$ (т. е. для f_1 степеней свободы) и $t_{1-p/2}(f_2)$. Наконец, вычислим величину

$$T = \frac{v_1 t_{1-p/2}(f_1) + v_2 t_{1-p/2}(f_2)}{\sqrt{v_1 + v_2}}.$$

Нулевая гипотеза $a_1 = a_2$ отклоняется, если окажется, что $|\bar{x} - \bar{y}| > T$. Заметим, что сформулированный критерий является двусторонним и превращается в односторонний при замене $\frac{p}{2}$ на p .

Рассмотрим в качестве примера результаты опытов, сделанных в начале нашего столетия английским физиком Рэлеем. Он сравнивал вес одного и того же объема азота, полученного после тщательнейшей химической очистки из азотистых соединений (в дальнейшем обозначается через x) и из воздуха (обозначается через y). Измерения проводились при неизменных условиях (15°C и 760 мм рт. ст.), результаты измерений приведены в таблице 6.2. Непосредственный подсчет средних дает $\bar{x} = 2299,47 \text{ мг}$ и $\bar{y} = 2310,16 \text{ мг}$, откуда $|\bar{x} - \bar{y}| = 10,69 \text{ мг}$.

Таблица 6.2

x	y	x	y
2301,43	2310,17	2299,40	2310,10
2298,90	2309,86	2298,49	2310,28
2298,16	2310,10	2298,89	2310,35
2301,82	2310,01		2310,26
2298,69	2310,24		2310,24

Таким образом, получается, что, несмотря на тщательнейшую очистку, азот воздуха почти на $\frac{1}{200}$ тяжелее азота, полученного из соединений.

Величина 10,69 невелика по сравнению с самими \bar{x} и \bar{y} , поэтому можно предположить, что она связана с неточностью измерений, с недостаточной очисткой азота и другими случайными причинами. Имея данные Рэлея, мы можем проверить, действительно ли расхождение $|\bar{x}-\bar{y}|$ случайно или оно значимо. В качестве уровня значимости возьмем самый низкий из имеющихся в таблице III Приложения, т. е. $p=0,001$. Этот уровень позволяет считать значимыми (неслучайными) примерно в 50 раз меньше событий, чем общепринятый уровень 0,05. Благодаря такому выбору уровня значимости мы будем максимально застрахованы от ошибки первого рода (п. 5.3) — возможности отвергнуть гипотезу о равенстве средних, в то время как она верна.

Итак, приступим к проверке гипотезы $a_1=a_2$. Непосредственный подсчет дает значения $s_1^2=1,9022$ и $s_2^2=0,01931$ с $f_1=7$ и $f_2=9$ степенями свободы. Отношение этих дисперсий $F = \frac{s_1^2}{s_2^2} = 98,5$ намного превосходит табличное значение

$F_{0,999}(7,9)=10,8$, поэтому дисперсии измерения нужно считать различными и для сравнения средних использовать приближенный критерий. Находим $v_1 = \frac{1,9022}{8} = 0,2388$, $v_2 = \frac{0,01931}{10} = 0,001931$. Из таблицы III Приложения берем значения $t_{0,9995}(7)=5,41$, $t_{0,9995}(9)=4,78$. Отсюда

$$T = \frac{0,2388 \cdot 5,41 + 0,001931 \cdot 4,78}{\sqrt{0,2388 + 0,001931}} = \frac{1,3011}{0,49} = 2,65,$$

и значит, $|\bar{x}-\bar{y}| > T$. Гипотезу о равенстве средних нужно отбросить (ее вероятность меньше $p=0,001$).

Из проведенного сравнения средних вытекает, что расхождение между весом азота из воздуха и из соединений не является случайным. Это заставило ученых более внимательно исследовать азот, полученный из воздуха, и вскоре было показано, что в нем содержится небольшое количество газа, более тяжелого, чем азот, — аргона.

Нам осталось рассмотреть вопрос о сравнении нескольких средних. Попарное сравнение здесь невозможно по тем же соображениям, что и при сравнении дисперсий (см. предыдущий пункт). Сравнение средних в целом является довольно трудоемкой задачей. Мы будем его проводить лишь в предположении, что дисперсии соответствующих выборок незначимо отличаются друг от друга (в целом); сравнение этих дисперсий можно произвести по критерию Бартлета или Кохрана.

Обозначим сравниваемые средние через $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$, соответствующие выборочные дисперсии — через $s_1^2, s_2^2, \dots, s_k^2$. Числа степеней свободы обозначим через f_1, f_2, \dots, f_k (как обычно, $f_i = n_i - 1$, где n_i — объем соответствующей выборки). Всем выборкам, по условию, соответствует единая генеральная дисперсия σ^2 ; в качестве ее оценки можно взять средневзвешенную дисперсию

$$s^2 = \frac{\sum_{i=1}^k f_i s_i^2}{f}, \quad f = \sum_{i=1}^k f_i,$$

которой соответствует f степеней свободы.

Если справедлива нулевая гипотеза о равенстве всех генеральных средних, то в качестве оценки этого единого генерального среднего α можно взять общее среднее всех элементов, как бы объединенных в одну выборку; обозначим это среднее через \bar{x} . Для дисперсии σ^2 можно теперь дать другую оценку,

$$\bar{s}^2 = \frac{1}{k-1} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2,$$

которой соответствует $k-1$ степеней свободы. Но тогда отношение дисперсий $\frac{\bar{s}^2}{s^2}$ должно быть распределено как $F(k-1, f)$, т. е. при доверительной вероятности $1-p$ должна быть справедлива оценка

$$\frac{\bar{s}^2}{s^2} \leq F_{1-p}(k-1, f).$$

Из сказанного немедленно вытекает критерий сравнения средних. А именно, на уровне значимости p гипотезу о равенстве генеральных средних в совокупности нужно отклонить, если

$$\frac{\bar{s}^2}{s^2} > F_{1-p}(k-1, f).$$

6.5. Проверка однородности наблюдений. Одним из важнейших условий правильного применения статистических оценок является отсутствие грубых ошибок при наблюдениях. Поэтому все грубые ошибки должны быть выявлены и исключены из рассмотрения в самом начале обработки наблюдений.

Единственным достаточно надежным способом выявления грубых ошибок является тщательный анализ условий самих испытаний. При этом наблюдения, проводившиеся в нарушенных условиях, должны отбрасываться, независимо от их результата. Например, если при проведении эксперимента, связанного с электричеством, в лаборатории на некоторое время был выключен ток, то весь эксперимент обязательно нужно проводить заново, хотя результат, быть может, не сильно отличается от предыдущих измерений. Точно так же отбрасываются результаты измерений на фотопластинках с поврежденной эмульсией и вообще на любых образцах с обнаруженным позднее дефектом.

На практике, однако, не всегда удастся провести подобный анализ условий испытания. Чаще всего приходится иметь дело с окончательным цифровым материалом, в котором отдельные данные вызывают сомнение лишь своим значительным отклонением от остальных. При этом сама «значительность» отклонения во многом субъективна — зачастую приходится сталкиваться со случаями, когда экспериментатор отбрасывает не понравившиеся ему наблюдения как ошибочные только потому, что они нарушают уже созданную им в воображении картину исследуемого процесса.

Строгий научный анализ готового ряда наблюдений может быть проведен лишь статистическим путем, причем должен быть достаточно хорошо известен характер распределения наблюдаемой случайной величины. Мы, как и всюду в настоящем параграфе, будем исходить из нормаль-

ного распределения. Каждая грубая ошибка будет соответствовать нарушению этого распределения, изменению его параметров, иными словами, нарушится однородность испытаний (или, как мы будем говорить, *однородность наблюдений*). Поэтому выявление грубых ошибок можно трактовать как проверку однородности наблюдений.

В самой общей статистической форме задача об однородности наблюдений формулируется следующим образом: совместимы ли (на уровне значимости p) элементы выборки x_1, x_2, \dots, x_n с гипотезой о том, что все они извлечены из одной и той же генеральной совокупности, имеющей нормальное распределение. Для решения этой задачи существует много методов. Если сомнение вызывает большое число элементов, то их нужно выделить в отдельную выборку и сравнить средние и дисперсии получившихся двух выборок; иногда такое разбиение приводит к образованию нескольких самостоятельных выборок. Если же сомнительными являются только один или два элемента (именно эта ситуация характерна при выявлении грубых ошибок), то задача об однородности наблюдений решается с помощью некоторых специальных распределений, о которых и пойдет речь ниже.

Если условия проведения испытаний не удастся контролировать на должном уровне, то единственным показателем ошибочности элемента в выборке может служить лишь его отклонение от остальных наблюдений. Поэтому сомнительными, как правило, бывают самый больший и самый малый элементы выборки. В силу симметричности нормального распределения они исследуются абсолютно одинаково, и в дальнейшем мы будем говорить просто о «крайнем» элементе выборки.

Предположим вначале, что нам известны оба параметра a и σ генеральной совокупности — такая ситуация встречается обычно при проведении контрольных испытаний. В этом случае при доверительной вероятности $1-p$ для первого же наблюдения x справедлива доверительная оценка *)

$$|x - a| \leq u_{1-p} \sigma, \quad (6.9)$$

*) Здесь применяется односторонняя оценка, так как по заданному наблюдению всегда можно увидеть, в какую сторону от a оно отклоняется.

где u_{1-p} — квантиль стандартного нормального распределения.

Таким образом, зная параметры a и σ , мы можем даже единственное наблюдение x считать грубо ошибочным (на уровне значимости p), если

$$|x - a| > u_{1-p} \sigma. \quad (6.10)$$

Анализ выборки из нескольких элементов будет уже более сложным. Его нельзя проводить, исследуя каждый элемент выборки отдельно по формуле (6.10), так как при увеличении объема выборки увеличивается вероятность больших отклонений. Это обстоятельство хорошо знакомо многим экспериментаторам: пока серия наблюдений невелика, отдельные наблюдения хорошо согласуются друг с другом; если же наблюдений становится много, появляются и большие расхождения между ними.

Указанное обстоятельство хорошо объясняется и теоретически. Действительно, если вероятность неравенства (6.9) для одного элемента выборки равна $1-p$, то вероятность того, что все n элементов выборки удовлетворяют тому же неравенству, есть вероятность совместного осуществления n независимых событий и потому (см. п. 1.3) равна $(1-p)^n$. Как бы ни был мал уровень значимости p , всегда $(1-p)^n \rightarrow 0$ при $n \rightarrow \infty$. А это значит, что при достаточно большом объеме выборки станет весьма вероятным появление наблюдений, не удовлетворяющих неравенству (6.9), т. е. отклоняющихся от a больше, чем на $u_{1-p} \sigma$.

Пусть теперь x — крайний элемент выборки объема n . Тогда, по выше доказанному, доверительной оценке (6.9) соответствует вероятность

$$(1-p)^n \approx 1 - np$$

(в силу малости p). Поэтому доверительной вероятности $1-p$ соответствует оценка

$$|x - a| \leq u_{1-p/n} \sigma.$$

Отсюда и вытекает критерий проверки крайнего элемента: элемент x считается грубо ошибочным (на уровне значимости p), если

$$|x - a| > u_{1-p/n} \sigma.$$

Одновременное знание обоих параметров a и σ редко встречается в экспериментальной практике. Если известен один параметр σ (например, по предыдущим «текущим измерениям»), то в качестве второго параметра a можно взять среднее \bar{x} исследуемой выборки. При этом возникает такое же «смещение», как и при вычислении выборочной дисперсии (п. 4.4), поэтому вводится поправочный множитель $\sqrt{\frac{n-1}{n}}$.

Крайний элемент x считается теперь несовместимым с остальными элементами выборки (грубо ошибочным), если выполняется неравенство

$$|x - \bar{x}| > u_{1-p/n} \sigma \sqrt{\frac{n-1}{n}}.$$

Перейдем к рассмотрению самого распространенного случая, когда неизвестны оба параметра, a и σ . Одновременная замена этих параметров выборочными параметрами \bar{x} и s заставляет отказаться от использования квантилей нормального распределения.

Если x — крайний элемент выборки, по которой подсчитываются \bar{x} и s , то величина

$$\tau = \frac{|x - \bar{x}|}{s},$$

называемая *максимальным относительным отклонением*, имеет специальное распределение, которое зависит только от объема выборки n . В таблице IX Приложения приведены квантили τ_{1-p} этого распределения при различных n .

С помощью τ -распределения можно получить критерий совместимости крайнего элемента с остальными, не используя никаких других сведений, кроме самой выборки. Согласно этому критерию крайнее значение x отбрасывается как грубо ошибочное (на уровне значимости p), если

$$\frac{|x - \bar{x}|}{s} > \tau_{1-p}.$$

Проиллюстрируем применение τ -критерия на следующем примере. При восьмикратном исследовании стального бруска на разрыв получено среднее значение предела прочности

$\bar{x}=62$ кг/мм², причем $s=1,3$. Минимальное значение 58,6, полученное при одном из наблюдений, вызывает сомнение. Для его проверки подсчитаем максимальное относительное отклонение

$$\tau = \frac{62 - 58,6}{1,3} = 2,62.$$

По таблице IX Приложения для $n=8$ находим $\tau_{0,05}=2,17$, т. е. $\tau > \tau_{0,05}$. Следовательно, на уровне значимости $p=0,05$ значение 58,6 нужно считать ошибочным, исключив его из соответствующих расчетов (в частности, необходимо заново вычислить \bar{x} и s).

Иногда сомнения вызывают одновременно два или даже три элемента выборки. Соответствующая проверка может быть проведена следующим образом. Для каждого из сомнительных элементов вычисляется значение максимального относительного отклонения τ и исследование начинается с элемента, имеющего наименьшее значение τ ; остальные сомнительные элементы временно исключаются из рассмотрения. В оставшейся выборке имеется теперь только один сомнительный элемент. Вычисляя для этой уменьшенной выборки заново \bar{x} и s , мы вычисляем новое значение τ для сомнительного элемента. Если это τ превосходит табличное значение τ_{1-p} (в качестве n берется первоначальный объем выборки до всех отбрасываний), то сомнительное значение является ошибочным. Тем более ошибочными будут и остальные, ранее отброшенные сомнительные элементы.

Если «наименее сомнительный» элемент не оказался ошибочным, то его присоединяют к выборке и начинают исследовать следующий по «сомнительности» элемент: снова вычисляют \bar{x} , s , τ и т. д.

Рассмотренные в этом пункте критерии однородности наблюдений можно применять не только для выявления грубых ошибок, но и для доказательства неслучайности некоторого наблюдения. Например, исследуя урожайность нескольких сортов пшеницы и получив на некотором сорте максимальный урожай, мы можем объединить все данные урожайности в одну выборку и исследовать максимальный урожай с помощью τ -критерия. Если окажется, что мак-

симальное значение урожайности несовместимо с остальными значениями, то это будет служить доказательством неслучайного появления такого максимума, т. е. доказательством преимуществ сорта с максимальной урожайностью.

Этот же t -критерий можно использовать для приближенного сравнения нескольких выборочных средних, найденных по выборкам одинакового объема. Действительно, если все выборки принадлежат к единой генеральной совокупности с параметрами a и σ , то и все средние должны принадлежать к единой генеральной совокупности с параметрами a и $\frac{\sigma}{\sqrt{n}}$. Проверяя с помощью максимального относительного отклонения крайнее значение среди всех выборочных средних, мы и сможем проверить гипотезу о случайном различии выборочных средних.

§ 7. АНАЛИЗ РАСПРЕДЕЛЕНИЯ НАБЛЮДЕНИЙ

7.1. Проверка основной гипотезы. Все оценки предыдущего параграфа существенно опираются на нормальность наблюдаемого распределения и в случае другого распределения могут быть не справедливы. Поэтому применение этих оценок допустимо лишь при достаточной уверенности, что наблюдаемое распределение нормально или по крайней мере близко к нормальному.

Гипотезу о нормальности изучаемого распределения мы выше (п. 4.2) называли *основной гипотезой*. Разумеется, самая надежная проверка этой гипотезы состоит в тщательном анализе условий испытаний; принципы такого анализа изложены в п. 4.2. Однако подобный анализ не всегда возможен, в других случаях он может вызывать сомнения. В такой ситуации гипотезу нормальности приходится проверять непосредственно по наблюдениям (по выборке), используя так называемые *критерии согласия*.

Основной принцип критериев согласия заключается в том, что заданная выборка сравнивается с некоторым заранее намеченным теоретическим распределением. Применение критериев согласия в той или иной форме зависит при этом от требований, предъявляемых к теоретическому распределению. Например, при исследовании согласия с нормальным распределением мы можем задать не только типом распределения, но и какими-то готовыми параметрами μ и σ . Совсем другая картина будет, если мы зададимся только нормальностью распределения, а параметры возьмем из выборки. Именно этот, второй случай наиболее характерен для практики, где для выяснения распределения нет никаких других данных, кроме самой выборки. Поэтому в дальнейшем мы в качестве параметров теоретического распределения всегда будем брать параметры заданной выборки: \bar{x} вместо μ и s вместо σ .

Простейшие критерии согласия основаны на сравнении некоторых генеральных параметров предполагаемого теоретического распределения и их оценок, полученных по исследуемой выборке. Нормальное распределение полностью определяется параметрами μ и σ , поэтому прочие параметры нормально распределенной генеральной совокупности выражаются через μ и σ , и их можно заранее теоретически рассчитать. В то же время для этих параметров можно найти оценки непосредственно по заданной выборке. Если выборочные оценки окажутся достаточно хорошо согласованными с теоретически вычисленными значениями параметров, то это может служить основанием для принятия основной гипотезы; в противном случае основную гипотезу нужно отвергнуть.

В качестве оцениваемых параметров удобнее всего брать моменты (п. 5.1), которые нетрудно вычислить для нормального распределения. В частности,

$$m_3 = 0, \quad m_4 = 3\sigma^4. \quad (7.1)$$

Величины

$$A = \frac{m_3}{\sigma^3}, \quad E = \frac{m_4}{\sigma^4} - 3$$

называются соответственно, *асимметрией* и *эксцессом* распределения. Из (7.1) сразу же следует, что для нормального распределения асимметрия и эксцесс равны нулю. Названия асимметрии и эксцесса отражают связь этих величин с формой графика плотности распределения. На рис. 24 и 25 приведены примеры графиков плотности распределений с ненулевыми асимметрией и эксцессом. Для сравнения штриховой линией изображена нормальная кривая с теми же математическим ожиданием μ и дисперсией σ^2 .

Если в качестве m_3 и m_4 взять моменты исследуемой выборки, то получатся *выборочные* асимметрия и эксцесс. Формулы для их расчета имеют вид

$$A = \frac{1}{ns^3} \sum_{i=1}^n (x_i - \bar{x})^3, \quad E = \frac{1}{ns^4} \sum_{i=1}^n (x_i - \bar{x})^4 - 3.$$

Выборочные асимметрия и эксцесс, как и все выборочные параметры, являются случайными величинами и поэтому даже для нормального распределения могут отличаться от

нуля. К сожалению, применить здесь общие методы критериев значимости трудно, так как распределения асимметрии и эксцесса очень сложны и мало изучены. Однако известны дисперсии этих величин,

$$D(A) = \frac{6(n-1)}{(n+1)(n+3)}, \quad D(E) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)},$$

где n — объем исследуемой выборки.

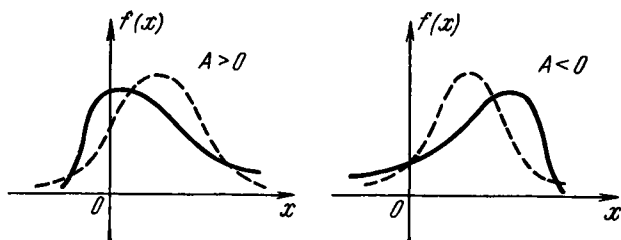


Рис. 24.

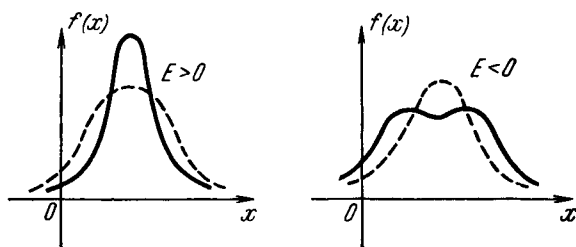


Рис. 25.

Дисперсия служит мерой рассеяния для любого распределения, в силу чего даже знание одной только дисперсии позволяет оценивать вероятности тех или иных значений исследуемой случайной величины (см. об этом подробнее в следующем пункте). Поэтому по дисперсиям $D(A)$ и $D(E)$ можно оценивать, значимо ли выборочные асимметрия и эксцесс отклоняются от своих математических ожиданий, т. е. от нуля.

Мы получаем следующий критерий согласия: *если выборочные асимметрия и эксцесс удовлетворяют неравенствам *)*

$$|A| \leq 3 \sqrt{D(A)}, \quad |E| \leq 5 \sqrt{D(E)},$$

то наблюдаемое распределение можно считать нормальным. В противном случае гипотезу нормальности следует отвергнуть или по крайней мере считать сомнительной.

Сформулированный критерий является весьма приближенным, поэтому его применение оправдано лишь при небольших объемах выборки. Если же выборка достаточно велика ($n \geq 20$), то рекомендуется применять более общие критерии согласия, использующие не только одну-две общие характеристики, но и все элементы выборки.

Простой критерий согласия получится, если использовать понятие эмпирического распределения (п. 4.3). Для этого по выборке строится выборочная функция распределения, согласно правилам, полученным в п. 4.3, и на миллиметровой бумаге вычерчивается ее график. Потом на этом же чертеже вычерчивается график функции нормального распределения

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x - \bar{x}}{s}\right),$$

где Φ есть функция Лапласа, значения которой даны в таблице I Приложения. Расхождение между эмпирической и теоретической функциями распределения оценивается на глаз; если оно невелико, то можно принять основную гипотезу.

Сравнение графиков очень удобно своей наглядностью, однако не является достаточно строгим из-за отсутствия надежной количественной оценки. Положение можно исправить, если воспользоваться *теоремой Колмогорова*. А. Н. Колмогоров доказал, что величина $\lambda = D\sqrt{n}$, где n — объем выборки, D — максимальная абсолютная величина разности между теоретической и эмпирической функциями распределения, имеет при больших n приближенную

*) Мотивировка выбора коэффициентов 3 и 5 дается в следующем пункте.

функцию распределения

$$K(y) = P\{\lambda < y\} = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 y^2} \quad (y > 0).$$

В таблице X Приложения приведены квантили λ_{1-p} этого распределения.

Мы получаем следующий критерий согласия. По чертежу или непосредственно по вычислениям нужно найти

$$D = \max |F_n(x) - F(x)|,$$

а затем $\lambda = D\sqrt{n}$. Если вычисленное значение λ меньше, чем найденное по выбранному уровню значимости p табличное значение λ_{1-p} , то основная гипотеза считается справедливой. В случае $\lambda \geq \lambda_{1-p}$ эта гипотеза отклоняется (или считается сомнительной).

Теорема Колмогорова справедлива не только для нормального, но и для любого теоретического распределения, лишь бы оно было непрерывным. В силу этой, а также ряда других причин в описанном критерии согласия берут очень «жесткие» уровни значимости $p=0,2$ или даже $p=0,3$.

Если число элементов в выборке очень велико (порядка сотни и выше), то все вычисления (в том числе вычисление \bar{x} и s) становятся очень громоздкими. В этом случае можно использовать «метод сгруппированных данных», который заключается в следующем. Вся область изменения выборки, т. е. отрезок между ее минимальным и максимальным элементами, разбивается на интервалы одинаковой длины h . Число интервалов k берут обычно в зависимости от объема выборки в пределах от 8 до 20.

В каждый получившийся интервал попадает какое-то число элементов выборки; число элементов, попавших в i -й интервал, обозначим через n_i . Вместо прежней выборки мы можем теперь рассматривать совокупность интервалов. Выбирая в качестве «представителя» интервала его середину, мы получим новую, «сгруппированную» выборку x_1, x_2, \dots, x_k ; объем этой выборки значительно меньше, чем у первоначальной, но каждый i -й элемент здесь нужно повторять n_i раз. Мы как бы перестаем различать элементы прежней выборки, отличающиеся меньше, чем на h , заменяя их все серединой соответствующего интервала.

Отметим, что подобная «группировка» происходит и в естественных условиях, так как у каждого измерительного прибора есть свой предел разрешающей способности, меньше которого он не различает. Благодаря этому в серии наблюдений часто появляются одинаковые наблюдения, хотя теоретически вероятность двух одинаковых наблюдений равна нулю. Отличие естественной группировки от описанной выше «искусственной» в том, что при естественной группировке величина интервала группирования не выше точности наблюдений, в то время как при «искусственной» группировке, объединяя отдельные наблюдения в один интервал, мы не меняем их точности, которая в результате оказывается намного меньше длины интервала h .

Среднее и дисперсию прежней выборки можно теперь с достаточной степенью точности вычислить по новым, «сгруппированным» данным с помощью формул

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k n_i x_i, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^k n_i (x_i - \bar{x})^2 - \frac{h^2}{12}.$$

Величина $\frac{h^2}{12}$ называется *поправкой Шеппарда*; она связана со смещением дисперсии при группировании. Причиной этого смещения как раз является превышение длины интервала h над точностью данных; именно поэтому поправка Шеппарда не нужна при естественной группировке.

Группирование данных значительно облегчает работу по вычислению \bar{x} и s^2 , а погрешность при этом, как правило, невелика. В связи с этим «метод сгруппированных данных» следует применять и в обычной обработке наблюдений, если объем этих наблюдений очень велик.

С группированием данных связан еще один, пожалуй, самый строгий и надежный критерий согласия, называемый обычно *критерием Пирсона*. Дело в том, что гипотеза о нормальном характере распределения позволяет вычислить теоретические значения для вероятностей p_i попасть в i -й интервал. Для этого используется выведенная в п. 3.2 формула

$$P\{\alpha \leq \xi \leq \beta\} = \Phi\left(\frac{\beta - \bar{x}}{s}\right) - \Phi\left(\frac{\alpha - \bar{x}}{s}\right);$$

здесь α и β — концы рассматриваемого интервала, а значения функции Лапласа Φ даны в таблице I Приложения. После того как найдено p_i , можно подсчитать наивероятнейшее число попаданий в i -й интервал — согласно п. 2.3 оно равно np_i .

Для сравнения эмпирического распределения с предполагаемым нормальным можно теперь сравнить числа n_i и np_i . Оказывается, при условии, что все $np_i \geq 5$, величина

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}$$

имеет приближенно χ^2 -распределение с $f=k-3$ степенями свободы *). Поэтому, выбрав уровень значимости p и найдя в таблице IV Приложения значение χ^2_{1-p} (с f степенями свободы), мы должны гипотезу нормальности отвергнуть при $\chi^2 \geq \chi^2_{1-p}$ и считать правильной при $\chi^2 < \chi^2_{1-p}$.

Отметим особенности группирования в критерии Пирсона. Интервалы здесь не обязательно брать одинаковые по длине (это, правда, затруднит вычисление \bar{x} и s). При подсчете теоретических вероятностей p_i нужно считать, что крайний левый интервал простирается до $-\infty$, крайний правый — до $+\infty$. Кроме того, несколько крайних интервалов, лежащих с одной стороны, объединяются в один, если в них по отдельности не выполняется неравенство $np_i \geq 5$.

В качестве иллюстрации применим все описанные критерии согласия к анализу распределения диаметров валиков (в мм), изготовленных на токарном станке. Для анализа отберем $n=200$ валиков; результаты сгруппируем в 10 интервалов длины $h=0,2$ мм. Сгруппированные данные приведены в таблице 7.1 в первых трех колонках. Остальные колонки служат для вычисления среднего, дисперсии, асимметрии и эксцесса; их данные позволяют вычислить s^2 , A и E , не вычисляя отклонений $x_i - \bar{x}$.

*) Действительно, здесь имеется три связи. Две из них — это \bar{x} и s , третья связь заключена в равенстве $\sum_{i=1}^k p_i = 1$.

Таблица 7.1

i	x_i	n_i	$n_i x_i$	$n_i x_i^2$	$n_i x_i^3$	$n_i x_i^4$
1	3,2	1	3,2	10,24	32,77	104,8
2	3,4	5	17,0	57,80	196,52	668,2
3	3,6	4	14,4	51,84	186,62	671,8
4	3,8	18	68,4	259,92	987,70	3753,2
5	4,0	86	344,0	1376,00	5504,00	22016,0
6	4,2	62	260,4	1093,68	4593,46	19292,5
7	4,4	14	61,6	271,04	1192,58	5247,3
8	4,6	6	27,6	126,96	584,02	2686,5
9	4,8	3	14,4	69,12	331,78	1592,5
10	5,0	1	5,0	25,00	125,00	625,0
Суммы		200	816,0	3341,60	13610,70	56657,8

Используя суммы, получившиеся в последней строке таблицы, находим

$$\begin{aligned}\bar{x} &= \frac{1}{n} \sum n_i x_i = \frac{816}{200} = 4,08, \\ s^2 &= \frac{1}{n-1} \left[\sum n_i x_i^2 - \frac{(\sum n_i x_i)^2}{n} \right] - \frac{h^2}{12} = \\ &= \frac{1}{199} \left(3341,6 - \frac{816^2}{200} \right) - \frac{0,04}{12} \approx 0,06,\end{aligned}$$

откуда $s \approx 0,25$. Асимметрию и эксцесс будем вычислять по следующим непосредственно проверяемым формулам:

$$\begin{aligned}A &= \frac{1}{s^3} \left[\frac{\sum n_i x_i^3}{n} - \frac{3 \sum n_i x_i^2 \sum n_i x_i}{n^2} + \frac{2 (\sum n_i x_i)^3}{n^3} \right], \\ E &= \frac{1}{s^4} \left[\frac{\sum n_i x_i^4}{n} - \frac{4 \sum n_i x_i^3 \sum n_i x_i}{n^2} + \frac{6 \sum n_i x_i^2 (\sum n_i x_i)^2}{n^3} - \right. \\ &\quad \left. - \frac{3 (\sum n_i x_i)^4}{n^4} \right] - 3.\end{aligned}$$

В нашем примере получаем

$$A = \frac{1}{0,25^3} \left[\frac{13610,7}{200} - \frac{3 \cdot 3341,6 \cdot 816}{200^2} + \frac{2 \cdot 816^3}{200^3} \right] = -217,5,$$

$$E = \frac{1}{0,25^4} \left[\frac{56657,8}{200} - \frac{4 \cdot 13610,7 \cdot 816}{200^2} + \frac{6 \cdot 3341,6 \cdot 816^2}{200^3} - \frac{3 \cdot 816^4}{200^4} \right] - 3 = 255,6.$$

В качестве первого критерия согласия сравним найденные асимметрию и эксцесс с их теоретическими дисперсиями:

$$D(A) = \frac{6(200-1)}{(200+1)(200+3)} = 0,0293, \quad \sqrt{D(A)} \approx 0,17,$$

$$D(E) = \frac{24 \cdot 200(200-2)(200-3)}{(200+1)^2(200+3)(200+5)} = 0,113, \quad \sqrt{D(E)} \approx 0,34.$$

Мы видим, что найденные асимметрия и эксцесс во много раз превосходят свои средние квадратичные отклонения. Поэтому нужно считать, что изучаемое распределение существенно отличается от нормального.

Проверим полученный вывод с помощью критериев Колмогорова и Пирсона. Составим новую таблицу 7.2, в которой первые три колонки оставим прежние, а остальные колонки используем для новых вычислений. Четвертая колонка содержит теоретические числа np_i попаданий в соответствующие интервалы. Вероятности p_i вычислены в предполо-

Таблица 7.2

i	x_i	n_i	np_i	$nF_n(x)$	$nF(x)$	$n F_n(x) - F(x) $	$\frac{(n_i - np_i)^2}{np_i}$
1	3,2	1	0,2	1	0,2	0,8	} 0,65
2	3,4	5	1,9	6	2,1	3,9	
3	3,6	4	10,8	10	12,9	2,9	
4	3,8	18	34,3	28	47,2	19,2	7,75
5	4,0	86	59,2	114	106,4	7,6	12,13
6	4,2	62	55,7	176	162,1	13,9	0,71
7	4,4	14	28,6	190	190,7	0,7	7,45
8	4,6	6	8,0	196	198,7	2,7	} 0,05
9	4,8	3	1,2	199	199,9	0,9	
10	5,0	1	0,1	200	200,0	0	

Итого: $\chi^2 = 28,74$

жении, что генеральная совокупность имеет нормальное распределение, причем в качестве параметров взяты найденные выше \bar{x} и s . Покажем, например, как вычислено np_4 , соответствующее середине четвертого интервала $x_4=3,8$. Границами четвертого интервала будут числа

$$\alpha = x_4 - \frac{h}{2} = 3,7, \quad \beta = x_4 + \frac{h}{2} = 3,9.$$

Отсюда по таблице I Приложения получим

$$\begin{aligned} p_4 &= P\{\alpha \leq \xi \leq \beta\} = \Phi\left(\frac{3,9-4,08}{0,25}\right) - \Phi\left(\frac{3,7-4,08}{0,25}\right) = \\ &= \Phi(-0,72) - \Phi(-1,52) = -0,2642 + 0,4357 = 0,1715 \end{aligned}$$

и после умножения,

$$np_4 = 200 \cdot 0,1715 = 34,3.$$

Пятая и шестая колонки содержат значения эмпирической и теоретической функций распределения, умноженные на $n=200$. Эти значения легко получить «накоплением» данных из третьей и четвертой колонок, т. е. складывая все n_i (или np_i) из рассматриваемого интервала и всех ему предшествующих. Например, в пятой строке колонки $nF_n(x)$ нужно записать сумму $1+5+4+18+86=114$, а в третьей строке колонки $nF(x)$ сумму $0,2+1,9+10,8=12,9$. Чтобы получить значения функций $F_n(x)$ и $F(x)$ в чистом виде, достаточно числа из пятой и шестой колонок поделить на $n=200$. Однако эти значения нам не понадобятся и поэтому в таблице 7.2 не приводятся.

Применим к исследуемому распределению критерий согласия Колмогорова. Для этого нужно вычислить разности $n|F_n(x) - F(x)|$ (они записаны в седьмой колонке таблицы 7.2) и среди них выбрать наибольшую, которая даст нам nD . В нашем примере $nD=19,2$. После этого находим

$$\lambda = \frac{nD}{\sqrt{n}} = \frac{19,2}{\sqrt{200}} = 1,36.$$

Мы уже говорили о том, что в критерии Колмогорова нужно брать «жесткий» уровень значимости, поэтому положим $p=0,2$. По таблице X Приложения находим $\lambda_{0,8}=1,07$, т. е. найденное выше λ значительно превосходит $\lambda_{0,8}$.

Следовательно, критерий Колмогорова снова подтверждает наличие существенных расхождений между исследуемым и нормальным распределениями.

Перейдем теперь к критерию Пирсона. Для этого нужно вычислить значения $\frac{(n_i - np_i)^2}{np_i}$, которые и записаны в последней колонке таблицы 7.2. При этом первые три интервала объединены в один, так как $np_1 = 0,2$, $np_2 = 1,9$, т. е. даже в сумме эти числа меньше 5. По этой же причине объединены в один последние три интервала. Величина χ^2 равна сумме чисел последней колонки; число степеней свободы здесь $f = 6 - 3 = 3$. Поэтому по таблице IV Приложения $\chi_{0,95}^2 = 7,8$ (в критерии Пирсона уже можно брать обычный 5%-ный уровень значимости). Так как найденное по выборке $\chi^2 = 28,74 > \chi_{0,95}^2$, то и критерий Пирсона требует отклонить гипотезу нормальности.

7.2. Неравенство Чебышева и его использование. Допустим, что применив к результатам наблюдений описанные выше критерии, мы получили несогласие с нормальным распределением. В этом случае методы анализа, изложенные в § 6, применять нельзя. Однако статистическую обработку таких наблюдений можно проводить, если воспользоваться некоторыми другими общими теоремами теории вероятностей.

Точность и надежность результатов статистической обработки при распределениях, отличных от нормального, зависит от того, насколько известен характер изучаемого распределения. Но и при совершенно неизвестном распределении генеральной совокупности удастся получать неплохие оценки, по крайней мере для генерального среднего a .

В силу теоремы Гливленко (см. пп. 4.3, 4.4) в качестве оценки генерального среднего a всегда можно выбрать выборочное среднее \bar{x} . Точность такой оценки можно указать, зная генеральную дисперсию σ^2 , вместо которой (опять-таки в силу теоремы Гливленко) можно взять выборочную дисперсию s^2 . Доверительные интервалы для среднего при неизвестном распределении генеральной совокупности вычисляются на основании *неравенства Чебышева*, имеющего самый общий вероятностный характер.

Докажем неравенство Чебышева для непрерывной случайной величины ξ , имеющей произвольную плотность распределения $f(x)$. Согласно определению математического ожидания и дисперсии (п. 2.3)

$$M\xi = \int_{-\infty}^{\infty} xf(x) dx, \quad D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 f(x) dx.$$

Для любого положительного числа ε

$$D\xi = \int_{-\infty}^{M\xi - \varepsilon} (x - M\xi)^2 f(x) dx + \int_{M\xi - \varepsilon}^{M\xi + \varepsilon} (x - M\xi)^2 f(x) dx + \\ + \int_{M\xi + \varepsilon}^{\infty} (x - M\xi)^2 f(x) dx.$$

Заметим, что

$$\int_{M\xi - \varepsilon}^{M\xi + \varepsilon} (x - M\xi)^2 f(x) dx > 0,$$

как интеграл от положительной функции. Поэтому

$$D\xi > \int_{-\infty}^{M\xi - \varepsilon} (x - M\xi)^2 f(x) dx + \int_{M\xi + \varepsilon}^{\infty} (x - M\xi)^2 f(x) dx.$$

Благодаря выбранным пределам интегрирования в обоих подынтегральных выражениях выполняется неравенство $(x - M\xi)^2 \geq \varepsilon^2$, откуда

$$D\xi > \varepsilon^2 \left[\int_{-\infty}^{M\xi - \varepsilon} f(x) dx + \int_{M\xi + \varepsilon}^{\infty} f(x) dx \right].$$

Но

$$\int_{-\infty}^{M\xi - \varepsilon} f(x) dx = P\{\xi \leq M\xi - \varepsilon\}, \quad \int_{M\xi + \varepsilon}^{\infty} f(x) dx = P\{\xi \geq M\xi + \varepsilon\},$$

а неравенства $\xi \leq M\xi - \varepsilon$ и $\xi \geq M\xi + \varepsilon$ являются частными случаями неравенства $|\xi - M\xi| \geq \varepsilon$. В п. 1.3 было доказано, что вероятность события равна сумме вероятностей его

частных случаев. Значит,

$$\int_{-\infty}^{M\xi-\varepsilon} f(x)dx + \int_{M\xi+\varepsilon}^{\infty} f(x)dx = P\{|\xi - M\xi| \geq \varepsilon\},$$

откуда и следует неравенство Чебышева:

$$P\{|\xi - M\xi| \geq \varepsilon\} < \frac{D\xi}{\varepsilon^2}.$$

Сохраняя обозначение a для генерального среднего и σ^2 для генеральной дисперсии, запишем неравенство Чебышева в виде $P\{|x - a| \geq \varepsilon\} < \sigma^2/\varepsilon^2$, где x — случайное наблюдение над величиной ξ . Полагая $\varepsilon = k\sigma$, получим, что

$$P\{|x - a| \geq k\sigma\} < \frac{1}{k^2}.$$

Мы видим, что знание генеральной дисперсии позволяет и в случае произвольного распределения находить доверительные интервалы или проверять статистические гипотезы. При доверительной вероятности $1 - p$ неравенство Чебышева дает для генерального среднего a доверительную оценку

$$x - \frac{\sigma}{\sqrt{p}} \leq a \leq x + \frac{\sigma}{\sqrt{p}},$$

которая позволяет оценивать генеральное среднее по одному наблюдению. Если вместо x взять среднее \bar{x} выборки из n элементов, то $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, откуда

$$\bar{x} - \frac{\sigma}{\sqrt{np}} \leq a \leq \bar{x} + \frac{\sigma}{\sqrt{np}}.$$

Полученные оценки аналогичны тем, которые были найдены в п. 6.1 для нормального распределения, только вместо квантилей $u_{1-p/2}$ здесь множители $\frac{1}{\sqrt{p}}$. Эта замена сильно ухудшает оценки, чего и следовало ожидать при всеобщности этих оценок. Приведем, например, общую доверительную оценку, соответствующую вероятности $1 - p = 0,95$:

$$x - 4,24\sigma \leq a \leq x + 4,24\sigma$$

(напомним, что в случае нормального распределения вместо 4,24 стояло 1,96).

В практических расчетах округляют множитель 4,24 до 5 (что соответствует доверительной вероятности 0,96). Считая отклонения с вероятностью $p < 0,04$ практически невозможными (значимыми), мы получаем следующее правило: *каково бы ни было распределение генеральной совокупности с дисперсией σ^2 , отклонение от генерального среднего больше, чем на 5σ практически невозможно*. Именно этим правилом и вызван коэффициент 5 в предыдущем пункте, когда мы оценивали выборочный эксцесс.

Если плотность $f(x)$ изучаемого распределения симметрично убывает по обе стороны от генерального среднего (так называемое *симметричное одновершинное распределение*), то неравенство Чебышева справедливо в усиленной форме:

$$P \{ |x - a| \geq k\sigma \} < \left(\frac{2}{3k} \right)^2.$$

При доверительной вероятности 0,95 отсюда получается оценка

$$\bar{x} - 2,82\sigma \leq a \leq \bar{x} + 2,82\sigma,$$

а отклонения от генерального среднего, превышающие 3σ , считаются практически невозможными. Этим, в частности, объясняется появление коэффициента 3 при оценке выборочной асимметрии (п. 7.1), которая как раз имеет симметричное одновершинное распределение.

7.3. Подбор плотности теоретического распределения. Мы уже упоминали, что оценки, устанавливаемые на основании неравенства Чебышева, намного уступают аналогичным оценкам, полученным для нормального распределения. Это не означает, однако, что нормальное распределение является каким-то исключительно хорошим распределением — здесь просто сказывается эффект большей информации. Действительно, при обработке нормального распределения мы полностью знаем плотность распределения изучаемой случайной величины. При использовании же неравенства Чебышева о плотности распределения вообще нет никаких сведений.

Если удастся получить какую-либо информацию о плотности изучаемого распределения или, тем более, известен

конкретный вид этой плотности, то появляются оценки, не уступающие нормальному распределению, а порой и превосходящие его. При этом можно использовать общие результаты п. 2.3, в силу которых доверительной оценке генерального среднего $A \leq a \leq B$ соответствует доверительная вероятность

$$1 - p = \int_A^B f(x) dx$$

(через $f(x)$ обозначена плотность изучаемого распределения).

Особенно просто получить оценку генерального среднего, зная квантили ξ_p наблюдаемой случайной величины ξ . Как было показано в п. 5.2, доверительной вероятности $1 - p$ всегда соответствует неравенство

$$\xi_{p/2} \leq \xi \leq \xi_{1-p/2}.$$

Для некоторых часто встречающихся на практике распределений составлены подробные таблицы квантилей; вспомним, например, t -, χ^2 - и F -распределения; для других распределений построены специальные графики и т. п. Чтобы подобрать теоретическое распределение, соответствующее тому, которое мы наблюдаем, можно воспользоваться критериями согласия Пирсона или Колмогорова, так как они справедливы (в слегка измененном виде) для любого теоретического распределения.

Разумеется, никакие самые полные таблицы не могут охватить всех встречающихся на практике распределений. Возможна и другая ситуация: общие рассуждения подсказывают нам тип распределения, но остаются неизвестными его параметры. Во всех этих случаях окончательный вид распределения (его плотность) приходится устанавливать непосредственно по выборке, используя некоторые специальные приемы.

Легче всего изучать распределения, которые связаны с нормальным распределением функциональной связью, т. е. такие распределения, которые становятся нормальными, если вместо наблюдений x рассматривать величину $z = \psi(x)$. Пересчитав все данные с помощью функции $\psi(x)$, мы можем исследовать величину z по всем правилам § 6, а затем окончательные результаты пересчитать обратно для величины x (см. об этом также в п. 3.3).

Преобразующую функцию $\psi(x)$ нетрудно найти в следующем, часто встречающемся на практике случае. Пусть отклонение экспериментального наблюдения от нормального вызвано тем, что в процессе наблюдений изменяется генеральная дисперсия σ^2 . Это бывает в тех случаях, когда в процессе наблюдений меняется методика испытаний или точность используемых приборов. Например, при длительной стрельбе из артиллерийского орудия начинает сильно накаляться ствол, что немедленно повышает рассеяние снарядов. Производя наблюдения короткими порциями (выборками), мы часто можем проследить изменение выборочного стандарта s_x и даже установить его зависимость от наблюдений: $s_x = \varphi(x)$. Конкретный вид функции $\varphi(x)$ можно найти, например, по методу наименьших квадратов, рассматриваемому ниже, в п. 9.3.

Произведем замену с неизвестной пока функцией: $z = \psi(x)$. Согласно формуле (4.2), тогда получим

$$s_z = \psi'(x) s_x = \psi'(x) \varphi(x).$$

Будем теперь подбирать преобразующую функцию $\psi(x)$ так, чтобы дисперсия величины z стала постоянной. Мы получим для $\psi(x)$ уравнение

$$\psi'(x) \varphi(x) = C \quad (C > 0),$$

откуда

$$\psi(x) = C \int \frac{dx}{\varphi(x)}.$$

Найденная преобразующая функция позволяет перейти к «косвенным измерениям» с постоянной дисперсией, т. е. рассматривать все наблюдения как выборку из единой генеральной совокупности. Такое преобразование называется *стабилизацией дисперсии*.

Рассмотрим следующий пример. При производстве аммиачной селитры применяется выпаривание для повышения концентрации раствора до нужной величины. В течение этого процесса производят анализ концентрации единым методом, повторяя анализ каждой пробы четыре раза. Было обнаружено, что дисперсия анализа (ошибка воспроизводимости) растет с увеличением концентрации:

концентрация x (в %): 20 30 40 50 60 70

стандарт анализа s_x : 0,58 0,82 1,10 1,34 1,57 1,85

Зависимость между стандартом и концентрацией выражается примерной формулой

$$s_x = 0,025x + 0,085,$$

поэтому данные анализов нужно преобразовать по формуле

$$z = C \int \frac{dx}{0,025x + 0,085} = \frac{C}{0,025} \ln(0,025x + 0,085)$$

или, положив $C = 0,025 \lg e$,

$$z = \lg(0,025x + 0,085).$$

Пересчитав данные анализов по найденной формуле, мы стабилизируем дисперсию, т. е. сможем считать, что после преобразования ошибки метода, применяемого при анализах концентрации селитры, подчиняются единому нормальному закону распределения.

Хорошо изученное и подробно затабулированное нормальное распределение привлекает исследователей и во многих других случаях. Нередко рассматриваются различные видоизменения и модификации этого распределения. Например, изучая распределение с большой асимметрией A и средним a , рассматривают не саму случайную ошибку $\xi - a$, а ее модуль $|\xi - a|$, после чего распределение сравнивают с распределением модуля нормальной величины, имеющим плотность

$$f(x) = \frac{2}{\sigma \sqrt{2\pi}} e^{-x^2/2\sigma^2} \quad (x \geq 0).$$

В подавляющем большинстве случаев для эмпирического распределения можно получить хорошее аналитическое приближение, если использовать плотность стандартного нормального распределения $\varphi(x)$ и некоторые ее производные. Плотность $f(x)$ и функция $F(x)$ искомого распределения вычисляются по формулам

$$f(x) \approx \frac{1}{\sigma} \left[\varphi(u) - \frac{A}{6} \varphi'''(u) + \frac{E}{24} \varphi^{IV}(u) \right],$$

$$F(x) \approx \frac{1}{2} + \Phi(u) - \frac{A}{6} \varphi''(u) + \frac{E}{24} \varphi'''(u),$$

где A и E — выборочные асимметрия и эксцесс изучаемого

распределения, $\Phi(u)$ — функция Лапласа, а

$$u = \frac{x-a}{\sigma} . \}$$

Полученное распределение носит название *распределения Шарлье*. Для удобства пользования этим распределением в таблице XI Приложения протабулированы функция $\varphi(u)$ и ее производные до четвертого порядка включительно.

Найдем, например, пользуясь распределением Шарлье, вероятность того, что отклонения случайной величины ξ от генерального среднего не превзойдут 2σ . Здесь $u=2$; по таблице XI Приложения находим $\varphi''(2)=0,1620$, $\varphi'''(2)=-0,1080$. В силу четности функции $\varphi(u)$ вторая производная $\varphi''(u)$ также должна быть четной, а третья $\varphi'''(u)$ должна быть нечетной:

$$\varphi''(-u) = \varphi''(u), \quad \varphi'''(-u) = -\varphi'''(u).$$

Поэтому $\varphi''(-2)=0,1620$, $\varphi'''(-2)=0,1080$. Кроме того, по таблице I Приложения находим $\Phi(2)=0,4772$. Окончательно получаем

$$\begin{aligned} P\{\Delta\xi < 2\sigma\} &= \left[\Phi(2) - \frac{A}{6} \varphi''(2) + \frac{E}{24} \varphi'''(2) \right] - \\ &\quad - \left[\Phi(-2) - \frac{A}{6} \varphi''(-2) + \frac{E}{24} \varphi'''(-2) \right] = \\ &= 2 \cdot 0,4772 - 2 \cdot \frac{E}{24} \cdot 0,1080 = 0,9544 - 0,009E. \end{aligned}$$

Распределение Шарлье является приближенным, в силу чего оно слабо отражает специфику изучаемого распределения. В некоторых задачах приходится сталкиваться с распределениями, настолько отличными от нормального, что положение не спасти никакими модификациями. Многие из таких распределений в настоящее время неплохо изучены, известны их определяющие свойства и признаки (см. п. 3.1). По этим признакам, однако, удастся устанавливать лишь тип распределения; параметры же распределения приходится оценивать по самой выборке (напомним, что такая же ситуация возникает при обработке нормально распределенной случайной величины, где параметры a и σ нужно оценивать по выборке). Так, широко используемое

в теории массового обслуживания *экспоненциальное распределение* с плотностью

$$f(x) = \lambda e^{-\lambda x}$$

содержит параметр λ ; используемое в теории надежности *гамма-распределение* с плотностью

$$f(x) = \frac{x^{m-1}}{c^m (m-1)!} e^{-x/c} \quad (x \geq 0)$$

содержит два параметра m и c .

Для того чтобы пользоваться подобными распределениями, необходимо знать формулы, которые позволяли бы по элементам выборки как можно точнее оценивать эти генеральные параметры. Такие формулы можно выводить с помощью общего приема, называемого *принципом максимума правдоподобия* и состоящего в следующем.

Предположим вначале, что плотность $f(x)$ теоретического распределения нам полностью известна, и пусть на опыте была получена выборка x_1, x_2, \dots, x_n . Окружим каждую точку x_i окрестностью длины ε . Мы получим некоторый набор интервалов. Вероятность попасть в первый интервал с границами $x_1 - \frac{\varepsilon}{2}, x_1 + \frac{\varepsilon}{2}$ приближенно равна $f(x_1)\varepsilon$. Точно так же вероятность попасть во второй интервал равна $f(x_2)\varepsilon$, в третий — $f(x_3)\varepsilon$ и т. д. Если произведены n наблюдений, то вероятность того, что одновременно первое наблюдение попадает в первый интервал, второе — во второй и т. д., есть вероятность совместного осуществления событий и в силу п. 1.3 равна произведению вероятностей:

$$p = f(x_1) f(x_2) \dots f(x_n) \varepsilon^n.$$

Если теперь известен только общий вид плотности $f(x)$, а параметры распределения неизвестны, то эти параметры нужно выбирать такими, чтобы вероятность p была максимальной. Действительно, событие с вероятностью p осуществилось на самом деле — ведь у нас есть выборка x_1, x_2, \dots, x_n , где каждый элемент попадает в нужный интервал. Естественно ожидать, что событию, осуществившемуся при первом же испытании, соответствует максимальная вероятность.

Вместо вероятности p рассматривается обычно функция

$$L = \ln \frac{p}{e^n} = \sum_{i=1}^n \ln f(x_i),$$

которая называется *функцией правдоподобия*. Используется при этом тот факт, что вероятность p и функция L имеют максимумы при одних и тех же значениях разыскиваемых параметров.

Максимум функции L можно искать по обычным правилам математического анализа. Если разыскиваемые параметры обозначить через $\alpha_1, \alpha_2, \dots, \alpha_k$, то для их отыскания применяются уравнения

$$\begin{cases} \frac{\partial L}{\partial \alpha_1} = 0, \\ \frac{\partial L}{\partial \alpha_2} = 0, \\ \dots \dots \dots \\ \frac{\partial L}{\partial \alpha_k} = 0. \end{cases}$$

Решения этой системы выражаются через x_1, x_2, \dots, x_n ; они и дадут нам необходимые оценки.

Рассмотрим, например, упоминавшееся выше экспоненциальное распределение. Для него функция правдоподобия имеет вид

$$L = \sum_{i=1}^n \ln (\lambda e^{-\lambda x_i}) = n \ln \lambda - \sum_{i=1}^n \lambda x_i.$$

Дифференцируя по λ и приравнявая к нулю, получим уравнение

$$\frac{n}{\lambda} - \sum_{i=1}^n x_i = 0,$$

откуда

$$\lambda = \frac{n}{\sum_{i=1}^n x_i}.$$

Получаемые с помощью принципа максимума правдоподобия соотношения имеют, конечно, некоторые ошибки.

Чтобы эти ошибки оценить (найти доверительные интервалы и т. п.), нужно знать распределения исследуемых параметров, что удается далеко не всегда. Более грубо эти ошибки можно оценить по неравенству Чебышева, найдя вначале дисперсии оцениваемых параметров. Для нахождения дисперсий используются те же формулы, что и для вычисления параметров; вычисляются эти дисперсии как дисперсии «косвенных измерений» (см. п. 4.4). Например, для найденного выше λ

$$s_{\lambda}^2 = \sum_{i=1}^n \left[\frac{\partial}{\partial x_i} \left(\frac{n}{\sum_{i=1}^n x_i} \right) \right]^2 s_x^2 = \frac{n^3 s_x^2}{\left(\sum_{i=1}^n x_i \right)^4},$$

где s_x^2 — дисперсия наблюдений x_1, x_2, \dots, x_n *).

7.4. Оценка вероятности случайного события. Общие методы статистических оценок можно использовать для исследования вероятностей случайных событий. С необходимостью такого исследования сталкиваются чаще всего при решении следующих задач.

1. Исследуется цифровой материал, связанный с событием, природа которого слабо изучена или вообще неизвестна. В этом случае вероятность события является наиболее доступной для изучения закономерностью, позволяющей делать объективные выводы о рассматриваемом событии. В качестве примера приведем такое событие, как формирование пола у будущего ребенка. Природа этого события еще не известна, и единственная возможность его изучения состоит пока в обработке статистического материала, определении различных вероятностей (в п. 1.2 приводилась, в частности, вероятность рождения мальчика, равная примерно 0,52).

*) Нетрудно видеть, что $\lambda = \frac{1}{\bar{x}}$, где \bar{x} — среднее выборки x_1, x_2, \dots, x_n . Поэтому от параметра λ в данном случае лучше перейти к параметру \bar{x} , дисперсия которого определяется очень легко и равна $\frac{s_x^2}{n}$ (см. п. 4.4).

2. Изучается событие, вероятность которого можно предсказать заранее на основании каких-либо абстрактных соображений (например, из соображений симметрии). В этом случае опытное вычисление вероятности позволило бы проверить исходные абстрактные предпосылки. Так, наблюдая частоту выпадения герба при бросании монеты, можно проверить гипотезу о ее симметричности. Другой пример дает теоретическая физика, где исходные гипотезы позволяют предсказывать вероятность той или иной формы распада элементарных частиц. Совпадение предсказанной вероятности с реально наблюдаемой явилось бы серьезным доводом в пользу исходной гипотезы.

3. Последняя задача связана с проверкой неизменности вероятности события в процессе испытаний. Можно вообще ставить задачу о сравнении двух или нескольких вероятностей различных событий.

При наблюдениях, предназначенных для вычисления вероятности, отмечается только качественный признак — появление или непоявление исследуемого события. Количественные результаты появляются при подсчете числа появлений события. Это число является случайной величиной ξ_n (n — число всех испытаний), имеющей биномиальное распределение (см. пп. 1.5, 2.2, 3.3). Напомним, что при биномиальном распределении случайная величина принимает лишь целые значения $k=0, 1, 2, \dots, n$, причем

$$P\{\xi_n = k\} = C_n^k p^k (1-p)^{n-k},$$

где p — вероятность появления события A в единичном испытании.

Как было доказано в п. 3.3 (теорема Бернулли), вероятность есть предел, к которому почти всегда (с вероятностью 1) стремится частота события A при $n \rightarrow \infty$. Следовательно, частота $\omega = \frac{k}{n}$ может быть при достаточно большом n принята в качестве оценки вероятности p .

Нетрудно видеть, что частота сама является случайной величиной и, следовательно, равенство $\omega = p$ содержит случайную ошибку. В связи с этим желательно уметь находить доверительные интервалы для неизвестной вероятности p . Соответствующим расчетам серьезно помогает то обстоятельство, что при больших n и при p , не очень близком

к 0 и 1, биномиальное распределение мало отличается от нормального (п. 3.3) с теми же математическим ожиданием $a=np$ и дисперсией $\sigma^2=np(1-p)$.

Из линейности нормального распределения вытекает, что распределение частоты ω также будет «почти» нормальным с параметрами $a_\omega = \frac{a}{n} = p$ и $\sigma_\omega = \frac{\sigma}{n} = \sqrt{\frac{p(1-p)}{n}}$. Поэтому при доверительной вероятности $1-p_0^*)$ справедлива оценка

$$\omega - u_{1-p_0/2} \sqrt{\frac{p(1-p)}{n}} \leq p \leq \omega + u_{1-p_0/2} \sqrt{\frac{p(1-p)}{n}},$$

где значение частоты ω берется из наблюдений, а $u_{1-p_0/2}$ — квантиль стандартного нормального распределения. В левую и правую части полученной оценки входит неизвестное p , и решение неравенств относительно p очень трудоемко. Поэтому величину p в формуле стандарта σ_ω приближенно заменяют опять-таки на найденное из наблюдений значение ω . Мы получим оценку

$$\omega - u_{1-p_0/2} \sqrt{\frac{\omega(1-\omega)}{n}} \leq p \leq \omega + u_{1-p_0/2} \sqrt{\frac{\omega(1-\omega)}{n}}.$$

Рассмотрим следующий пример. В научно-исследовательском институте психологии, изучавшем возможность телепатии, т. е. передачи мыслей на расстояние без посредства органов чувств, был проделан опыт. В одной комнате сидел экспериментатор, который по сигналу брал один из лежавших перед ним трех картонных прямоугольников: черный, белый и полосатый. Сидящий в соседней комнате второй экспериментатор по тому же сигналу выбирал один из таких же прямоугольников. Легко видеть, что вероятность случайного совпадения прямоугольников равна $\frac{1}{3}$, следовательно, такой же примерно должна быть и частота совпадений на опыте. Если же между экспериментаторами существует телепатическая связь, то вероятность совпадений может неслучайным образом повыситься.

*) Мы обозначаем здесь уровень значимости через p_0 , чтобы не путать его с разыскиваемой вероятностью p .

Произведя 100 наблюдений, экспериментаторы получили 39 совпадений, т. е. частоту $\omega=0,39$. При доверительной вероятности 0,95 это даст для неизвестной вероятности p оценку

$$0,39 - 1,96 \sqrt{\frac{0,39 \cdot 0,61}{100}} \leq p \leq 0,39 + 1,96 \sqrt{\frac{0,39 \cdot 0,61}{100}}$$

или, после вычислений, $0,294 \leq p \leq 0,486$.

Мы видим, что $\frac{1}{3}$ попадает в полученный интервал, значит, отклонение частоты от $\frac{1}{3}$ вполне могло быть случайным (на уровне значимости $p_0=0,05$).

Условия задачи позволяют применить односторонний критерий, однако это не меняет вывода. Только взяв уровень значимости $p_0=0,15$, мы получим односторонний интервал, не содержащий $\frac{1}{3}$. Таким образом, наличие телепатии в указанном эксперименте приходится считать весьма сомнительным.

Изучавшаяся в § 6 нормально распределенная случайная величина имела среднее a и стандарт σ , не зависящие друг от друга. Наоборот, среднее a_ω и стандарт σ_ω зависят друг от друга и связаны формулой

$$\sigma_\omega = \sqrt{\frac{a_\omega(1-a_\omega)}{n}};$$

этим и вызвана трудность получения точной доверительной оценки вероятности p . Формулы предыдущего пункта позволяют перейти от ω к такой случайной величине $\tilde{\omega}$, дисперсия которой постоянна и не зависит от $a_\omega=p$. Преобразующая функция здесь равна

$$\psi(x) = \int \frac{dx}{\sqrt{\frac{x(1-x)}{n}}} = 2\sqrt{n} \arcsin \sqrt{x}.$$

Это значит, что случайная величина

$$\tilde{\omega} = 2\sqrt{n} \arcsin \sqrt{\omega} \quad (7.2)$$

имеет дисперсию, равную 1. Распределение величины $\tilde{\omega}$ по-прежнему будет нормальным, поэтому ее легко применять для проверки различных гипотез о вероятности p .

Используем величину $\tilde{\omega}$ для сравнения меткости стрельбы двух стрелков. Оба стрелка сделали по 500 выстрелов; первый поразил цель 385 раз, второй — 360. Здесь $n=500$, $\omega_1 = \frac{385}{500} = 0,77$, $\omega_2 = \frac{360}{500} = 0,72$. Если принять гипотезу об одинаковой меткости стрелков, то вероятности попадания в цель p_1 и p_2 должны быть равны и, значит, разность $\tilde{\omega}_1 - \tilde{\omega}_2$ должна иметь нормальное распределение с параметрами $a=0$, $\sigma = \sqrt{2}$. Поэтому при уровне значимости $p_0=0,05$ и одностороннем критерии критическим значением нулевой гипотезы будет число $u_{0,95}\sigma = 1,64\sqrt{2} = 2,33$. В то же время реальный подсчет разности $\tilde{\omega}_1 - \tilde{\omega}_2$ в нашем случае дает

$$\tilde{\omega}_1 - \tilde{\omega}_2 = 2\sqrt{500} (\arcsin \sqrt{0,77} - \arcsin \sqrt{0,72}) = 2,46.$$

Мы получили, что $\tilde{\omega}_1 - \tilde{\omega}_2$ больше критического значения 2,33. Следовательно, различие вероятностей p_1 и p_2 нужно считать неслучайным (значимым) — первый стрелок был более метким, чем второй.

Для сравнения нескольких вероятностей также можно использовать величину $\tilde{\omega}$. Пусть проведены k серий независимых испытаний. Число испытаний в i -й серии обозначим через n_i , найденную из опыта частоту события A — через ω_i . Перейдем к величинам $\tilde{\omega}_i = 2\sqrt{n_i} \arcsin \sqrt{\omega_i}$ и найдем среднее $\bar{\omega} = \frac{\tilde{\omega}_1 + \dots + \tilde{\omega}_k}{k}$. Тогда сумма

$$\chi^2 = \sum_{i=1}^k (\tilde{\omega}_i - \bar{\omega})^2 = \sum_{i=1}^k \tilde{\omega}_i^2 - \frac{\left(\sum_{i=1}^k \tilde{\omega}_i\right)^2}{k}$$

имеет приближенно χ^2 -распределение с $f=k-1$ степенями свободы *).

Для проверки гипотезы о равенстве всех вероятностей $p_1=p_2=\dots=p_k$ можно теперь использовать квантили χ^2 -распределения, приведенные в таблице IV Приложения. Если

*) Написанная сумма совпадает с обычным определением величины χ^2 (см. п. 6.2), так как здесь все дисперсии $\sigma_{\omega_i}^2 = 1$.

при выбранном уровне значимости p_0 получится, что $\chi^2 > \chi^2_{1-p_0}$, то гипотезу о равенстве вероятностей нужно отклонить. Если же $\chi^2 \leq \chi^2_{1-p_0}$, то для отклонения гипотезы оснований нет.

Описанный критерий применяют чаще всего для проверки того, что вероятность некоторого события A не изменилась в процессе длительных исследований. Рассмотрим пример из метеорологических наблюдений. За 65 лет нашего столетия в городе N отмечалось 26 лет с майскими снегопадами. За весь XIX век таких лет было 58, за весь XVIII век — 62. Одновременно по летописям установлено, что в период 1390—1442 гг. (т. е. за 52 года) было 44 года со снегопадами в мае. Мы видим, что частота события «год с майским снегопадом» уменьшается со временем. Проверим на уровне значимости $p_0 = 0,05$, не является ли это уменьшение случайным.

По данным наблюдений имеем $n_1 = 65, \omega_1 = \frac{2}{5}$; $n_2 = 100, \omega_2 = \frac{29}{50}$; $n_3 = 100, \omega_3 = \frac{31}{50}$; $n_4 = 52, \omega_4 = \frac{11}{13}$. Используя тригонометрические таблицы, найдем $\tilde{\omega}_1 = 11,04$; $\tilde{\omega}_2 = 17,31$; $\tilde{\omega}_3 = 18,13$; $\tilde{\omega}_4 = 16,87$. Отсюда $\chi^2 = 31,5$. По таблице IV Приложения находим, что при $f = 3$ степенях свободы $\chi^2_{0,95} = 7,8$, т. е. найденное нами $\chi^2 > \chi^2_{0,95}$, тем самым гипотеза о неизменной вероятности майских снегопадов отвергается.

Если вероятность p очень мала, а n велико, то дисперсия $\sigma^2 \approx np = a$, т. е. дисперсия и среднее практически совпадают. Если при этом $np < 9$, то (п. 3.3) замена биномиального распределения нормальным приводит к большим погрешностям. Как отмечалось в п. 3.3, в этом случае биномиальное распределение можно приближенно заменять распределением Пуассона:

$$P \{ \xi_n = k \} = \frac{a^k e^{-a}}{k!}.$$

Распределение Пуассона является дискретным, поэтому связанные с ним расчеты довольно трудоемки. Тем не менее, они проще, чем непосредственное использование формул

биномиального распределения, уже хотя бы потому, что в распределении Пуассона участвует только один параметр a .

Для оценки параметра a берется само число появлений события $k=n\omega$, при этом число всех испытаний n на оценку не влияет. Распределение k как случайной величины в случае распределения Пуассона оказывается тесно связанным с χ^2 -распределением. Это и позволяет получать достоверные интервалы для параметра a . Приведем без доказательства соответствующие формулы.

При достоверной вероятности $1-p_0$ справедлива оценка

$$\frac{1}{2} \chi_{p_0/2}^2(f_1) \leq a \leq \frac{1}{2} \chi_{1-p_0/2}^2(f_2),$$

где степени свободы $f_1=2k$ и $f_2=2(k+1)$. Значение квантилей величины χ^2 нужно брать, как обычно, из таблицы IV Приложения. Если число степеней свободы окажется равным нулю, то и соответствующая величина χ^2 тождественно равна нулю.

Подчеркнем еще раз, что в оценке параметра a число испытаний n не участвует — оно лишь подразумевается очень большим. Поэтому при нахождении закона распределения это число можно вообще не знать. Если же цель исследования — определение вероятности p , то n потребуется, ибо $p = \frac{a}{n}$. Для вероятности p получится достоверная оценка

$$\frac{1}{2n} \chi_{p_0/2}^2(f_1) \leq p \leq \frac{1}{2n} \chi_{1-p_0/2}^2(f_2).$$

Рассмотрим такой пример. При исследовании космических лучей следы тяжелых мезонов были отмечены на трех снимках из 1000. Необходимо вычислить вероятность появления указанных частиц. Здесь $k=3$, поэтому $f_1=6$, $f_2=8$. По таблице IV Приложения находим $\chi_{0,05}^2(6)=1,63$, $\chi_{0,95}^2(8)=15,5$. Следовательно, при достоверной вероятности 0,90 справедлива оценка $\frac{1,63}{2000} \leq p \leq \frac{15,5}{2000}$ или, после окончательных вычислений, $0,0008 \leq p \leq 0,0078$.

Распределение Пуассона применимо к изучению редких событий с числом появлений $k < 9$ (при достаточно большом n). Если $k \geq 9$, то распределение Пуассона становится близ-

ким к нормальному. В связи с этим даже очень редкие события можно изучать, не прибегая к распределению Пуассона, если удалось зарегистрировать больше 9—10 появлений события.

Различные гипотезы о вероятности p в случае распределения Пуассона проверяются точно так же, как и при нормальном распределении, если вместо (7.2) перейти к величине $\hat{\omega} = 2\sqrt{k}$, дисперсия которой, как нетрудно проверить, также равна 1.

7.5. Использование оценок вероятности для анализа распределения. Пусть имеется выборка x_1, x_2, \dots, x_n из генеральной совокупности с неизвестным распределением. Будем рассматривать в качестве случайного события A наличие некоторого качества у элементов x_i . Высказывая ту или иную гипотезу о генеральном распределении, мы на основании этой гипотезы можем обычно предсказывать теоретически вероятность p события A . Одновременно с этим можно, проверив наличие качества A у элементов заданной выборки, найти эмпирическую оценку этой вероятности (методами предыдущего пункта). Сравнивая эмпирическое и теоретическое значения вероятности, мы получим подтверждение или опровержение высказанной ранее гипотезы.

Рассмотрим такой пример. В результате наблюдений была получена выборка из 16 элементов:

12 16 10 11 15 13 10 14 9 12 15 14 8 9 12 11.

Высказывается гипотеза о том, что наблюдаемое распределение симметрично. Подтверждение этой гипотезы было бы очень полезным, так как позволило бы применять усиленное неравенство Чебышева (см. п. 7.2). Для проверки найдем вначале среднее выборки $\bar{x} = 11,85$. Если справедлива гипотеза о симметричности, то вероятность значений $x < \bar{x}$ равна $\frac{1}{2}$.

Заданную выборку можно рассматривать как последовательность испытаний по проверке качества A , состоящего в том, что элемент выборки меньше, чем \bar{x} . Число всех испытаний здесь $n = 16$, число появлений качества A находится непосредственно по выборке и равно $k = 6$. Отсюда частота

$\omega = \frac{3}{8}$, и при доверительной вероятности $1 - p_0 = 0,95$ получается оценка $\frac{3}{8} - 1,96 \sqrt{\frac{3}{8} \left(1 - \frac{3}{8}\right) \frac{1}{16}} \leq p \leq \frac{3}{8} + 1,96 \sqrt{\frac{3}{8} \left(1 - \frac{3}{8}\right) \frac{1}{16}}$ или, после вычислений, $0,13 \leq p \leq 0,62$.

Мы видим, что значение $p = \frac{1}{2}$ попадает в доверительный интервал. Следовательно, наблюдаемая выборка не противоречит гипотезе о симметричности генерального распределения.

Сопоставление эмпирической и теоретической вероятности может быть использовано и для сравнения двух генеральных распределений по выборкам. Такое сравнение проводится с учетом всех элементов заданных выборок и иногда (правда, очень редко) оказывается более чувствительным, чем рассматривавшееся в п. 6.4 сравнение средних. Другим достоинством этого способа является его применимость к любым видам распределений.

Сравним, например, две выборки из 15 элементов каждая:

x : 0,42 0,47 0,44 0,42 0,45 0,44 0,45 0,46 0,48
 0,41 0,45 0,47 0,44 0,42 0,41
 y : 0,40 0,42 0,43 0,44 0,39 0,41 0,44 0,43 0,45
 0,42 0,43 0,46 0,42 0,40 0,42.

В качестве события A будем рассматривать тот факт, что элемент x больше соответствующего элемента y . Если обеим выборкам соответствует единое генеральное среднее, то событие A должно иметь вероятность $\frac{1}{2}$. Сравнивая элементы заданных выборок (первый с первым, второй со вторым и т. д.), найдем, что частота события A равна $\omega = \frac{12}{15} = 0,8$. Поэтому при доверительной вероятности $1 - p_0 = 0,95$ справедлива оценка

$$0,8 - 1,96 \sqrt{\frac{0,8 \cdot 0,2}{15}} \leq p \leq 0,8 + 1,96 \sqrt{\frac{0,8 \cdot 0,2}{15}}$$

или, после вычислений, $0,6 \leq p \leq 1,0$.

Значение $p = \frac{1}{2}$ не попадает в этот доверительный интервал, следовательно, гипотезу о равенстве генеральных средних нужно отбросить.

Точно так же можно сравнивать дисперсии выборок. Для этого нужно вместо x_i выписать квадраты $(x_i - \bar{x})^2$ отклонений x_i от среднего \bar{x} и то же самое сделать для y_i . После этого квадраты $(x_i - \bar{x})^2$ и $(y_i - \bar{y})^2$ сравниваются между собой так же, как раньше x_i и y_i .

Применение описанного метода сравнения допустимо только в том случае, если обе выборки расположены совершенно случайно (например, в порядке естественного получения). Малейшая упорядоченность сразу же исказит результат.

Проведенное выше сравнение выборок учитывает только знак разности $x_i - y_i$, но не ее величину. Поэтому оно может быть полезно лишь там, где один из знаков подавляюще преобладает над другим, а это бывает редко. Но зато при наличии такого преобладания указанное сравнение выявляет любые различия, как бы малы они ни были, почему и может оказаться чувствительнее, чем сравнение средних по критерию Стьюдента.

Нередко при обработке наблюдений приходится пользоваться случайностью заданной последовательности элементов (например, в разбиравшемся выше сравнении выборок). Если эта последовательность возникла естественным путем как результат действия случайных факторов, то ее случайность редко вызывает сомнения. Однако, в процессе наблюдений могут появиться некоторые доминирующие факторы, вызывающие неслучайное смещение результатов, и это уже требует специальной проверки. Встречается и такая ситуация, когда последовательность создается искусственно, например, смешивается из нескольких различных последовательностей. В этом случае обязательно нужно выяснить, осталась ли она случайной.

В зависимости от конкретных обстоятельств применяются различные критерии случайности. Мы рассмотрим только случай, когда заданная последовательность содержит одинаковое количество элементов двух видов и проверяется гипотеза о том, что взаимное расположение элементов

различных видов обязано своим происхождением чисто случайным причинам. Разумеется, это лишь весьма частный случай поставленной выше задачи, но к нему сводится несколько важных проблем обработки наблюдений, часть из которых мы рассмотрим ниже.

Итак, пусть последовательность из n элементов содержит $\frac{n}{2}$ элементов типа A и столько же элементов типа B . Если вместо каждого элемента отмечать только его тип, то заданная последовательность даст некоторую комбинацию букв A и B , например,

$$AAABABBBABAABBB.$$

В предположении случайности все такие комбинации должны быть абсолютно равновозможны. Число этих комбинаций равно $C_n^{\frac{n}{2}} = \frac{n!}{\left(\frac{n}{2}!\right)^2}$, следовательно, вероятность каж-

дой комбинации равна $\frac{\left(\frac{n}{2}!\right)^2}{n!}$. Для того чтобы выяснить, могла ли та или иная комбинация возникнуть случайно, нужно, согласно общему правилу проверки статистических гипотез, связать с этими комбинациями некоторую случайную величину, которая позволяла бы выделять маловероятные комбинации.

В качестве такой случайной величины обычно рассматривают число серий R . *Серией* называется любая последовательность элементов одинакового типа, граничащая с элементами другого типа. В приведенной выше комбинации вначале идет серия AAA , затем серия из одного элемента B , затем A , BB и т. д. Всего серий в указанном примере $R=8$.

Число серий R при каждом фиксированном n является конечнозначной случайной величиной и точный расчет всех ее вероятностей довольно громоздок. Если же n достаточно велико (по крайней мере больше 10), то распределение величины R становится близким к нормальному с параметрами

$$a_R = \frac{n+1}{2}, \quad \sigma_R^2 = \frac{n-1}{4}.$$

Поэтому для R можно написать доверительную оценку

$$a_R - u_{1-p/2} \sigma_R \leq R \leq a_R + u_{1-p/2} \sigma_R,$$

где $u_{1-p/2}$ — квантиль стандартного нормального распределения, соответствующий уровню значимости p . После подстановки значений a_R и σ_R получается оценка

$$\frac{1}{2}(n+1-u_{1-p/2}\sqrt{n-1}) \leq R \leq \frac{1}{2}(n+1+u_{1-p/2}\sqrt{n-1}).$$

Поскольку R может быть только целым, полученные доверительные границы округляют до целых, уменьшая нижнюю и увеличивая верхнюю. Например, при $n=50$ и уровне значимости $p=0,05$ находим

$$\frac{1}{2}(50+1-1,96\sqrt{49}) \leq R \leq \frac{1}{2}(50+1+1,96\sqrt{49})$$

или, после вычислений и округления, $18 \leq R \leq 33$.

Число серий в заданной комбинации можно теперь использовать в качестве критерия ее случайности. Если это число лежит в заранее вычисленных доверительных пределах, то комбинацию можно считать случайной. Если же найденное по комбинации R выходит за доверительные пределы (попадает в критическую область гипотезы), то случайность комбинации нужно признать практически невозможной. В рассматривавшейся выше комбинации число всех элементов $n=14$, поэтому при уровне значимости $p=0,05$ число серий R должно лежать в пределах от 3 до 12. Поскольку фактическое число серий $R=8$ лежит в этих пределах, комбинацию нужно признать случайной.

Укажем некоторые применения метода серий. Допустим, сравниваются две выборки численностью $m = \frac{n}{2}$. Объединим обе выборки в одну (объема n) и расположим все элементы объединенной выборки в возрастающем порядке. Если прежние выборки различались лишь случайно, то в полученной последовательности чередование элементов, взятых из первой выборки (тип A) и из второй (тип B) должно быть чисто случайным. Последнее обстоятельство и проверяется с помощью серий.

Другое применение находит метод серий в тех случаях, когда нужно проверить, является ли наблюдаемое изменение

(флуктуации) результатов чисто случайным. Рассмотрим такой пример. При обработке валиков заданного диаметра на токарном станке получились следующие 20 значений диаметров (в мм):

22,6 22,8 22,8 22,7 23,0 22,8 23,0 23,0 23,1 23,2
23,4 23,2 23,3 23,4 23,4 23,1 22,9 23,6 23,4 23,6.

Требуется проверить, нет ли в диаметрах неслучайного изменения, связанного с каким-либо постоянным фактором (например, затуплением резца).

Найдем вначале выборочную медиану (см. п. 5.1). Для этого заданную выборку нужно расположить в возрастающем порядке и в получившейся последовательности взять средний (по номеру) элемент. Так как у нас число элементов четное ($n=20$), то средних элементов будет два — 10-й и 11-й, и в качестве медианы нужно брать их полусумму. В нашем примере оба эти элемента оказываются одинаковыми, в силу чего медианой выборки будет их общее значение 23,1.

Вернемся теперь к первоначальной выборке. Будем элемент обозначать буквой *A*, если он меньше медианы, и буквой *B* в противном случае (элементы, равные медиане, вообще исключим из рассмотрения). Мы получим комбинацию букв

ААААААААВВВВВВВВВВ.

В этой комбинации четыре серии. В то же время для $n=18$ элементов (столько их у нас сейчас осталось) доверительные пределы числа серий R при уровне значимости $p=0,05$ равны 5 и 14. Мы видим, что получившуюся комбинацию нельзя считать случайной. А это заставляет признать, что в диаметрах изготовленных валиков наблюдается неслучайное изменение.

Подведем некоторые общие итоги параграфа. Обработка наблюдений всегда должна начинаться с анализа наблюдаемого распределения (статистического или непосредственно по условиям испытаний с применением теоремы Ляпунова). Если распределение оказывается отличным от нормального, то нужно попытаться найти его плотность, определяя все нужные параметры по принципу максимума правдоподобия,

Вместе с тем, многие оценки можно получать, не прибегая к детальному изучению распределения — методами пп. 7.2 и 7.5. Подобные методы носят название *непараметрической статистики*. Хотя непараметрическая статистика и обладает высокой универсальностью, но применять ее нужно осторожно, так как достаточно надежные результаты получаются лишь при очень больших n .

И последнее. При обработке небольшого цифрового материала (микростатистика) можно, как правило, всегда пользоваться критериями нормального распределения, изложенными в § 6, так как отклонения различных распределений друг от друга практически не заметны на малых выборках.

§ 8. ДИСПЕРСИОННЫЙ АНАЛИЗ

8.1. Постановка задачи. В предыдущем изложении, как правило, предполагалось, что наблюдаемый разброс результатов связан лишь со случайными причинами (факторами). Именно это предположение лежало в основе всех проверявшихся нулевых гипотез. Все подконтрольные испытателю факторы поддерживались на одном и том же уровне. Обработка материала сводилась фактически к тому, чтобы определить точность применяемого метода исследования, его ошибку воспроизводимости. В качестве меры такой точности рассматривалась дисперсия.

Перейдем теперь к рассмотрению второй (по счету, но не по важности) задачи математической обработки наблюдений. Задача эта состоит в том, что один или несколько основных факторов начинают заданным образом изменяться. Эти изменения могут повлиять на результаты наблюдений. Степень такого влияния, его качественные и количественные характеристики и будут объектами наших рассмотрений в этом и следующем параграфах.

Настоящий параграф будет посвящен вопросу общей оценки действующего переменного фактора, его сравнения с другими факторами. Важность подобных общих оценок можно подчеркнуть хотя бы следующими двумя примерами.

Земной шар, помимо суточного вращения вокруг своей оси и годового вращения вокруг Солнца, подвержен еще некоторым слабым колебаниям и перемещениям. Эти перемещения удастся выявить, наблюдая за звездами. Однако перемещения так малы, что могут быть приняты просто за ошибки наблюдений, связанные с неточной работой астрономических инструментов (последнее, кстати, и явилось причиной того, что перемещения, о которых идет речь, были

обнаружены лишь сравнительно недавно). Прежде чем проводить детальное изучение перемещений, выяснять их причины и т. д., необходимо проверить в целом, действительно ли эти перемещения существенны на фоне случайных погрешностей приборов.

Второй пример возьмем из проблемы контроля за производством. От завода обычно требуется выпуск однородной по своим качествам продукции. Процесс производства включает в себе несколько стадий, каждая из которых вносит свою «лепту» в фактическую неоднородность конечного продукта. Одновременное усовершенствование всех стадий, как правило, слишком трудоемко и дорого. Поэтому вначале нужно выяснить, какая из стадий дает наибольшую неоднородность, и усовершенствование начинать с нее. При исследовании может даже оказаться, что некоторые стадии дают лишь незначительную неоднородность, так что затраты на их усовершенствование вообще были бы неоправданными.

В дальнейшем мы все время будем считать, что случайные ошибки наблюдений имеют нормальное распределение. Влияние изучаемых факторов может быть двояким: они могут изменять как истинный результат (среднее) наблюдений α , так и дисперсию этих наблюдений σ^2 . Мы, однако, все время будем предполагать, что дисперсия наблюдений σ^2 остается неизменной. Это предположение обычно оправдывается, если для наблюдений используется одна и та же методика, одни и те же приборы. Если же стабильность дисперсии вызывает сомнение, нужно провести специальное исследование с помощью критериев Бартлета или Кохрана. В случае значимого изменения дисперсии в процессе наблюдений нужно попытаться ее стабилизировать, подобрав соответствующую преобразующую функцию (п. 7.3).

Таким образом, в настоящем параграфе будет изучаться лишь влияние переменных факторов на генеральное среднее наблюдаемого распределения. Для того чтобы сравнивать влияние различных факторов, нужно найти какой-нибудь достаточно надежный и универсальный показатель этого влияния. Допустим, изучаемый фактор A на различных уровнях привел к серии истинных результатов:

$$a_1, a_2, \dots, a_k.$$

В качестве показателя влияния фактора A мы будем тогда брать число

$$\sigma_A^2 = \frac{1}{k} \sum_{i=1}^k (a_i - \bar{a})^2,$$

где \bar{a} — среднее арифметическое чисел a_i :

$$\bar{a} = \frac{a_1 + a_2 + \dots + a_k}{k}.$$

Число σ_A^2 называется *дисперсией фактора A* . Это название дано по аналогии с обычной дисперсией, но нужно помнить, что числа a_1, a_2, \dots, a_k не являются случайными, поэтому σ_A^2 не связана ни с какой случайной величиной.

Выбор σ_A^2 удобен по двум причинам. Во-первых, дисперсия является простейшей мерой рассеивания (п. 2.3). Во-вторых, показатель влияния фактора A определен теперь аналогично показателю влияния случайного фактора (т. е. обычной дисперсии σ^2), что позволит непосредственно сравнивать фактор A с эффектом случайности. Изучение переменных факторов по их дисперсиям называется *дисперсионным анализом* *).

Рассмотрим самый простой случай, когда дисперсия наблюдений σ^2 известна заранее и исследуется один переменный фактор A . Пусть при изменении фактора A получились результаты наблюдений x_1, x_2, \dots, x_k . Найдем выборочную дисперсию

$$s^2 = \frac{1}{k-1} \sum (x_i - \bar{x})^2 = \frac{1}{k-1} \left[\sum x_i^2 - \frac{(\sum x_i)^2}{k} \right].$$

Сравним эту дисперсию, имеющую $f = k-1$ степеней свободы, с генеральной дисперсией σ^2 (метод такого сравнения см. в п. 6.3). Если s^2 от σ^2 отличается незначимо, то и влияние

*) Дисперсионный анализ был впервые разработан в двадцатых годах нашего столетия английским статистиком Р. Фишером; дальнейшее существенное развитие метод получил в трудах Изйтса. Предназначавшийся вначале для нужд сельскохозяйственной статистики дисперсионный анализ в настоящее время превратился в мощное орудие обработки самых различных наблюдений. Он охватывает большое число детально разработанных приемов планирования и обработки экспериментов, которых мы лишь слегка коснемся в настоящей книге.

фактора A нужно признать незначимым, так как он не сумел существенно увеличить случайный разброс наблюдений.

Если же s^2 отличается значимо от σ^2 , то это может быть вызвано только влиянием фактора A , которое теперь нужно признать значимым. Для того чтобы оценить σ_A^2 , воспользуемся тем, что дисперсия суммы двух независимых случайных величин равна сумме их дисперсий (см. п. 2.3). В нашем случае складываются эффект случайности (с дисперсией σ^2) и эффект воздействия фактора A (с дисперсией σ_A^2), которые очевидным образом независимы. Поэтому общая дисперсия наблюдений должна быть равна $\sigma^2 + \sigma_A^2$. А величина s^2 является оценкой этой общей дисперсией. Следовательно,

$$\sigma_A^2 \approx s^2 - \sigma^2.$$

Расчеты по указанной схеме весьма несложны. Пусть, например, изучается сопротивление электролита в зависимости от его температуры в пределах 10° — 60° С. Делая замеры через каждые 10° , получаем данные (в ом): 0,16; 0,19; 0,18; 0,20; 0,21; 0,23. Ошибка воспроизводимости известна из предыдущих опытов и равна $\sigma=0,02$.

Находим выборочную дисперсию $s^2=0,0006$. Отсюда $F = \frac{0,0006}{0,0004} = 1,5$. По таблице VII Приложения находим $F_{0,95}(5, \infty) = 2,2$. Найденное нами выше $F = 1,5 < 2,2$, следовательно, влияние изучаемого фактора (температуры) нужно считать незначимым. Иными словами, слишком высокий стандарт $\sigma=0,02$ не позволяет считать существенным рост наблюдаемого сопротивления — он вполне может оказаться случайным.

Подчеркнем в заключение, что мы оцениваем лишь влияние фактора в целом и не выясняем количественных соотношений. Поэтому замеры сопротивления можно было делать и не через равные промежутки температуры; можно вообще не отмечать температуру, лишь бы она менялась в нужных пределах.

8.2. Однофакторный дисперсионный анализ. Перейдем к анализу данных в случае, когда значение генеральной дисперсии σ^2 заранее неизвестно. Нужно найти такую схему анализа, которая позволила бы одновременно дать оценку и дисперсий изучаемых факторов, и дисперсии σ^2 .

Начнем с анализа одного фактора A , изучаемого на уровнях A_1, A_2, \dots, A_k . Для того чтобы дать оценку неизвестной дисперсии воспроизводимости σ^2 , нужно обязательно иметь дублирующие наблюдения при каждом уровне фактора A . Здесь можно поступать по-разному: можно на первом же уровне A_1 провести достаточно много наблюдений, чтобы сразу же определить дисперсию σ^2 и использовать найденное значение для изучения остальных уровней (как это было сделано в предыдущем пункте). Лучше, однако, равномерно повторять наблюдения на всех уровнях, ибо при этом появляется дополнительная возможность контроля за неизменностью дисперсии σ^2 .

Наиболее простые расчеты получаются в случае, когда на каждом уровне фактора A проводится одинаковое число наблюдений. Будем обозначать серию наблюдений на уровне A_i через $x_{i1}, x_{i2}, \dots, x_{in}$ (n — число повторных наблюдений на каждом уровне). Через \bar{x}_i будем обозначать среднее значений наблюдений на i -м уровне:

$$\bar{x}_i = \frac{x_{i1} + x_{i2} + \dots + x_{in}}{n}.$$

Кроме того, нам понадобится среднее всех наблюдений по всем уровням:

$$\bar{\bar{x}} = \frac{1}{kn} \sum_{i=1}^k \sum_{j=1}^n x_{ij} = \frac{1}{k} \sum_{i=1}^k \bar{x}_i$$

(легко видеть, что общее число всех наблюдений равно kn).

Определим общую выборочную дисперсию всех наблюдений:

$$\begin{aligned} s^2 &= \frac{1}{kn-1} \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{\bar{x}})^2 = \\ &= \frac{1}{kn-1} \left[\sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - \frac{\left(\sum_{i=1}^k \sum_{j=1}^n x_{ij} \right)^2}{kn} \right]. \end{aligned}$$

Эта дисперсия обязана своим появлением всем действующим факторам — как фактору A , так и фактору случайности на каждом уровне. Основная задача, которую решает диспер-

сионный анализ — это разложение общей дисперсии s^2 на составляющие, которые характеризовали бы фактор A и фактор случайности в отдельности.

Фактор случайности оценить нетрудно благодаря наличию повторных наблюдений на каждом уровне. Для уровня A_i выборочная дисперсия равна

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 = \frac{1}{n-1} \left[\sum_{j=1}^n x_{ij}^2 - \frac{\left(\sum_{j=1}^n x_{ij} \right)^2}{n} \right].$$

Пользуясь этой формулой, получаем серию выборочных дисперсий $s_1^2, s_2^2, \dots, s_k^2$, характеризующих фактор случайности на всех уровнях A_i . Если у нас нет априорной уверенности в том, что генеральная дисперсия воспроизводимости σ^2 одинакова на всех уровнях, то дисперсии $s_1^2, s_2^2, \dots, s_k^2$ нужно сравнить методами п. 6.3.

Если между дисперсиями s_i^2 нет значимых различий, то их все можно использовать для оценки генеральной дисперсии σ^2 по принципу «текущих измерений» (см. п. 4.4). Мы получим оценку

$$s_0^2 = \frac{1}{k} \sum_{i=1}^k s_i^2 = \frac{1}{k(n-1)} \left[\sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - \frac{1}{n} \sum_{i=1}^k \left(\sum_{j=1}^n x_{ij} \right)^2 \right],$$

имеющую $k(n-1)$ степеней свободы.

После того как найдена дисперсия s_0^2 , связанная со случайностью, можно уже дать приближенную оценку для дисперсии фактора A :

$$\sigma_A^2 \approx s^2 - s_0^2.$$

Эта оценка, однако, слишком груба из-за погрешностей величин s^2 и s_0^2 . Более точную оценку для σ_A^2 находят из следующих соображений. Влияние фактора A наиболее заметно на изменении средних \bar{x}_i по отдельным уровням. Действительно, дисперсия случайности для средних значений в n раз меньше, чем для отдельных наблюдений (см. п. 4.4). Поэтому

$$\frac{1}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2 \approx \sigma_A^2 + \frac{\sigma^2}{n} \approx \sigma_A^2 + \frac{s_0^2}{n}.$$

Отсюда легко определить дисперсию фактора A :

$$\sigma_A^2 \approx \frac{1}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2 - \frac{s_0^2}{n},$$

причем точность этой оценки выше, чем у предыдущей благодаря тому, что дисперсия s_0^2 входит в нее, деленная на n .

Введем обозначение:

$$s_A^2 = \frac{n}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2 \approx n\sigma_A^2 + s_0^2.$$

Эта дисперсия имеет $k-1$ степень свободы. Для того чтобы влияние фактора A было значимым, необходимо и достаточно, чтобы дисперсия s_A^2 значительно отличалась от s_0^2 . Сравнение s_A^2 и s_0^2 проводится по критерию Фишера, т. е. влияние фактора A признается значимым на уровне значимости p , если

$$\frac{s_A^2}{s_0^2} > F_{1-p},$$

где F -распределение рассматривается с $f_1 = k-1$ и $f_2 = k(n-1)$ степенями свободы*). Если же $\frac{s_A^2}{s_0^2} \leq F_{1-p}$, то влияние фактора

A нужно считать незначимым. В этом случае общая дисперсия s^2 связана только с фактором случайности и поэтому $\sigma^2 \approx s^2$. Последняя оценка дисперсии случайности лучше, чем s_0^2 , так как имеет $kn-1$ степень свободы (на $k-1$ степень больше, чем у s_0^2).

Проведение дисперсионного анализа связано с большими вычислениями, поэтому желательно применять вычислительную технику, упрощающие приемы. Перед началом вычислений все данные можно уменьшить на одну и ту же величину — на дисперсиях это не отразится.

Вычисления ведут по следующим этапам:

а) все данные располагают в таблицу 8.1, где через X_i обозначены суммы наблюдений по столбцам;

*) Здесь применяется односторонний критерий, так как $s_A^2 \geq s_0^2$.

Таблица 8.1

Номер наблюдения j	Уровни фактора			
	A_1	A_2	...	A_k
1	x_{11}	x_{21}	...	x_{k1}
2	x_{12}	x_{22}	...	x_{k2}
.
.
n	x_{1n}	x_{2n}	...	x_{kn}
Итого	X_1	X_2	...	X_k

б) находят сумму квадратов всех наблюдений:

$$Q_1 = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2;$$

в) находят сумму квадратов итогов по столбцам, деленную на число параллельных наблюдений:

$$Q_2 = \frac{1}{n} \sum_{i=1}^k X_i^2;$$

г) находят квадрат общего итога, деленный на число всех наблюдений:

$$Q_3 = \frac{1}{kn} \left(\sum_{i=1}^k X_i \right)^2;$$

д) вычисляют дисперсии s_A^2 и s_0^2 по формулам

$$s_0^2 = \frac{Q_1 - Q_2}{k(n-1)}, \quad s_A^2 = \frac{Q_2 - Q_3}{k-1}.$$

Проделав указанные вычисления, можно перейти к сравнению дисперсий s_A^2 и s_0^2 . Если их различие оказывается незначимым, то получаем оценку генеральной дисперсии

$$s^2 = \frac{Q_1 - Q_3}{kn - 1},$$

имеющую $kn-1$ степень свободы. Если же различие s_A^2 и s_0^2 оказывается значимым, то находим оценку влияния фактора A :

$$\sigma_A^2 \approx \frac{s_A^2 - s_0^2}{n}.$$

В качестве примера применим дисперсионный анализ к исследованию влияния нескольких различных катализаторов на выход конечного продукта заданной химической реакции. Обозначая катализаторы через A_1, A_2, \dots, A_k , получим уровни общего «фактора катализа» A . В таблице 8.2 приведены данные по выходу продукта реакции в граммах.

Таблица 8.2

Номер наблюдения	Катализаторы				
	A_1	A_2	A_3	A_4	A_5
1	3,2	2,6	2,9	3,7	3,0
2	3,1	3,1	2,6	3,4	3,4
3	3,1	2,7	3,0	3,2	3,2
4	2,8	2,9	3,1	3,3	3,5
5	3,3	2,7	3,0	3,5	2,9
6	3,0	2,8	2,8	3,3	3,1
Итоги	18,5	16,8	17,4	20,4	19,1

Непосредственные вычисления дают $Q_1=285,6$; $Q_2=284,7$; $Q_3=283,4$, откуда $s_0^2 = \frac{285,6-284,7}{25} = 0,036$, $s_A^2 = \frac{284,7-283,4}{4} = 0,325$. Найденные дисперсии сравним по критерию Фишера: $F = \frac{0,325}{0,036} = 9,03$. По таблице VII Приложения находим $F_{0,95}(4,25) = 2,8$. Мы видим, что $F > 2,8$, поэтому различие катализаторов следует признать значимым. Нетрудно оценить дисперсию «фактора катализа»:

$$\sigma_A^2 \approx \frac{0,325 - 0,036}{6} = 0,048.$$

Дисперсионный анализ фактически представляет собой сравнение нескольких средних, о котором мы уже говорили в п. 6.4. Это сравнение проводится в целом и поэтому является лишь первым этапом исследования. Обнаружив, что уровни фактора A в целом значимо различаются, мы можем перейти к попарному сравнению уровней с помощью критерия Стьюдента (см. п. 6.4). Так, в рассмотренном примере мы можем поставить вопрос, значимо ли различаются катализатор A_4 (дающий самый высокий выход) и следующий за ним катализатор A_5 . Рекомендуем читателям самостоятельно проверить, что различие между A_4 и A_5 значимо при $p=0,05$ и незначимо при $p=0,025$ (т. е., по существу, сомнительно).

Не всегда удается провести на каждом уровне фактора A одинаковое число наблюдений. При этом можно, конечно, ориентируясь на уровень с наименьшим числом повторных наблюдений, отбросить лишние наблюдения в остальных уровнях. Такое отбрасывание, однако, нежелательно, так как резко снизит точность проводимого анализа. Тем более, что однофакторный дисперсионный анализ с успехом можно проводить и при неравных столбцах (кстати, именно этот общий случай рассматривался в п. 6.4 при сравнении нескольких средних). Соответствующая схема вычислений лишь немногим отличается от случая равных столбцов.

Итак, пусть на уровне A_i было проведено n_i параллельных наблюдений и пусть $N = \sum_{i=1}^k n_i$ есть общее число всех наблюдений. Проводим следующие вычисления:

а) находим сумму квадратов всех наблюдений:

$$Q_1 = \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^2;$$

б) находим сумму квадратов итогов по столбцам, деленных на число наблюдений в соответствующем столбце:

$$Q_2 = \sum_{i=1}^k \frac{X_i^2}{n_i};$$

в) находим квадрат общего итога, деленный на число всех наблюдений:

$$Q_3 = \frac{1}{N} \left(\sum_{i=1}^k X_i \right)^2;$$

г) вычисляем дисперсии s_A^2 и s_0^2 по формулам:

$$s_0^2 = \frac{Q_1 - Q_2}{N - k}, \quad s_A^2 = \frac{Q_2 - Q_3}{k - 1}.$$

После указанных вычислений проверяют значимость отношения $\frac{s_A^2}{s_0^2}$ по критерию Фишера, исходя из $f_1 = k - 1$, $f_2 = N - k$ степеней свободы. Если это отношение незначимо, то получаем оценку генеральной дисперсии

$$s^2 = \frac{Q_1 - Q_3}{N - 1},$$

имеющую $N - 1$ степень свободы. Если же это отношение значимо, то вычисляем дисперсию фактора A по формуле

$$\sigma_A^2 \approx \frac{(k - 1) N}{N^2 - \sum_{i=1}^k n_i^2} (s_A^2 - s_0^2).$$

8.3. Двухфакторный дисперсионный анализ. Дисперсионный анализ особенно эффективен при одновременном изучении нескольких факторов. При классическом методе исследования подобное изучение проводят, варьируя лишь один фактор, а остальные оставляя неизменными. В связи с этим затрачивается много времени, ибо для каждого фактора проводится своя серия наблюдений, не используемая в дальнейшем при изучении других факторов. Есть у такого способа и еще один недостаток — он не позволяет изучать взаимодействие факторов при одновременном их изменении.

Всех этих недостатков лишен дисперсионный анализ, при котором каждое наблюдение служит для одновременной оценки всех факторов и их взаимодействий. Особенно ценно то, что при этом можно зачастую не делать параллельных наблюдений, ограничиваясь лишь одним наблюдением для каждого сочетания уровней изучаемых факторов.

Полное описание многофакторного анализа занимает очень много места. С добавлением каждого нового фактора усложняются таблицы и формулы для расчетов. Эти таблицы и формулы имеются в специальной литературе по дисперсионному анализу (см. указатель в конце книги), к которой мы и отошлем читателя, нуждающегося в них. В настоящей

же книге мы ограничимся лишь рассмотрением случая двух факторов, раскрыв на его примере основные идеи дисперсионного анализа *).

Итак, пусть изучаются одновременно два фактора A и B на уровнях A_1, A_2, \dots, A_k и B_1, B_2, \dots, B_m . Результаты наблюдений занесем в таблицу 8.3. Через X_i здесь обозначены

Таблица 8.3

$B \backslash A$	A_1	A_2	\dots	A_k	И т о г и
B_1	x_{11}	x_{21}	\dots	x_{k1}	X'_1
B_2	x_{12}	x_{22}	\dots	x_{k2}	X'_2
\dots	\dots	\dots	\dots	\dots	\dots
B_m	x_{1m}	x_{2m}	\dots	x_{km}	X'_m
И т о г и	X_1	X_2	\dots	X_k	

итоги данных по столбцам, через X'_j — по строкам. Через \bar{x}_i и \bar{x}'_j обозначим средние по столбцам и строкам, т. е.

$$\bar{x}_i = \frac{X_i}{m}, \quad \bar{x}'_j = \frac{X'_j}{k}.$$

Посмотрим, из каких компонент складывается рассеивание средних по строкам или по столбцам. Очевидно, на каждое такое рассеивание оказывает влияние лишь один из факторов A и B , так как все уровни второго фактора усреднены. Так, рассеивание \bar{x}_i (средние по столбцам) не зависит от фактора B , рассеивание \bar{x}'_j (средние по строкам) не зависит от фактора A . Кроме того, на всех рассеиваниях сказывается влияние фактора случайности с дисперсией

*) Об одном важном частном случае многофакторного анализа говорится в следующем пункте.

$\frac{\sigma^2}{m}$ для \bar{x}_i и $\frac{\sigma^2}{k}$ для \bar{x}'_j . Окончательно получаем, что

$$\frac{1}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2 \approx \sigma_A^2 + \frac{\sigma^2}{m}, \quad \frac{1}{m-1} \sum_{j=1}^m (\bar{x}'_j - \bar{\bar{x}})^2 \approx \sigma_B^2 + \frac{\sigma^2}{k}, \quad (8.1)$$

где через $\bar{\bar{x}}$ обозначено среднее всех данных таблицы.

Получившиеся равенства позволяют оценить дисперсии факторов A и B , если будет известна оценка генеральной дисперсии наблюдений σ^2 . Казалось бы, для оценки последней дисперсии нет никаких предпосылок, так как полностью отсутствуют параллельные наблюдения. И тем не менее, дисперсию σ^2 удастся оценить, сравнивая рассеивание средних с рассеиванием самих наблюдений.

Найдем дисперсию наблюдений по i -му столбцу:

$$s_i^2 = \frac{1}{m-1} \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2.$$

Эта дисперсия, очевидно, возникла под влиянием фактора B и фактора случайности, поэтому $s_i^2 \approx \sigma_B^2 + \sigma^2$. Равенство станет более точным, если s_i^2 заменить средневзвешенной дисперсией по всем столбцам, т. е.

$$\sigma_B^2 + \sigma^2 \approx \frac{1}{k} \sum_{i=1}^k s_i^2 = \frac{1}{k(m-1)} \sum_{i=1}^k \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2. \quad (8.2)$$

Полученная оценка содержит одновременно обе дисперсии, σ^2 и σ_B^2 . Ранее у нас была получена еще одна оценка, содержащая эти же дисперсии — это равенство (8.1). Вычитая (8.1) из (8.2), получим, что

$$\sigma^2 - \frac{\sigma^2}{k} \approx \frac{1}{k(m-1)} \sum_{i=1}^k \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2 - \frac{1}{m-1} \sum_{j=1}^m (\bar{x}'_j - \bar{\bar{x}})^2,$$

откуда

$$\sigma^2 \approx \frac{1}{(k-1)(m-1)} \left[\sum_{i=1}^k \sum_{j=1}^m (x_{ij} - \bar{x}_i)^2 - k \sum_{j=1}^m (\bar{x}'_j - \bar{\bar{x}})^2 \right].$$

Полученное равенство дает оценку искомой дисперсии σ^2 через некоторую величину, зависящую от всех наблюдений.

По своей природе эта величина является выборочной дисперсией с $(k-1)(m-1)$ степенями свободы; будем эту дисперсию обозначать через s_0^2 . Кроме того, введем обозначения

$$s_A^2 = \frac{m}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2 \approx m\sigma_A^2 + s_0^2,$$

$$s_B^2 = \frac{k}{m-1} \sum_{j=1}^m (\bar{x}'_j - \bar{\bar{x}})^2 \approx k\sigma_B^2 + s_0^2.$$

Величины s_A^2 и s_B^2 также можно считать выборочными дисперсиями с $k-1$ и $m-1$ степенями свободы соответственно.

Итак, мы получили все необходимые данные для проведения анализа. Окончательно схему вычислений можно представить следующим образом:

а) находят сумму квадратов всех наблюдений:

$$Q_1 = \sum_{i=1}^k \sum_{j=1}^m x_{ij}^2;$$

б) находят сумму квадратов итогов по столбцам, деленную на число наблюдений в столбце:

$$Q_2 = \frac{1}{m} \sum_{i=1}^k X_i^2;$$

в) находят сумму квадратов итогов по строкам, деленную на число наблюдений в строке:

$$Q_3 = \frac{1}{k} \sum_{j=1}^m X_j'^2;$$

г) находят квадрат общего итога, деленный на число всех наблюдений:

$$Q_4 = \frac{1}{mk} \left(\sum_{i=1}^k X_i \right)^2 = \frac{1}{mk} \left(\sum_{j=1}^m X_j' \right)^2$$

(наличие двойной формулы можно использовать для проверки правильности вычислений);

д) вычисляют дисперсии s_0^2 , s_A^2 , s_B^2 по формулам

$$s_0^2 = \frac{Q_1 + Q_4 - Q_2 - Q_3}{(k-1)(m-1)}, \quad s_A^2 = \frac{Q_2 - Q_4}{k-1}, \quad s_B^2 = \frac{Q_3 - Q_4}{m-1}.$$

После того как будут проделаны все необходимые вычисления, можно приступить к непосредственному анализу влияний факторов A и B . Для того чтобы влияние фактора A можно было признать значимым, нужно, чтобы s_A^2 значительно отличалось от s_0^2 ; то же самое справедливо для фактора B и дисперсии s_B^2 . Сравнение дисперсий, как обычно, проводится по критерию Фишера, т. е. вычисляются отношения $\frac{s_A^2}{s_0^2}$ и $\frac{s_B^2}{s_0^2}$ и сравниваются с табличными значениями F -распределения.

Допустим, что выбран уровень значимости p . Влияние фактора A признается значимым, если

$$\frac{s_A^2}{s_0^2} > F_{1-p},$$

где в F -распределении берутся $f_1 = k - 1$, $f_2 = (k - 1)(m - 1)$ степеней свободы. Дисперсия фактора A оценивается в этом случае равенством

$$\sigma_A^2 \approx \frac{s_A^2 - s_0^2}{m}.$$

Аналогично, влияние фактора B считается значимым, если

$$\frac{s_B^2}{s_0^2} > F_{1-p},$$

где на этот раз берутся $f_1 = m - 1$, $f_2 = (k - 1)(m - 1)$ степеней свободы. Дисперсия фактора B оценивается при этом равенством

$$\sigma_B^2 \approx \frac{s_B^2 - s_0^2}{k}.$$

Если $\frac{s_A^2}{s_0^2} \leq F_{1-p}$, то влияние фактора A нужно признать незначимым. В этом случае обе дисперсии s_A^2 и s_0^2 можно использовать для оценки генеральной дисперсии σ^2 . Это приведет к равенству

$$\sigma^2 \approx \frac{(k-1)s_A^2 + (k-1)(m-1)s_0^2}{(k-1) + (k-1)(m-1)} = \frac{Q_1 - Q_3}{m(k-1)},$$

которое дает более точную оценку для σ^2 , чем s_0^2 , ибо у этой оценки $m(k-1)$ степеней свободы (на $k-1$ степень больше, чем у s_0^2).

Аналогично, влияние фактора B признается незначимым, если $\frac{s_B^2}{s_0^2} \leq F_{1-p}$, и это дает оценку

$$\sigma^2 \approx \frac{(m-1)s_B^2 + (k-1)(m-1)s_0^2}{(m-1) + (k-1)(m-1)} = \frac{Q_1 - Q_2}{k(m-1)}.$$

Если же незначимыми оказываются влияния обоих факторов одновременно, то для дисперсии σ^2 получается самая лучшая оценка:

$$\sigma^2 \approx \frac{(k-1)s_A^2 + (m-1)s_B^2 + (k-1)(m-1)s_0^2}{(k-1) + (m-1) + (k-1)(m-1)} = \frac{Q_1 - Q_4}{mk-1},$$

имеющая $mk-1$ степень свободы. Точно такую же оценку мы получили бы, если бы отвлеклись от факторов A и B и все наблюдения рассматривали как единую выборку.

Изложенный анализ является исчерпывающим, если факторы A и B независимы. Если же между ними существует взаимодействие, то этому взаимодействию как фактору присуща своя дисперсия σ_{AB}^2 . В указанной выше схеме расчетов дисперсия σ_{AB}^2 как составная часть войдет в дисперсию s_0^2 . Выделить эту часть из дисперсии s_0^2 , к сожалению, невозможно без наличия параллельных наблюдений *).

Допустим теперь, что при каждом сочетании уровней факторов A и B проводится n параллельных наблюдений. Тогда при любых i и j на пересечении i -го столбца и j -й строки появится целая серия наблюдений $x_{ij1}, x_{ij2}, \dots, x_{ijn}$. Сохраним обозначение x_{ij} за средним каждой такой серии. Каждая серия имеет свою выборочную дисперсию,

$$s_{ij}^2 = \frac{1}{n-1} \sum_{v=1}^n (x_{ijv} - x_{ij})^2,$$

которой соответствует $n-1$ степень свободы. Поскольку на протяжении серии факторы A и B остаются неизменными,

*) При анализе трех и более факторов отдельные эффекты взаимодействия удается оценить и без параллельных наблюдений.

то s_{ij}^2 является оценкой дисперсии воспроизводимости σ^2 . Еще лучшей оценкой будет средневзвешенная дисперсия по всем сериям

$$s^2 = \frac{1}{mk} \sum_{i=1}^k \sum_{j=1}^m s_{ij}^2,$$

которой соответствуют уже $mk(n-1)$ степеней свободы.

Для того чтобы выделить теперь дисперсию взаимодействия, найдем, в каких пропорциях входят σ_{AB}^2 и s^2 в дисперсию s_0^2 . Дисперсия s_0^2 соответствует рассеянию средних x_{ij} , поэтому она складывается из дисперсии взаимодействия σ_{AB}^2 и дисперсии случайности, которая для средних x_{ij} равна $\frac{\sigma^2}{n}$ (см. п. 4.4). Таким образом,

$$s_0^2 \approx \sigma_{AB}^2 + \frac{\sigma^2}{n} \approx \sigma_{AB}^2 + \frac{s^2}{n}, \text{ откуда } \sigma_{AB}^2 \approx s_0^2 - \frac{s^2}{n}.$$

Разумеется, эффект взаимодействия факторов A и B может оказаться незначимым. Для проверки этого составляют отношение $\frac{ns_0^2}{s^2}$ и сравнивают его с табличным значением F_{1-p} , где берутся $f_1 = (k-1)(m-1)$ и $f_2 = mk(n-1)$ степеней свободы.

Наличие параллельных наблюдений используется только для отделения эффекта взаимодействия от дисперсии случайности σ^2 . Поэтому общая схема вычислений здесь лишь слегка дополняет ту схему, что приводилась выше. Вначале вычисляем сумму квадратов всех наблюдений

$$Q_5 = \sum_{i=1}^k \sum_{j=1}^m \sum_{v=1}^n x_{ijv}^2.$$

После этого отдельные наблюдения серий нам уже не нужны, и в таблицу можно занести только средние серий x_{ij} , так что получится вновь таблица 8.3. Над числами x_{ij} проводятся теперь все вычисления прежней схемы, т. е. вычисляются Q_1, Q_2, Q_3, Q_4 , а по ним дисперсии s_A^2, s_B^2 и s_0^2 . Так же, как описано выше, производится оценка значимости факторов A и B и вычисляются, в случае необходимости, дисперсии σ_A^2 и σ_B^2 .

Новые вычисления появляются лишь при оценке взаимодействия факторов A и B . Для этого находим

$$s^2 = \frac{Q_5 - nQ_1}{mk(n-1)},$$

затем сравниваем ns_0^2 и s^2 по F -критерию. Если отличие этих дисперсий окажется значимым, то нужно признать значимым эффект взаимодействия факторов A и B , причем

$$\sigma_{AB}^2 \approx \frac{ns_0^2 - s^2}{n}.$$

Если же отличие оказывается незначимым, то взаимодействием факторов можно пренебречь и оценки s_0^2 и s^2 объединить в одну:

$$\sigma^2 \approx \frac{(k-1)(m-1)s_0^2 + mk(n-1)s^2}{(k-1)(m-1) + mk(n-1)} = \frac{Q_4 + Q_5 - Q_2 - Q_3}{mkn - m - k + 1}.$$

Приведем пример дисперсионного анализа с двумя факторами. В четырех различных лабораториях производились испытания трех типов опреснителей морской воды. Каждое конкретное испытание дублировалось три раза, результаты испытаний выражены в % остаточной засоленности воды. Обозначая опреснители буквами A_1, A_2, A_3 и лаборатории — буквами B_1, B_2, B_3, B_4 , мы получаем два фактора: «фактор опреснителя» A и «фактор лаборатории» B . Для того чтобы выбрать лучший опреснитель, нужно оценить фактор A ; правильная интерпретация получающихся при этом результатов требует также учета фактора B , тем более, что различие данных по лабораториям может быть связано с различием местных условий (тип морской воды), с которым нужно считаться. Наконец, между факторами A и B может быть взаимодействие, так как опреснитель, «хороший» для одной лаборатории, может оказаться «плохим» для другой.

Таким образом, данная задача представляет собой классический образец для применения дисперсионного анализа. Составим вначале полную таблицу из всех данных (таблица 8.4). Определим сумму квадратов всех наблюдений $Q_5 = 431,32$.

Таблица 8.4

$\begin{matrix} A \\ B \end{matrix}$	A_1			A_2			A_3		
B_1	3,6	3,8	4,1	2,9	3,1	3,0	2,7	2,5	2,9
B_2	4,2	4,0	4,1	3,3	2,9	3,2	3,7	3,5	3,6
B_3	3,8	3,5	3,6	3,6	3,7	3,5	3,2	3,0	3,4
B_4	3,4	3,2	3,2	3,4	3,6	3,5	3,6	3,8	3,7

Таблица 8.5

$\begin{matrix} A \\ B \end{matrix}$	A_1	A_2	A_3	Итоги
B_1	3,83	3,00	2,70	9,53
B_2	4,10	3,13	3,50	10,73
B_3	3,63	3,60	3,20	10,43
B_4	3,27	3,50	3,70	10,47
Итоги	14,83	13,23	13,10	41,16

В дальнейшем нам понадобятся только средние параллельных наблюдений. Составим новую таблицу 8.5. В этой таблице на пересечении уровней A_i и B_j вместо прежних трех наблюдений стоит их среднее. Так, на пересечении уровней A_1 и B_1 стоит число

$$\frac{3,6 + 3,8 + 4,1}{3} = 3,83.$$

В нижней строке таблицы 8.5 записаны итоги по столбцам X_i , в правом столбце — итоги по строкам X'_j . В самом нижнем

правом углу записана общая сумма всех элементов:

$$\sum_{i=1}^k \sum_{j=1}^m x_{ij} = \sum_{i=1}^k X_i = \sum_{j=1}^m X'_j.$$

Легко видеть, что в нашем примере $k=3$, $m=4$.

Обработку полученной таблицы из средних ведем по всем правилам двухфакторного дисперсионного анализа без повторений. Для этого находим:

$$Q_1 = 3,83^2 + 3,00^2 + 2,70^2 + 4,10^2 + \dots + 3,70^2 = 142,83;$$

$$Q_2 = \frac{14,83^2 + 13,23^2 + 13,10^2}{4} = 141,64;$$

$$Q_3 = \frac{9,53^2 + 10,73^2 + 10,43^2 + 10,47^2}{3} = 141,45;$$

$$Q_4 = \frac{41,16^2}{3 \cdot 4} = 141,18.$$

Найдем теперь дисперсию воспроизводимости

$$s^2 = \frac{431,32 - 3 \cdot 142,83}{4 \cdot 3 \cdot 2} = 0,118$$

и совместную дисперсию воспроизводимости и взаимодействия

$$s_0^2 = \frac{142,83 + 141,18 - 141,64 - 141,45}{2 \cdot 3} = 0,153.$$

Чтобы проверить значимость эффекта взаимодействия, сравним дисперсии $ns_0^2 = 3 \cdot 0,153 = 0,459$ и $s^2 = 0,118$ по критерию Фишера. Имеем:

$$\frac{ns_0^2}{s^2} = \frac{0,459}{0,118} = 3,9,$$

в то время как табличное значение $F_{0,95}$ со степенями свободы $f_1=6$, $f_2=24$ равно 2,5. Таким образом, эффект взаимодействия факторов «опреснителя» и «лаборатории» оказывается значимым и оценивается дисперсией

$$\sigma_{AB}^2 = 0,153 - \frac{0,118}{3} = 0,114.$$

Оценим теперь влияние факторов A и B по отдельности. Для этого находим

$$s_A^2 = \frac{141,64 - 141,18}{2} = 0,23; \quad s_B^2 = \frac{141,45 - 141,18}{3} = 0,09.$$

Сравнивая по критерию Фишера дисперсии s_A^2 и s_0^2 , находим, что

$$\frac{s_A^2}{s_0^2} = \frac{0,23}{0,153} = 1,5,$$

в то время как табличное значение $F_{0,95}$ со степенями свободы $f_1=2$, $f_2=6$ равно 5,1. Это значит, что влияние фактора A (различие опреснителей) надо признать незначимым. Тем более надо признать незначимым влияние фактора B (различие лабораторий), так как $s_B^2 < s_0^2$.

Проведенный анализ показывает, что из всех факторов значимым является только фактор взаимодействия. Это значит, что лучший опреснитель можно выделить только для той или иной конкретной лаборатории, а в целом все опреснители одинаковы.

8.4. Планирование эксперимента при дисперсионном анализе. Дисперсионный анализ тесно связан с соответствующим планированием эксперимента. Удачно спланированный эксперимент, выявляя все необходимые эффекты, оказывается всегда либо более точным, либо менее трудоемким по сравнению с непродуманным экспериментом.

Если на результат эксперимента действуют одновременно несколько факторов, то наилучший эффект дает одновременный дисперсионный анализ всех этих факторов (многофакторный анализ). Можно, разумеется, проводить и попарное сравнение факторов, при котором все прочие факторы игнорируются. Однако такое построение эксперимента не позволяет выявить многие эффекты взаимодействий. Вдобавок, на такое исследование уйдет больше времени, так как переходя к каждой новой паре факторов, мы должны все испытания начинать заново, тогда как при дисперсионном анализе любое испытание служит одновременно исследованием всех факторов.

Эксперимент, спланированный так, что в нем встречаются все возможные сочетания уровней изучаемых факторов, называется *полным факторным экспериментом* (ПФЭ). Число испытаний при таком эксперименте равно произведению количеств уровней изучаемых факторов. Например, при исследовании трех факторов с 3, 5 и 6 уров-

ниями соответственно, общее число испытаний при ПФЭ равно $3 \cdot 5 \cdot 6 = 90$.

Методы дисперсионного анализа позволяют исследовать и такой случай, когда некоторые сочетания уровней пропущены. Такой эксперимент называется *дробным факторным экспериментом* (ДФЭ). Планирование при ДФЭ приобретает особо важную роль, ибо пропущенные сочетания уровней не так-то просто нейтрализовать.

ДФЭ дают возможность при неизменном числе испытаний исследовать гораздо большее число факторов, чем ПФЭ. Разумеется, при этом часть информации теряется, однако при удачно спланированном ДФЭ теряется именно та информация, которая по тем или иным причинам в данный момент несущественна.

Особенно широко используются ДФЭ, в которых теряется лишь информация о взаимодействиях изучаемых факторов. Такой ДФЭ можно применять в тех случаях, когда эффект взаимодействий заведомо отсутствует или настолько мал, что его можно не учитывать. ДФЭ, не учитывающие взаимодействий, подробно изучались Фишером и Иэйтсом; мы в этом пункте рассмотрим лишь простейшие случаи.

Пусть проводится двухфакторный анализ, рассматривавшийся в предыдущем пункте, причем число уровней для обоих факторов одинаково и равно k , т. е. таблица исходных данных (таблица 8.3) представляет собой квадрат. Эффектом взаимодействия мы не интересуемся, поэтому наблюдения при каждом сочетании уровней факторов A и B не дублируются. Этот эксперимент является ПФЭ, число испытаний равно k^2 .

Попытаемся, не меняя числа испытаний, присоединить к этим факторам третий фактор C , также имеющий k уровней. В результате получится ДФЭ; наша задача спланировать его так, чтобы можно было оценить эффекты действия всех трех факторов.

Очевидно, к каждому сочетанию уровней факторов A и B можно добавлять только один уровень фактора C , так что сравнивать различные сочетания уровней факторов A и B для изучения этих факторов станет невозможно: скажется разница и в факторе C . Но ведь мы в предыдущем пункте изучали факторы A и B не по отдельным сочетаниям уровней, а в целом — по строкам или столбцам. Каждая

строка (столбец) содержит уже k наблюдений, так что среди них можно разместить все k уровней фактора C . Тогда на всей строке (столбце) в целом фактор C не скажется, и средние по строкам и столбцам по-прежнему можно будет использовать для изучения факторов A и B .

Сразу же возникает первая задача: как расположить уровни фактора C , чтобы в каждой строке и каждом столбце встречались все уровни (и значит, каждый по разу). Такие расположения существуют и притом не единственные; согласно Фишеру их называют *латинскими квадратами* *). Эти расположения приводятся в специальных справочниках;

Таблица 8.6

	A_1	A_2	...	A_{k-1}	A_k
B_1	C_1	C_2	...	C_{k-1}	C_k
B_2	C_2	C_3	...	C_k	C_1
...
B_{k-1}	C_{k-1}	C_k	...	C_{k-2}	C_{k-1}
B_k	C_k	C_1	...	C_{k-2}	C_{k-1}

Таблица 8.7

	A_1	A_2	A_3	A_4	A_5
B_1	C_1	C_2	C_3	C_4	C_5
B_2	C_2	C_3	C_4	C_5	C_1
B_3	C_3	C_4	C_5	C_1	C_2
B_4	C_4	C_5	C_1	C_2	C_3
B_5	C_5	C_1	C_2	C_3	C_4

один вид такого квадрата легко построить без всяких справочников по типу таблицы 8.6. Например, при $k=5$ получается таблица 8.7. Рассматривая пересечение столбца с уровнем A_i и строки с уровнем B_j , мы видим, какой уровень фактора C к ним нужно добавлять.

Схема расчетов для латинского квадрата очень похожа на обычный двухфакторный анализ. А именно, обозначая,

*) Фишер обозначал уровни добавляемого фактора латинскими буквами; отсюда и получилось название квадрата.

как и ранее, через x_{ij} наблюдение, полученное при уровнях A_i и B_j , через X_i и X'_j — итоги по столбцам и строкам соответственно, мы получим следующий порядок анализа:

а) находят сумму квадратов всех наблюдений:

$$Q_1 = \sum_{i=1}^k \sum_{j=1}^k x_{ij}^2;$$

б) находят сумму квадратов по столбцам, деленную на число наблюдений в столбце:

$$Q_2 = \frac{1}{k} \sum_{i=1}^k X_i^2;$$

в) находят сумму квадратов итогов по строкам, деленную на число наблюдений в строке:

$$Q_3 = \frac{1}{k} \sum_{j=1}^k X_j'^2;$$

г) находят квадрат общего итога, деленный на число всех наблюдений:

$$Q_4 = \frac{1}{k^2} \left(\sum_{i=1}^k X_i \right)^2 = \frac{1}{k^2} \left(\sum_{j=1}^k X_j' \right)^2.$$

Все эти вычисления проводились и при обычном двухфакторном анализе. Однако теперь, в связи с необходимостью оценивать также фактор C , появляются и новые вычисления. Обозначим через Y_v сумму всех наблюдений, проводившихся при уровне C_v фактора C (независимо от того, какими при этом были уровни факторов A и B). К перечисленным выше этапам анализа добавляется еще один:

д) находят сумму квадратов итогов по уровням фактора C , деленную на число уровней:

$$Q_5 = \frac{1}{k} \sum_{v=1}^k Y_v^2.$$

Перейдем теперь к вычислению и оценке значимости дисперсий. Формулу для оценки дисперсии s_0^2 , связанной со случайностью наблюдений (дисперсия воспроизводимости), приведем, опуская несложный, но громоздкий

Вывод:

$$s_0^2 = \frac{Q_1 + 2Q_4 - Q_2 - Q_3 - Q_5}{(k-1)(k-2)};$$

число степеней свободы этой дисперсии равно $(k-1)(k-2)$. Отметим, что при наличии эффекта взаимодействия между факторами влияние взаимодействия войдет как составная часть в дисперсию s_0^2 .

После того как дисперсия s_0^2 вычислена, влияние факторов A и B оценивается точно так же, как и при простом двухфакторном анализе — ведь влияние фактора C у нас полностью нейтрализовано. Это значит, что вычисляются дисперсии

$$s_A^2 = \frac{Q_2 - Q_4}{k-1}, \quad s_B^2 = \frac{Q_3 - Q_4}{k-1},$$

имеющие по $k-1$ степеней свободы. Если эти дисперсии значимо (по критерию Фишера) отличаются от дисперсии s_0^2 , то действие факторов A и B оценивается дисперсиями

$$\sigma_A^2 \approx \frac{s_A^2 - s_0^2}{k}, \quad \sigma_B^2 \approx \frac{s_B^2 - s_0^2}{k}.$$

По этой же схеме оценивается и фактор C . А именно, вначале нужно вычислить дисперсию

$$s_C^2 = \frac{Q_5 - Q_4}{k-1},$$

имеющую $k-1$ степени свободы, затем сравнить ее с s_0^2 по критерию Фишера. Если отличие будет значимым, то

$$\sigma_C^2 \approx \frac{s_C^2 - s_0^2}{k}.$$

При анализе вещества (или продукции) к латинскому квадрату прибегают чаще всего тогда, когда в исследуемом материале имеется неоднородность. В этом случае факторы A и B связаны с самим исследованием (например, два этапа анализа или две стадии обработки продукции), а в качестве третьего фактора C рассматривается неоднородность материала.

Анализ с помощью латинского квадрата требует, чтобы число уровней всех факторов было одинаковым. Этого всегда

можно добиться, повторяя при эксперименте какие-либо из уровней недостающее число раз.

К исследуемым трем факторам A , B и C , не меняя общего числа испытаний k^2 , можно попытаться добавить еще и четвертый фактор D . Это нетрудно сделать, если удастся найти такое расположение уровней факторов C и D , при котором в каждой строке и каждом столбце имеются все k уровней фактора C и все k уровней фактора D и в то же время никакие два уровня факторов C и D не встречаются

Таблица 8.8

C_1 D_1	C_2 D_2	C_3 D_3
C_2 D_3	C_3 D_1	C_1 D_2
C_3 D_2	C_1 D_3	C_2 D_1

 $k=3$

C_1 D_1	C_2 D_2	C_3 D_3	C_4 D_4
C_2 D_3	C_1 D_4	C_4 D_1	C_3 D_2
C_3 D_4	C_4 D_3	C_1 D_2	C_2 D_1
C_4 D_2	C_3 D_1	C_2 D_4	C_1 D_3

 $k=4$

C_1 D_1	C_2 D_2	C_3 D_3	C_4 D_4	C_5 D_5
C_2 D_3	C_3 D_4	C_4 D_5	C_5 D_1	C_1 D_2
C_3 D_5	C_4 D_1	C_5 D_2	C_1 D_3	C_2 D_4
C_4 D_2	C_5 D_3	C_1 D_4	C_2 D_5	C_3 D_1
C_5 D_4	C_1 D_5	C_2 D_1	C_3 D_2	C_4 D_3

 $k=5$

по всей таблице вместе больше одного раза. Расположение такого типа называется *квадратом второго порядка*.

Строить квадраты второго порядка довольно трудно; приведем примеры (таблица 8.8) для случаев $k=3,4,5$ (обозначения уровней факторов A и B для краткости опущены).

Анализ для квадрата второго порядка ведется так же, как и при обычном латинском квадрате, только нужно еще ввести в рассмотрение суммы Z_μ наблюдений при фиксированных уровнях D_μ . К перечисленным выше суммам Q_1, \dots, Q_5 добавится

$$Q_6 = \frac{1}{k} \sum_{\mu=1}^k Z_\mu^2.$$

Дисперсия воспроизводимости оценится равенством

$$s_0^2 = \frac{Q_1 + 3Q_4 - Q_2 - Q_3 - Q_5 - Q_6}{(k-1)(k-3)},$$

дисперсии s_A^2 , s_B^2 и s_C^2 — так же, как и раньше. Добавится новая дисперсия:

$$s_D^2 = \frac{Q_6 - Q_4}{k-1},$$

по которой, в случае ее значимого отличия от s_0^2 , можно оценить действие фактора D :

$$\sigma_D^2 \approx \frac{s_D^2 - s_0^2}{k}.$$

Фишер и Иэйтс показали, что при k уровнях и число изучаемых факторов может быть доведено до k . Исключение составляет лишь $k=6$, для которого, кроме обычного латинского квадрата, нельзя построить никакого квадрата более высокого порядка.

Существует много других способов планирования дробного факторного эксперимента: неполные сбалансированные блоки, смешивание и т. д. О них можно прочесть в специальной литературе по дисперсионному анализу.

§ 9. ЗАВИСИМОСТЬ МЕЖДУ СЛУЧАЙНЫМИ ВЕЛИЧИНАМИ

9.1. Корреляция. В предыдущем изложении изучались наблюдения над одной случайной величиной. Между тем для выяснения тех или иных причинно-следственных связей в окружающей природе необходимо вести одновременные наблюдения над целым рядом случайных величин, чтобы по полученным данным изучать взаимоотношения этих величин. При каждом испытании основные факторы одинаковы для всех наблюдаемых величин, однако случайные факторы для каждой величины могут быть свои. В силу этого, зависимости между случайными величинами оказываются сильно «завуалированными» влиянием «своих» случайных факторов, и их выяснение возможно лишь методами математической статистики.

В настоящей книге мы ограничимся случаем, когда одновременно наблюдаются две случайные величины. Зависимости между большим числом величин можно изучать, объединяя их попарно. Изучение «совокупных» зависимостей представляет значительные технические трудности; в то же время принципиально новые методы (по сравнению с теорией двух величин) здесь почти не появляются.

В математическом анализе зависимость между двумя величинами выражается понятием функции $y=f(x)$, где каждому допустимому значению одной переменной соответствует одно и только одно значение другой переменной. Такая зависимость носит название *функциональной*; она обнаруживается с помощью строгих логических доказательств и не нуждается в опытной проверке. Если $y=\text{const}$ при изменении x , то говорят, что y не зависит от x ; всякое изменение y есть проявление зависимости от x . Так, например, угол правильного многоугольника зависит от числа сторон, но не зависит от их длины.

Гораздо сложнее обстоит дело с понятием зависимости случайных величин: если при изменении x изменилось y , мы не можем сразу сказать, является ли это изменение результатом зависимости от x или оно обязано лишь влиянию случайных факторов. Правда, и между случайными величинами может существовать строгая функциональная зависимость, устанавливаемая логическим путем. Например, число мужчин ξ и число женщин η на 1000 человек населения являются случайными величинами, однако всегда $\xi + \eta = 1000$. Подобного рода зависимость между случайными величинами обычно известна из теоретических соображений заранее, до всяких наблюдений. На практике она проявляется в том случае, когда для вычисления двух случайных величин используются одни и те же наблюдения (случай косвенных измерений), например, при нахождении числа мужчин и женщин на 1000 человек населения или при вычислении площади квадрата по измеренной стороне. Если же для вычисления каждой из случайных величин используются свои наблюдения, то на эти случайные величины действуют разные случайные факторы, и функциональная зависимость между ними уже невозможна. Попробуйте, например, отдельно измерять сторону a и площадь S квадрата и вы сами убедитесь, что на практике не всегда $S = a^2$ *).

Как правило, между случайными величинами может существовать лишь связь особого рода, при которой с изменением одной величины меняется распределение другой — такая связь называется *стохастической*. Изменение случайной величины η , соответствующее изменению величины ξ , разбивается при этом на две компоненты: *стохастическую* (связанную с зависимостью η от ξ) и *случайную* (связанную с влиянием «собственных» случайных факторов величин ξ и η). Если первая компонента отсутствует, то величины η и ξ независимы. Если же стохастическая компонента не равна нулю, то между η и ξ есть стохастическая связь. При этом соотношение между стохастической и случайной

*) От строгой функциональной следует отличать *ложную функциональную зависимость*, нередко возникающую на практике из-за того, что влияние случайных факторов оказывается слабее точности проводимых наблюдений.

компонентами определяет *силу связи* (понятие, лишенное смысла для функциональной зависимости). Наконец, отсутствие второй компоненты дает функциональную зависимость.

Выявление стохастической связи и оценка ее силы представляют важную и трудную задачу математической статистики. Достаточно сказать, что эта задача в общем виде не решена. Существуют показатели, оценивающие те или иные стороны стохастической связи. Из них важнейшим является *коэффициент корреляции*, рассматриваемый ниже.

В пункте 2.3. указывалось, что дисперсия суммы двух независимых величин равна сумме дисперсий этих величин. Поэтому если для двух случайных величин ξ и η окажется, что

$$D(\xi + \eta) \neq D\xi + D\eta,$$

то это служит верным признаком наличия зависимости между ξ и η . Таким образом, сравнивая дисперсию $D(\xi + \eta)$ с $D\xi + D\eta$, мы получаем первый критерий стохастической связи между ξ и η . Непосредственно из свойств дисперсии и математического ожидания (п. 2.3) вытекает, что

$$\begin{aligned} D(\xi + \eta) &= M[\xi + \eta - M(\xi + \eta)]^2 = \\ &= M[(\xi - M\xi)^2 + 2(\xi - M\xi)(\eta - M\eta) + (\eta - M\eta)^2] = \\ &= M(\xi - M\xi)^2 + 2M[(\xi - M\xi)(\eta - M\eta)] + M(\eta - M\eta)^2. \end{aligned}$$

Но

$$M(\xi - M\xi)^2 = D\xi, \quad M(\eta - M\eta)^2 = D\eta,$$

поэтому

$$D(\xi + \eta) - (D\xi + D\eta) = 2M[(\xi - M\xi)(\eta - M\eta)].$$

Итак, зависимость между ξ и η немедленно вытекает из неравенства

$$M[(\xi - M\xi)(\eta - M\eta)] \neq 0. \quad (9.1)$$

К сожалению, обратное утверждение несправедливо и из равенства $M[(\xi - M\xi)(\eta - M\eta)] = 0$ независимость ξ и η не вытекает. Это значит, что на дисперсии суммы сказывается не всякая стохастическая связь между слагаемыми. Может быть и так, что $D(\xi + \eta) \neq D\xi + D\eta$, но в этом неравенстве «повинна» лишь часть имеющейся связи между ξ и η .

Та часть стохастической связи между ξ и η , которая сказывается на отличии $D(\xi + \eta)$ от $D\xi + D\eta$, называется *корреляцией*. Необходимым и достаточным условием корреляции служит неравенство (9.1), в связи с чем величина $M[(\xi - M\xi)(\eta - M\eta)]$ носит название *корреляционного момента*. Корреляционный момент зависит от единиц измерения величин ξ и η . Поэтому на практике чаще используется безразмерная величина

$$\rho = \frac{M[(\xi - M\xi)(\eta - M\eta)]}{\sqrt{D\xi D\eta}},$$

которая называется *коэффициентом корреляции*.

Отметим некоторые важные свойства коэффициента корреляции. Прежде всего, коэффициент корреляции независимых величин равен нулю. Но, как уже указывалось, он может быть равен нулю и для некоторых зависимых величин, которые в этом случае называются *некоррелированными*. Коэффициент корреляции не меняется от прибавления к ξ и η каких-либо постоянных (неслучайных) слагаемых, от умножения ξ и η на положительные числа. Если же одну из величин, не меняя другой, умножить на -1 , то на -1 умножится и коэффициент корреляции.

Зная числа $D\xi$, $D\eta$ и ρ , можно найти дисперсии суммы и разности величин ξ и η :

$$D(\xi + \eta) = D\xi + D\eta + 2\rho \sqrt{D\xi D\eta},$$

$$D(\xi - \eta) = D\xi + D\eta - 2\rho \sqrt{D\xi D\eta}.$$

Из перечисленных выше свойств коэффициента корреляции вытекает, что этот коэффициент не изменится при переходе от ξ и η к нормированным величинам

$$\xi_0 = \frac{\xi - M\xi}{\sqrt{D\xi}}, \quad \eta_0 = \frac{\eta - M\eta}{\sqrt{D\eta}}.$$

Для нормированных величин $D\xi_0 = D\eta_0 = 1$ (см. п. 3.2), поэтому

$$D(\xi_0 + \eta_0) = 2 + 2\rho, \quad D(\xi_0 - \eta_0) = 2 - 2\rho.$$

Но дисперсия любой величины есть неотрицательное число, следовательно, справедливы два важных неравенства

$$1 + \rho \geq 0, \quad 1 - \rho \geq 0,$$

откуда

$$-1 \leq \rho \leq 1.$$

Если коэффициент корреляции отличен от нуля, то он своей величиной характеризует не только наличие, но и силу стохастической связи между ξ и η , точнее, той части этой связи, которую мы выше называли корреляцией. Чем больше абсолютная величина ρ , тем сильнее корреляция между ξ и η . Максимальная корреляция соответствует значениям $\rho = \pm 1$. Оказывается, это возможно только в случае, когда $\xi_0 = \pm \eta_0$, т. е. когда между величинами ξ_0 и η_0 (а значит, и между ξ и η) существует строгая функциональная связь. Действительно, при $\rho = \pm 1$ справедливо одно из равенств

$$D(\xi_0 + \eta_0) = 0, \quad D(\xi_0 - \eta_0) = 0,$$

а это значит, что либо $\xi_0 + \eta_0$, либо $\xi_0 - \eta_0$ есть постоянный нуль.

Итак, мы получили первый (и важнейший в силу своей простоты) показатель зависимости между случайными величинами ξ и η . Из коэффициента корреляции можно извлечь и еще одну информацию: если $\rho > 0$, то величины ξ и η с точностью до случайных погрешностей одновременно возрастают или убывают, если же $\rho < 0$, то с возрастанием одной величины другая убывает.

Тем не менее, коэффициент корреляции как показатель зависимости обладает серьезными недостатками. Мы уже упоминали, что из равенства $\rho = 0$ не следует независимость величин ξ и η . Оказывается, и крайние значения $\rho = \pm 1$ не очень полезны, так как соответствуют не всякой функциональной зависимости, а только строгой линейной связи между ξ и η . Действительно, из равенства $\xi_0 = \pm \eta_0$ следует, что

$$\frac{\xi - M\xi}{\sqrt{D\xi}} = \pm \frac{\eta - M\eta}{\sqrt{D\eta}},$$

т. е.

$$\eta = a\xi + b,$$

где

$$a = \pm \frac{\sqrt{D\eta}}{\sqrt{D\xi}}, \quad b = M\eta \mp \frac{\sqrt{D\eta}}{\sqrt{D\xi}} M\xi.$$

Таким образом, зависимость между ξ и η может быть строго функциональной (например, квадратичной), без следа случайности, а коэффициент корреляции все еще будет меньше 1 (по абсолютной величине), и корреляция будет неполной.

Исходя из сказанного, можно считать, что коэффициент корреляции есть показатель того, насколько связь между случайными величинами близка к строгой линейной зависимости. Он одинаково отмечает и слишком большую долю случайности, и слишком большую криволинейность этой связи.

Существуют, однако, такие случайные величины, для которых коэффициент корреляции является достаточно полным показателем зависимости. Сюда относятся в первую очередь величины, между которыми заранее, из общих соображений, можно предсказать линейную зависимость. Например, измеряя в электрической цепи одновременно напряжение и силу тока, мы должны, по закону Ома, ожидать между ними линейной зависимости (пропорциональности). Поэтому сильное отличие коэффициента корреляции ρ от 1 будет свидетельствовать о недостатках измерительных приборов или о наличии переменного сопротивления в цепи.

Сильно повышается ценность коэффициента корреляции и для величин, собственные случайные колебания которых подчиняются нормальному закону. Для таких величин, как это можно показать строго математически, отсутствие корреляции, т. е. равенство $\rho=0$, означает одновременно и отсутствие всякой зависимости.

Последнее свойство для нас особенно важно, так как мы при обработке наблюдений преимущественно сталкиваемся с нормальными распределениями. Возникает вопрос, как оценить коэффициент корреляции по данным наблюдений.

Допустим, что проведено m испытаний и при каждом отмечались значения двух случайных величин. В результате получатся m пар выборочных значений $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$. Для наглядности эти пары значений можно рассматривать как координаты точек на плоскости. Образовавшаяся совокупность точек сразу же даст нам представление о силе корреляции. На рис. 26 приведены примеры

совокупностей точек, соответствующих сильной (а), слабой (б) корреляции и полному ее отсутствию (с).

Выборочный коэффициент корреляции r вычисляется по такой же формуле, что и генеральный коэффициент ρ ,



Рис. 26.

только здесь берутся выборочные математические ожидания (средние) и дисперсии. Если через \bar{x} и \bar{y} обозначить средние значения для x_i и y_i :

$$\bar{x} = \frac{1}{m} \sum x_i, \quad \bar{y} = \frac{1}{m} \sum y_i,$$

то выборочный корреляционный момент равен

$$\frac{1}{m-1} \sum (x_i - \bar{x})(y_i - \bar{y}),$$

откуда

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(m-1) s_x s_y},$$

где через s_x и s_y обозначены выборочные дисперсии

$$s_x^2 = \frac{1}{m-1} \sum (x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{m-1} \sum (y_i - \bar{y})^2.$$

При вычислениях удобно пользоваться формулами

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{\sum x_i \sum y_i}{m},$$

$$(m-1) s_x^2 = \sum x_i^2 - \frac{1}{m} \left(\sum x_i \right)^2,$$

$$(m-1) s_y^2 = \sum y_i^2 - \frac{1}{m} \left(\sum y_i \right)^2.$$

При достаточно большом объеме выборки m выборочный коэффициент корреляции r приближенно равен генеральному коэффициенту ρ . Однако оценить возникающую при этом погрешность очень трудно. Это и не обязательно, так как точное значение ρ в расчетах почти не используется и нужно нам лишь как показатель силы связи. На практике выборочный коэффициент корреляции используют в основном для проверки общей гипотезы о наличии корреляции между наблюдаемыми величинами, не вдаваясь в подробные оценки силы этой корреляции.

В связи со случайностью выборки выборочный коэффициент корреляции r может быть отличен от нуля, даже если между наблюдаемыми величинами нет корреляции. Следовательно, для проверки гипотезы об отсутствии корреляции необходимо проверять, значимо ли отличается r от нуля. А для этого нужно знать распределение r как случайной величины. Это распределение, как и следовало ожидать, зависит от генерального коэффициента корреляции ρ , который нам неизвестен. Но если мы в качестве нулевой гипотезы возьмем равенство $\rho=0$ (отсутствие корреляции), то нам потребуется лишь r -распределение, соответствующее $\rho=0$. Такое распределение оказывается уже зависящим только от объема выборки m . Легко видеть, что оно симметрично относительно нуля и сосредоточено на отрезке $[-1, 1]$. В таблице XII Приложения приведены квантили $r_{1-p/2}$ этого распределения для некоторых уровней значимости p и объемов m .

Зная r -распределение, можно заранее предсказать доверительные границы для выборочного коэффициента корреляции r в предположении, что генеральная корреляция отсутствует. А именно, с вероятностью $1-p$ должна быть справедлива оценка

$$-r_{1-p/2} \leq r \leq r_{1-p/2}.$$

Поэтому если окажется, что найденный по выборке коэффициент корреляции удовлетворяет неравенству

$$|r| > r_{1-p/2},$$

то его нужно признать значимым, т. е. нужно считать, что нулевая гипотеза неверна. А это значит, что $\rho \neq 0$ и между наблюдаемыми величинами есть корреляция. Корреляция

эта будет тем сильнее, чем значительнее $|r|$ превышает $r_{1-p/2}$ и приближается к 1.

Оценка зависимости случайных величин по выборочному коэффициенту корреляции называется *корреляционным анализом*. Пример корреляционного анализа будет дан ниже, в п. 9.4. Сейчас лишь отметим в заключение, что наличие корреляции (да и всякой стохастической связи) еще не означает наличия причинно-следственной зависимости между величинами. Корреляция может возникнуть и в том случае, когда обе величины являются следствиями единой причины, не отмечавшейся при наблюдении. В качестве примера укажем на первые опыты Беккереля, который заметил, что интенсивность рентгеновских лучей тем выше, чем сильнее люминесцирует катодная трубка, испускающая эти лучи, и на основании этого ошибочно предполагал, что рентгеновские лучи являются следствием всякой люминесценции.

9.2. Регрессия. Корреляционный анализ служит фактически той же цели, что и рассматривавшийся в § 8 дисперсионный анализ — установлению значимости (неслучайности) изменения наблюдаемой случайной величины в процессе испытаний. При этом корреляционный анализ представляет собой как бы новую ступень, ибо одновременно оценивает и степень неслучайности наблюдаемых изменений.

Следующей, еще более высокой ступенью должно явиться выяснение точных количественных характеристик изменения случайной величины. Подобно тому как совокупность значений случайной величины описывалась набором неслучайных параметров, так и стохастическую, т. е. содержащую элемент случайности, связь нужно научиться выражать через строгие функциональные (неслучайные) соотношения.

Прежде всего покажем, что одну из наблюдаемых величин всегда можно считать неслучайной. Действительно, отмечая одновременно значения двух величин, ξ и η , мы при сопоставлении этих величин можем отнести все ошибки лишь к величине η . Таким образом, ошибка наблюдения $\Delta\eta$ будет складываться из собственной случайной ошибки величины η и из «ошибки сопоставления», возникающей из-за того, что с величиной η сопоставляется не совсем то значение ξ , которое имело место на самом деле.

Поясним сказанное на примере. Допустим, что изучается зависимость сопротивления электролита от температуры, и пусть при некотором испытании отмечены температура $t=20^\circ$ и сопротивление $R=15 \text{ ом}$. Поскольку температура и сопротивление измеряются отдельно своими приборами, то оба найденных числа содержат свои случайные ошибки, т. е. не являются истинными значениями. Тем не менее, мы можем считать, что температура в нашем опыте в точности равна 20° и случайным является лишь сопротивление. При этом значение $R=15 \text{ ом}$ будет приписано не «своей» температуре, ибо реальная температура равна, очевидно, какому-то $t_1 \neq t$. Если R окажется не зависящим от t , то для R безразлично, какой температуре его припишут. Если же R зависит от t , то малое изменение температуры вызовет соответствующее изменение сопротивления, что войдет как составная часть в общую ошибку величины R .

Итак, вместо сопоставления двух случайных величин мы можем теперь вести речь о зависимости некоторой случайной величины η от неслучайного параметра ξ . При каждом значении параметра случайная величина имеет какое-то распределение, которое называется *условным*. Если это распределение меняется от одного значения параметра к другому, то, значит, между ξ и η имеет место стохастическая связь.

Распределение случайной величины η при любом фиксированном значении x параметра ξ полностью определяется функцией распределения $F(y)$, которая в общем случае зависит и от x :

$$F(y) = F(x, y).$$

Зная функцию двух переменных $F(x, y)$, мы получаем полное описание зависимости η от ξ , притом выраженное в точных функциональных соотношениях.

Функция $F(x, y)$ носит название *условной функции распределения*; ее отыскание, как правило, оказывается весьма трудной задачей. Гораздо проще исследовать зависимость между ξ и η , когда условное распределение величины η при каждом значении параметра ξ подчинено нормальному закону. Нормальное распределение полностью определено средним и дисперсией, благодаря чему для описания зави-

симости η от ξ достаточно лишь указать, как при изменении параметра ξ изменяются среднее и дисперсия величины η .

Будем в дальнейшем значения параметра ξ обозначать через x , значения величины η — через y . Обозначая через a_y и σ_y^2 среднее и дисперсию величины η , мы приходим к необходимости отыскания двух функций:

$$a_y = f_1(x), \quad \sigma_y^2 = f_2(x).$$

Зависимость дисперсии σ_y^2 от параметра x носит название *скедастической зависимости*. Она характеризует изменение точности методики наблюдений при изменении параметра и используется редко. На практике обычно довольствуются «сводной» дисперсией, характеризующей рассеяние y по всем значениям параметра — нечто вроде средневзвешенной дисперсии по всем отдельным наблюдениям.

Гораздо важнее знать точную зависимость средних a_y от x , которая носит название *регрессии*. Поскольку средние a_y представляют собой истинные значения наблюдаемой величины η (п. 4.2), то регрессия дает *истинную зависимость* величин ξ и η , лишенную всех случайных наслоений. Поэтому идеальной целью всяких исследований зависимых величин является отыскание уравнения регрессии, а дисперсия используется лишь для оценки точности полученного результата.

Точное уравнение регрессии можно написать только, зная средние a_y для всех допустимых значений x параметра ξ . В практических наблюдениях такая ситуация, конечно, невозможна. Более того, даже отдельные значения средних a_y не могут быть найдены точно, а допускают лишь приближенные оценки. В связи с этим и мы можем искать лишь уравнения *приближенной регрессии*, оценивая тем или иным способом величину и вероятность этой приближенности.

В общем виде задача ставится так: по данной выборке $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$ найти уравнение приближенной регрессии и оценить допускаемую при этом ошибку. В дальнейшем уравнение приближенной регрессии мы будем записывать в виде $y=f(x)$, понимая под этим, что $a_y \approx f(x)$.

Уравнение приближенной регрессии существенно зависит от выбираемого нами принципа приближенности. Подбирать это уравнение мы должны по значениям (x_i, y_i) .

Поэтому здесь может возникнуть ошибка двоякого рода: во-первых, ошибка, связанная с тем, что y_i не совпадают с соответствующими средними a_j ; во-вторых, ошибка интерполяции, ибо самое точное уравнение для нескольких точек (а их у нас m штук) не гарантирует точности в промежуточных точках. Из сказанного видно, что бессмысленно гоняться за такими уравнениями, которые на всех парах чисел (x_i, y_i) давали бы точные равенства. Последним, к сожалению, грешат очень многие экспериментаторы — построив на координатной плоскости несколько найденных из опыта точек, они в качестве графика зависимости берут плавную кривую, непременно проходящую через все построенные точки!

Итак, метод интерполяции с заданных точек, весьма полезный при обработке неслучайных данных, в теории случайных величин неприменим. Необходимо искать принцип приближенности, который бы использовал существенные статистические свойства самого уравнения регрессии. В качестве такого принципа берут обычно *принцип наименьших квадратов*.

Принцип наименьших квадратов вытекает из общего положения математической статистики, согласно которому в качестве меры рассеяния всегда берется дисперсия (среднее из суммы квадратов отклонений). Этот принцип позволяет сравнивать, какое из двух произвольных уравнений дает лучшее приближение регрессии. Пусть, например, заданы два уравнения: $y=f_1(x)$ и $y=f_2(x)$. Согласно первому уравнению, значению x_i должно соответствовать значение $f_1(x_i)$, на самом же деле задано значение y_i , т. е. получается отклонение $y_i - f_1(x_i)$. Общей мерой рассеяния всех y_i вокруг функции $f_1(x)$ будет величина

$$D_1 = \frac{1}{m-l_1} \sum [y_i - f_1(x_i)]^2,$$

где l_1 — число связей, накладываемых на выборку функцией $f_1(x)$ (см. об этом ниже). Аналогично получается мера рассеяния вокруг функции $f_2(x)$:

$$D_2 = \frac{1}{m-l_2} \sum [y_i - f_2(x_i)]^2.$$

Из функций $f_1(x)$ и $f_2(x)$ мы должны теперь выбрать ту, у которой меньше соответствующая мера рассеяния D_1 или D_2 .

Принцип наименьших квадратов пригоден для сравнения любого числа функций. При этом удобнее всего сравнивать функции, накладывающие на выборку одинаковое число связей l , так как при этом можно сравнивать просто суммы квадратов отклонений (отсюда и название принципа). В общем виде этот принцип формулируется следующим образом.

Пусть задан некоторый класс функций $f(x)$, накладывающих на выборку одинаковое число связей. Тогда наилучшее уравнение приближенной регрессии дает та функция из рассматриваемого класса, для которой сумма квадратов

$$S = \sum [y_i - f(x_i)]^2$$

имеет наименьшее значение.

Чем шире выбранный класс, тем большее число функций участвует в сравнении и тем, казалось бы, надежнее будет сделанный выбор. Оказывается, это справедливо лишь до тех пор, пока не изменяется число связей l , накладываемых рассматриваемыми функциями на выборку. Если же класс расширяется за счет числа связей l , то это приводит к уменьшению числа степеней свободы $m-l$ и может ухудшить дисперсию

$$D = \frac{S}{m-l}.$$

В качестве класса сравниваемых функций берут обычно совокупность функций зафиксированного типа, но с произвольными коэффициентами, например, совокупность всех многочленов фиксированной степени k . Чтобы выбрать среди всех этих функций лучшую, нужно найти ее коэффициенты. Эти коэффициенты вычисляются по заданной выборке и поэтому каждый из них представляет собой некоторую связь, наложенную на выборку. Отсюда вытекает, что число связей l , накладываемых функцией $f(x)$ на выборку, равно числу неопределенных коэффициентов, входящих в аналитическое выражение этой функции. Например, для многочленов степени k число связей $l=k+1$.

Производя вычисления по принципу наименьших квадратов, мы сталкиваемся, таким образом, со следующим противоречием. С одной стороны, желательно рассматривать более широкий класс функций, например, многочлены с более высокой степенью. С другой стороны, такое расширение приводит к увеличению числа связей l и, значит, к увеличению дисперсии D . Выход отсюда может быть только один: провести тщательное предварительное исследование изучаемой зависимости с тем, чтобы в уравнение приближенной регрессии вошло лишь минимальное необходимое число неопределенных коэффициентов. В большинстве случаев удастся ограничиться уравнением первой степени

$$y = a + \beta x,$$

накладываемым на выборку лишь две связи.

Обычно уравнение приближенной регрессии ищут не сразу, а по этапам: сначала определяют один-два коэффициента, добавляя к ним постепенно новые члены. Например, вначале ищут уравнение первого порядка, затем добавляют квадратный член, член третьего порядка и т. д. Добавляемые члены получаются, как правило, со все меньшими коэффициентами, пока, наконец, эти добавки не перестанут быть значимыми по сравнению со случайными ошибками наблюдений. Кроме того, при каждой новой добавке нужно проверять, чтобы не увеличилась дисперсия D (за счет увеличения числа связей l), в противном случае от добавки нужно отказаться. Различные критерии, позволяющие проверять необходимость вводимых добавок, будут рассмотрены в следующем пункте.

9.3. Вычисление и анализ приближенной регрессии. Принцип наименьших квадратов позволяет полностью вычислить уравнение приближенной регрессии заданного типа (с неопределенными коэффициентами). Для этого вначале составляется сумма

$$S = \sum [y_i - f(x_i)]^2,$$

где функция $f(x)$ записана со всеми неопределенными коэффициентами $\alpha, \beta, \gamma, \dots$. Величину S можно теперь рассматривать как функцию от этих неопределенных коэффициентов. Задача состоит в том, чтобы найти набор коэффициентов $\alpha, \beta, \gamma, \dots$, минимизирующий величину S .

В математической статистике, как правило, рассматриваются лишь такие функции $f(x)$, которые дифференцируемы по всем своим коэффициентам. При этом условии отыскание минимизирующего набора коэффициентов превращается в несложную задачу математического анализа. Как известно, необходимым условием минимума дифференцируемой функции многих переменных $S(\alpha, \beta, \gamma, \dots)$ является выполнение равенств

$$\frac{\partial S}{\partial \alpha} = 0, \quad \frac{\partial S}{\partial \beta} = 0, \quad \frac{\partial S}{\partial \gamma} = 0, \quad \dots$$

Эти равенства можно рассматривать как уравнения относительно $\alpha, \beta, \gamma, \dots$; в математической статистике они называются *нормальными уравнениями*.

Если система нормальных уравнений имеет несколько решений, то не все решения будут соответствовать минимуму S . Здесь могут быть и точки максимума, и различные виды минимакса, так что для выделения минимума потребуются дополнительные исследования. Поскольку подобная ситуация встречается в математической статистике очень редко, мы не будем здесь заниматься ее подробным анализом.

Заметим, что величина $S \geq 0$ при любых $\alpha, \beta, \gamma, \dots$, и значит, у нее обязательно должен существовать хотя бы один минимум. Поэтому если система нормальных уравнений имеет единственное решение, то оно и является минимальным для величины S , и никаких дополнительных исследований проводить не нужно.

Используя правила дифференцирования, нормальным уравнениям можно придать следующий вид:

$$\left\{ \begin{array}{l} \sum 2[y_i - f(x_i)] \frac{\partial f(x_i)}{\partial \alpha} = 0, \\ \sum 2[y_i - f(x_i)] \frac{\partial f(x_i)}{\partial \beta} = 0, \\ \dots \dots \dots \end{array} \right.$$

или, после небольших изменений,

$$\left\{ \begin{array}{l} \sum y_i \frac{\partial f(x_i)}{\partial \alpha} - \sum f(x_i) \frac{\partial f(x_i)}{\partial \alpha} = 0, \\ \sum y_i \frac{\partial f(x_i)}{\partial \beta} - \sum f(x_i) \frac{\partial f(x_i)}{\partial \beta} = 0, \\ \dots \dots \dots \end{array} \right.$$

Покажем для примера, как составляются нормальные уравнения для регрессии второго порядка $y = \alpha + \beta x + \gamma x^2$. В этом случае

$$\frac{\partial f(x)}{\partial \alpha} = 1, \quad \frac{\partial f(x)}{\partial \beta} = x, \quad \frac{\partial f(x)}{\partial \gamma} = x^2.$$

Поэтому нормальные уравнения имеют такой вид:

$$\begin{cases} \sum y_i - \sum (\alpha + \beta x_i + \gamma x_i^2) = 0, \\ \sum y_i x_i - \sum (\alpha x_i + \beta x_i^2 + \gamma x_i^3) = 0, \\ \sum y_i x_i^2 - \sum (\alpha x_i^2 + \beta x_i^3 + \gamma x_i^4) = 0. \end{cases}$$

Относительно неизвестных коэффициентов α , β , и γ получилась линейная система уравнений третьего порядка; ее нетрудно решить, например, с помощью определителей.

После того как уравнение приближенной регрессии найдено, его надо подвергнуть статистическому анализу. Этот анализ состоит из двух существенно различающихся частей. Во-первых, нужно оценить ошибку от замены истинной регрессии приближенной, проверить значимость всех слагаемых найденного регрессионного уравнения в сравнении со случайной ошибкой наблюдений σ_y^2 — такое исследование носит название *регрессионного анализа*. Во-вторых, нужно оценить силу связи, т. е. долю регрессии в общем рассеянии значений y_i — такое исследование по причинам, указываемым в следующем пункте, называется *корреляционным анализом*.

Чтобы оценить степень приближенности найденного уравнения регрессии, находят оценки для всех коэффициентов α , β , γ , ... , входящих в это уравнение. Такая задача оказывается в большинстве случаев очень трудной, так как коэффициенты α , β , γ , ... как случайные величины (при случайной выборке) имеют весьма сложные распределения. В следующем пункте мы укажем, как оценивать коэффициенты α и β в регрессии вида $y = \alpha + \beta x$.

В тех случаях, когда число наблюдений m достаточно велико, распределения коэффициентов получаются близкими к нормальному и поэтому достаточно лишь оценить их дисперсии. Это можно сделать, если считать α , β , γ , ... косвенными измерениями и выразить их дисперсии через дисперсии величин x и y (п. 4.4).

Если все коэффициенты уравнения приближенной регрессии определяются сразу, регрессионный анализ заканчивается на оценке коэффициентов. Однако, как уже указывалось в предыдущем пункте, уравнение регрессии находят чаще всего не сразу, а путем последовательных уточнений. В регрессионный анализ входят при этом различные критерии, позволяющие принять или отвергнуть вводимую поправку.

Допустим, что первоначальное уравнение приближенной регрессии $y=f(x)$ предлагается заменить уточненным уравнением $y=f(x)+\varphi(x)$. Поправка $\varphi(x)$ может быть отвергнута по следующим двум причинам:

1) у нее может оказаться слишком много связей, накладываемых на выборку;

2) она может быть слишком мала по сравнению со случайной ошибкой наблюдений σ_y .

Проверка первой возможности является основным критерием и должна проводиться всегда. Вторая возможность поддается проверке только в том случае, когда есть достаточно надежная оценка для случайной ошибки наблюдений σ_y . Серьезным преимуществом второй проверки перед первой является то, что в первом случае мы испытываем лишь конкретную добавку $\varphi(x)$ и, в частности, должны ее вначале рассчитать по принципу наименьших квадратов, хотя она впоследствии может оказаться лишней. В то же время во втором случае проверяется только сама функция $f(x)$ — нужны ли ей вообще добавки? Поэтому наиболее рациональна следующая последовательность проверок. Сначала проверяют вторую возможность (если, конечно, для этого есть необходимые данные). Если проверка покажет излишность дальнейших добавок, то уравнение $y=f(x)$ считают окончательным. Если же окажется, что функция $f(x)$ нуждается в дальнейших уточнениях, переходят к анализу конкретных добавок $\varphi(x)$ по первой возможности.

Рассмотрим теперь методику проверки каждой из возможностей. Пусть функция $f(x)$ содержит l_1 связей, накладываемых на выборку, функция $f(x)+\varphi(x)$ содержит l_2 связей. Вычисляем дисперсии

$$D_1 = \frac{1}{m-l_1} \sum [y_i - f(x_i)]^2,$$

$$D_2 = \frac{1}{m-l_2} \sum [y_i - f(x_i) - \varphi(x_i)]^2.$$

Согласно общим положениям п. 9.2 поправка $\varphi(x)$ должна быть безусловно отвергнута, если окажется, что $D_2 \geq D_1$. Но даже и уменьшение дисперсии не означает, что поправку $\varphi(x)$ нужно принять — это уменьшение должно быть значимым (неслучайным). Значимость уменьшения дисперсии проверяется по критерию Фишера. А именно, уменьшение дисперсии нужно считать значимым (и, следовательно, принять поправку $\varphi(x)$), если окажется, что

$$\frac{D_1}{D_2} > F_{1-p}$$

где число F_{1-p} берется из таблицы VII Приложения соответственно степеням свободы $f_1 = m - l_1$, $f_2 = m - l_2$ и выбранному уровню значимости p .

Отметим, что указанный критерий способен отвергнуть лишь конкретную поправку $\varphi(x)$ — вместо нее может оказаться необходимой другая. Поэтому все намечаемые поправки необходимо заранее расположить в порядке их предполагаемого убывания. Это позволяет прекращать анализ после первой же отвергнутой поправки, так как остальные будут еще меньше. В частности, подбирая регрессию в виде многочлена, слагаемые нужно проверять в порядке возрастания степеней.

Перейдем к анализу второй возможности. Для этого анализа необходимо знать дисперсию $\sigma^2 = \sigma_y^2$, связанную со случайными ошибками наблюдений y_i (мы предполагаем, что все наблюдения ведутся с одинаковой точностью). Эту дисперсию можно оценить, если дублировать наблюдения y_i при каждом фиксированном x_i . Каждая i -я серия дублирующих наблюдений даст свою оценку дисперсии s_i^2 , после чего можно взять средневзвешенную дисперсию по всем i (см. п. 4.4). Обозначим найденную таким образом дисперсию через s^2 , ее число степеней свободы — через f . Это f равно числу всех наблюдений минус число серий.

Если f очень велико, то с достаточной степенью точности можно считать, что $s^2 = \sigma^2$. Может оказаться и так, что дисперсия σ^2 известна из большого числа предыдущих опытов. Во всех этих случаях мы и за генеральной дисперсией σ^2 сохраним прежнее обозначение s^2 , но будем считать, что число ее степеней свободы $f = \infty$.

Рассмотрим вновь дисперсию D_1 . В образовании этой дисперсии участвуют два фактора: рассеяние y_i вокруг истинной линии регрессии (т. е. вокруг своих средних a_y), описываемое дисперсией s^2 , и погрешность в определении приближенной регрессии $y=f(x)$, которой соответствует некоторая дисперсия $s_{\text{рег}}^2$. Поскольку эти факторы независимы, то

$$D_1 = s^2 + s_{\text{рег}}^2$$

Для того чтобы $s_{\text{рег}}^2$ была отлична от нуля, необходимо, чтобы дисперсия D_1 была больше s^2 , причем не просто больше, а, как обычно в математической статистике, больше значимо. Значимость эта вновь проверяется по критерию Фишера, т. е. ошибка в определении регрессии признается значимой только, если

$$\frac{D_1}{s^2} > F_{1-p},$$

где число F_{1-p} берется из таблицы VII Приложения соответственно степеням свободы $f_1=m-l_1$, $f_2=f$ и выбранному уровню значимости p . В противном случае нужно признать, что дисперсия D_1 образована только за счет случайных ошибок наблюдений и, следовательно, никакие добавки к функции $f(x)$ не способны эту дисперсию уменьшить.

В качестве иллюстрации к полученным правилам рассмотрим следующий пример. Дана выборка

$x:$	1	2	3	4	5	6	7	8	9	10
$y:$	2,2	3,1	4,1	5,0	5,8	6,9	7,8	9,0	10,2	11,1.

Из большого числа предыдущих наблюдений известна дисперсия случайных ошибок в наблюдениях y_i , равная $s^2=0,01$; согласно общему правилу, введенному выше, мы должны этой дисперсии приписать $f=\infty$ степеней свободы. Первоначальное уравнение приближенной регрессии имеет вид

$$y = 1 + x, \quad (9.2)$$

к нему предлагается добавить квадратичный член $0,001 x^2$. Требуется проверить необходимость этой поправки.

Проверим вначале, нуждается ли уравнение (9.2) вообще в каких-либо поправках. Для этого определим дисперсию

$$D_1 = \frac{1}{8} \sum (y_i - 1 - x_i)^2 = \frac{0,20}{8} = 0,025.$$

Выбрав уровень значимости $p=0,05$, сравним дисперсии D_1 и s^2 по критерию Фишера, учитывая, что у нас $f_1=8$, $f_2=\infty$. По таблице VII Приложения находим $F_{0,95}=2,0$. В то же время по нашим данным

$$\frac{D_1}{s^2} = \frac{0,025}{0,01} = 2,5 > 2,0.$$

Следовательно, отличие D_1 от s^2 является значимым и уравнение (9.2) нуждается в уточнениях.

Проверим теперь, дает ли требуемое уточнение предлагаемая поправка $0,001x^2$, для чего найдем

$$D_2 = \frac{1}{7} \sum (y_i - 1 - x_i - 0,001 x_i^2)^2 = \frac{0,21}{7} = 0,03.$$

Мы получаем, что $D_2 > D_1$, и, значит, предлагаемая поправка приводит даже к ухудшению уравнения регрессии. Отсюда можно сделать вывод, что поправка к уравнению (9.2), необходимость которой уже была доказана, должна иметь какой-то другой, более сложный вид. Методы отыскания лучшей поправки заданной степени будут изложены в п. 9.6.

Перейдем теперь к анализу силы связи между произвольными величинами ξ и η . Предположим, что уже найдено уравнение регрессии $y=f(x)$ и вычислена соответствующая дисперсия

$$D = \frac{1}{m-1} \sum [y_i - f(x_i)]^2.$$

Если считать, что уравнение регрессии найдено с достаточной точностью, то дисперсия D обусловлена только случайными ошибками наблюдений и $D \approx \sigma^2$.

Вычислим средние всех наблюдений x_i и y_i :

$$\bar{x} = \frac{1}{m} \sum x_i, \quad \bar{y} = \frac{1}{m} \sum y_i$$

и соответствующие «полные» дисперсии

$$s_x^2 = \frac{1}{m-1} \sum (x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{m-1} \sum (y_i - \bar{y})^2.$$

В то время как дисперсия s_x^2 соответствует только разбросу истинных значений x (в силу неслучайности ξ), дисперсия s_y^2 содержит и разброс истинных значений a_y , и случайный разброс наблюдений y_i , определяемый дисперсией σ^2 . Чем меньше доля σ^2 в общей дисперсии s_y^2 , тем сильнее связь между ξ и η , ибо тем меньше доля случайности в этой связи. Поэтому силу связи можно характеризовать числом

$$\zeta = \frac{(m-1) D}{(m-1) s_y^2}.$$

Легко видеть, что всегда $0 \leq \zeta \leq 1$, причем связь между ξ и η тем сильнее, чем меньше число ζ . При этом $\zeta=0$ только, если все $y_i = f(x_i)$, т. е. ни одна из заданных точек не отклоняется от линии регрессии. В случае большого числа точек m такую связь практически можно считать строго функциональной.

Если $\zeta \neq 0$, то в зависимости η от ξ появляется элемент случайности, который достигает своего максимума при $\zeta=1$, когда $\sigma^2 = s_y^2$. Однако и в этом, крайнем случае величины ξ и η нельзя считать независимыми, ибо зависимость между ними, не сказываясь на дисперсиях, может проявлять себя в моментах более высоких порядков. Разумеется, подобной ситуации не может быть в случае нормального распределения, где все моменты выражаются через среднее и дисперсию. Это значит, что для нормально распределенных случайных величин число ζ позволяет полностью решать вопрос об их зависимости.

9.4. Линейная регрессия. В этом пункте мы рассмотрим важный случай регрессии первого порядка

$$y = \alpha + \beta x,$$

которая называется *линейной регрессией*. Линейной регрессии следует ожидать в очень многих исследованиях, особенно там, где справедлив «закон равномерного накопления». Например, может быть известно, что изменение величины y связано с непосредственным изменением параметра x , но не зависит от того, какое «количество» параметра x уже накопилось. Пропорциональная зависимость величин часто вытекает из математических соображений или физических аналогий — короче говоря, исследователь почти всегда

должен представлять себе, с какой формой зависимости он имеет дело.

Если же об изучаемом явлении нет никаких косвенных сведений, кроме наблюдений, проводимых в настоящий момент, то предварительный характер зависимости можно выяснить, нанося данные в виде точек на координатной плоскости.

Наконец, во всех сложных случаях, когда регрессия заведомо будет нелинейной, изучение линейной регрессии можно использовать как первый этап исследования, с тем, чтобы в дальнейшем внести в нее необходимые поправки.

Пользуясь принципом наименьших квадратов, легко составить нормальные уравнения линейной регрессии:

$$\begin{cases} \sum y_i - \sum (\alpha + \beta x_i) = 0, \\ \sum y_i x_i - \sum (\alpha + \beta x_i) x_i = 0. \end{cases}$$

Делая простые преобразования, приводим эту систему к виду

$$\begin{cases} m\alpha + \beta \sum x_i = \sum y_i, \\ \alpha \sum x_i + \beta \sum x_i^2 = \sum y_i x_i. \end{cases}$$

Число β называется *коэффициентом регрессии*; его легко найти с помощью определителей:

$$\beta = \frac{m \sum x_i y_i - \sum x_i \sum y_i}{m \sum x_i^2 - (\sum x_i)^2}.$$

Число α называется *свободным членом регрессии*. Его тоже нетрудно найти с помощью определителей, но проще выразить его из первого уравнения через найденное уже β :

$$\alpha = \frac{\sum y_i - \beta \sum x_i}{m}. \quad (9.3)$$

Полученные формулы полностью определяют линейную регрессию по заданной выборке.

Равенство (9.3) для свободного члена регрессии можно переписать в виде

$$\alpha = \frac{1}{m} \sum y_i - \beta \frac{1}{m} \sum x_i = \bar{y} - \beta \bar{x},$$

откуда

$$\bar{y} = \alpha + \beta \bar{x}.$$

Мы получили, что *средняя точка* (\bar{x}, \bar{y}) *совместного распределения изучаемых величин всегда лежит на линии регрессии*. Отсюда вытекает, что для определения линии регрессии достаточно знать лишь ее угловой коэффициент β .

В определении линии регрессии участвуют две случайные величины: среднее \bar{y} (заменяющее α) и коэффициент регрессии β . Распределение \bar{y} есть обычное распределение среднего, его дисперсия в m раз меньше средневзвешенной дисперсии по всем отдельным наблюдениям y_i . Если наблюдения y_i не дублировались, то их дисперсия σ^2 должна быть известна из предыдущих наблюдений. В этом случае дисперсия для \bar{y} также равна $\frac{\sigma^2}{m}$.

Более трудно оценить коэффициент регрессии β , так как он выражается через элементы выборки с помощью довольно сложной формулы. Распределение коэффициента регрессии и некоторых связанных с ним величин изучалось многими авторами. Для нас будет удобен результат, полученный Бартлетом. А именно, он доказал, что величина

$$t = \frac{s_x \sqrt{m-2}}{s_y \sqrt{1-r}} (\beta - \beta_0),$$

где s_x и s_y — корни квадратные из дисперсий выборок x_i и y_i вокруг своих средних \bar{x} и \bar{y} , r — выборочный коэффициент корреляции; β_0 — коэффициент истинной регрессии, имеет распределение Стьюдента с $f=m-2$ степенями свободы. Используя квантили t -распределения, мы можем отсюда найти доверительные пределы для коэффициента истинной регрессии β_0 , ибо все остальные величины s_x , s_y , r , β определяются непосредственно по выборке. При доверительной вероятности $1 - p$ справедливы неравенства

$$-t_{1-p/2} \leq \frac{s_x \sqrt{m-2}}{s_y \sqrt{1-r}} (\beta - \beta_0) \leq t_{1-p/2},$$

откуда

$$\beta - t_{1-p/2} \frac{s_y \sqrt{1-r}}{s_x \sqrt{m-2}} \leq \beta_0 \leq \beta + t_{1-p/2} \frac{s_y \sqrt{1-r}}{s_x \sqrt{m-2}}.$$

После того как найдены доверительные пределы для

среднего \bar{y} (методами п. 6.1) и коэффициента регрессии, можно построить *доверительную область*, в которой с вероятностью $(1-p)^2$ лежит линия истинной регрессии. Обозначим найденные доверительные пределы для среднего через \bar{y}' и \bar{y}'' , для коэффициента регрессии — через β' и β'' . Через каждую из точек (\bar{x}, \bar{y}') и (\bar{x}, \bar{y}'') нужно провести две прямые с угловыми коэффициентами β' и β'' . Максимальная область, охватываемая этими прямыми, и представляет собой требуемую доверительную область. Пример построения дан на рис. 27.

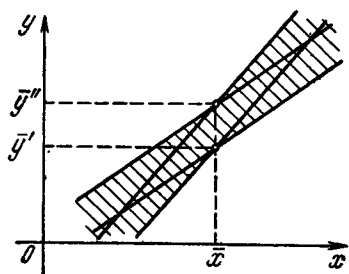


Рис. 27.

Перейдем к оценке силы найденной связи. Тот факт, что исследуемая зависимость предполагается линейной, позволяет использовать для оценки силы связи выборочный коэффициент корреляции r . Формула для вычисления r была указана в п. 9.1. Эту формулу можно упр-

простить, если использовать найденное уже значение коэффициента регрессии β .

Действительно,

$$\beta = \frac{m \sum x_i y_i - \sum x_i \sum y_i}{m \sum x_i^2 - (\sum x_i)^2} = \frac{\sum x_i y_i - m \bar{x} \bar{y}}{\sum (x_i - \bar{x})^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = r \frac{s_y}{s_x},$$

откуда

$$r = \frac{\beta s_x}{s_y}$$

или, в развернутом виде,

$$r = \beta \sqrt{\frac{m \sum x_i^2 - (\sum x_i)^2}{m \sum y_i^2 - (\sum y_i)^2}}.$$

Если коэффициент корреляции был вычислен ранее, то можно использовать обратную замену β на r . Мы получим уравнение регрессии в виде

$$y = a + r \frac{s_y}{s_x} x$$

или, заменяя α на $\bar{y} - \beta \bar{x}$, в виде

$$y - \bar{y} = r \frac{s_y}{s_x} (x - \bar{x}).$$

Из этого равенства хорошо видна роль коэффициента корреляции: чем меньше r , тем ближе линия регрессии к горизонтальному положению, т. е. тем ближе к состоянию неизменности будут средние a_y наблюдений y_i .

Линейная регрессия служит частным случаем регрессии вообще, поэтому для оценки ее силы можно также использовать введенное в предыдущем пункте число ζ . Возникает заманчивая перспектива исследовать силу линейной связи двумя методами, взаимно уточняя их результаты. К сожалению, это невозможно, так как для линейной регрессии число ζ однозначно выражается через коэффициент корреляции r .

Действительно, для линейной регрессии

$$\begin{aligned} D &= \frac{1}{m-2} \sum (y_i - \alpha - \beta x_i)^2 = \frac{1}{m-2} \sum [(y_i - \bar{y}) - \beta (x_i - \bar{x})]^2 = \\ &= \frac{1}{m-2} \left[\sum (y_i - \bar{y})^2 - 2\beta \sum (y_i - \bar{y})(x_i - \bar{x}) + \beta^2 \sum (x_i - \bar{x})^2 \right] = \\ &= \frac{1}{m-2} [(m-1)s_y^2 - 2\beta(m-1)r s_x s_y + \beta^2(m-1)s_x^2]. \end{aligned}$$

Заменяя коэффициент регрессии β его выражением через коэффициент корреляции r , получим

$$D = \frac{m-1}{m-2} [s_y^2 - 2r^2 s_y^2 + r^2 s_y^2] = \frac{m-1}{m-2} s_y^2 (1 - r^2),$$

откуда $\zeta = 1 - r^2$.

Таким образом, в случае линейной регрессии число $\theta = \sqrt{1 - \zeta}$ просто совпадает с коэффициентом корреляции. Поэтому и в случае произвольной регрессии вместо ζ предпочитают рассматривать число θ , которое называют тогда *корреляционным отношением*. Корреляционное отношение в криволинейной регрессии играет ту же роль, что и коэффициент корреляции в линейной, показывая тесноту группировки данных вокруг линии регрессии. Именно по этой причине анализ силы связи по θ называют *корреляционным*, какова бы ни была изучаемая регрессия.

Т а б л и ц а 9.1

Температура	0	10	20	30	40	50	60	70	80
Растворимость	33,5	37,0	41,2	46,1	50,0	52,9	56,8	64,3	69,9

Т а б л и ц а 9.2

Номер п/п	x	y	x^2	xy	y^2	$x+y$	$(x+y)^2$
1	0	33,5	0	0	1122,25	33,5	1122,25
2	10	37,0	100	370	1369,00	47,0	2209,00
3	20	41,2	400	824	1697,44	61,2	3745,44
4	30	46,1	900	1383	2125,21	76,1	5791,21
5	40	50,0	1600	2000	2500,00	90,0	8100,00
6	50	52,9	2500	2645	2798,41	102,9	10588,41
7	60	56,8	3600	3408	3226,24	116,8	13642,24
8	70	64,3	4900	4501	4134,49	134,3	18036,49
9	80	69,9	6400	5592	4886,01	149,9	22470,01
Суммы	360	451,7	20400	20723	23859,05		85705,05

Применим все описанные выше правила к исследованию зависимости растворимости тиосульфата натрия (в %) от температуры (в градусах), если на опыте были получены данные, приведенные в таблице 9.1. Обозначим температуру через x , растворимость через y . Уравнение регрессии будем искать в виде

$$y = a + \beta x.$$

Данные и результаты вычислений расположим в таблицу 9.2. Последние два столбца этой таблицы используются лишь для проверки вычислений по очевидной формуле

$$\sum (x_i + y_i)^2 = \sum x_i^2 + 2 \sum x_i y_i + \sum y_i^2.$$

В нашем случае

$$85705,05 = 20400 + 2 \cdot 20723 + 23859,05,$$

т. е. вычисления проведены правильно.

По выведенным ранее формулам найдем теперь β , α и r :

$$\beta = \frac{9 \cdot 20723 - 360 \cdot 451,7}{9 \cdot 20400 - 360^2} = \frac{23895}{54000} = 0,44;$$

$$\alpha = \frac{451,7 - 0,44 \cdot 360}{9} = \frac{293,3}{9} = 32,6;$$

$$r = 0,44 \sqrt{\frac{9 \cdot 20400 - 360^2}{2 \cdot 23859,05 - 451,7^2}} = 0,44 \sqrt{\frac{54000}{10699}} = 0,99.$$

Мы видим, что коэффициент корреляции очень близок к единице. А это значит, что зависимость между x и y является практически линейной, и окончательным уравнением регрессии нужно признать равенство

$$y = 32,6 + 0,44 x.$$

В заключение отметим, что все расчеты можно проводить, рассматривая вместо x_i , y_i отклонения

$$\Delta x_i = x_i - \bar{x}, \quad \Delta y_i = y_i - \bar{y}.$$

Уравнение регрессии примет тогда вид

$$y - \bar{y} = \beta (x - \bar{x}),$$

где для коэффициента регрессии получается простая формула

$$\beta = \frac{\sum \Delta x_i \Delta y_i}{\sum \Delta x_i^2}.$$

9.5. Нелинейная регрессия. Основным способом отыскания уравнения нелинейной регрессии (так же как и линейной) служит принцип наименьших квадратов. Это значит, что уравнение ищется в заданном классе функций и выборочные числовые данные используются лишь для определения неизвестных коэффициентов из системы нормальных уравнений. При этом различаются два случая: тип уравнения фиксируется сразу, так что принцип наименьших квадратов используется лишь один раз, или же уравнение регрессии в дальнейшем подвергается уточнениям, для чего принцип наименьших квадратов последовательно используется несколько раз.

Уравнение регрессии может быть известно заранее из соображений аналогии, из теоретических рассуждений или

из сравнения эмпирических данных с известными функциями. Разумеется, никакая уверенность в типе регрессии не освобождает от регрессионного и корреляционного анализов найденного уравнения. *Уравнение регрессии можно считать окончательным лишь тогда, когда соответствующая ему дисперсия D незначимо отличается от дисперсии случайных наблюдений σ^2* (см. п. 9.3).

Из курса математического анализа известно, что любая непрерывная функция может быть со сколь угодно высокой точностью заменена многочленом, при этом повышение точности достигается за счет повышения степени многочлена. Поэтому на практике любую регрессию можно искать в виде многочлена, находя его степень путем последовательных подсчетов. Эта степень, однако, может оказаться очень высокой, из-за чего в уравнении регрессии будет очень много неопределенных коэффициентов. Мы уже упоминали, что каждый такой коэффициент накладывает лишнюю связь на выборку, а это приводит, в конечном счете, к увеличению дисперсии D .

Если объем выборки n очень велик по сравнению с числом связей l , то увеличением числа связей можно пренебрегать, отыскивая регрессию именно в виде многочлена. Способ отыскания многочленной (параболической) регрессии будет дан ниже, в п. 9.6.

При малых выборках каждая лишняя связь заметно сказывается на точности приближенной регрессии. Поэтому здесь желательно уменьшать число неопределенных коэффициентов, используя другие (не степенные) элементарные функции. При этом может возникнуть регрессия показательного, логарифмического, дробно-степенного, тригонометрического и т. д. типов. Число коэффициентов в такой регрессии чаще всего не превышает двух — трех, что позволяет уменьшить дисперсию D .

Если для подбора многочленной регрессии существуют специальные приемы, то выбор типа *трансцендентной регрессии* сложен. Хорошо, если удастся использовать физические соображения и аналогии; в противном случае тип регрессии можно определить лишь путем кропотливого подбора.

Вычисление трансцендентной регрессии упрощается, если провести замену переменных, превращающую регрессию

в линейную. Например, зависимость показательного типа

$$y = \alpha \beta^x$$

превращается в линейную путем логарифмирования

$$\lg y = \lg \alpha + x \lg \beta.$$

С помощью логарифмирования становится линейной и зависимость дробно-степенного типа

$$y = \alpha x^\beta,$$

так как

$$\lg y = \lg \alpha + \beta \lg x.$$

Чтобы использовать здесь все приемы исследования линейной регрессии, исходные данные сразу же пересчитывают по формулам $\tilde{y}_i = \lg y_i$ для первого уравнения, $\tilde{y}_i = \lg y_i$, $\tilde{x}_i = \lg x_i$ — для второго. Если при возрастании x вторая переменная y убывает, то часто помогает замена $\tilde{x}_i = \frac{1}{x_i}$ или $\tilde{y}_i = \frac{1}{y_i}$.

Наиболее трудной задачей является подбор типа регрессии непосредственно по эмпирическим данным, когда теоретические предпосылки изучаемой зависимости совершенно неизвестны. Сведение к линейной зависимости здесь также оказывается одним из самых удобных приемов.

Рассмотрим в качестве примера изменение давления водяных паров P (в мм ртутного столба) в зависимости от температуры $t^\circ \text{C}$.

Соответствующие данные приведены в первых двух столбцах таблицы 9.3.

Измерения температуры проведены через равные промежутки $\Delta t = 10^\circ$. Если бы давление P выражалось через t линейно, то разности между соседними значениями P также должны были бы быть равными. Непосредственно данными таблицы это предположение отвергается.

Наша задача теперь — найти такое преобразование $\bar{P} = \varphi(P)$, после которого все разности между соседними значениями P станут одинаковыми. Тогда мы получим, что линейной зависимостью связаны \bar{P} и t ; найдем и исследуем

Таблица 9.3

t	P	\sqrt{P}	$\lg P$	$\lg^2 P$	$(1+0,01t) \lg P$	$(1+0,005t) \lg P$	Δ
0	0,46	0,68	-0,34	—	-0,34	-0,34	
10	0,91	0,95	-0,04	—	-0,04	-0,04	0,30
20	1,74	1,32	0,24	0,06	0,29	0,26	0,30
30	3,15	1,77	0,50	0,25	0,65	0,57	0,31
40	5,49	2,34	0,74	0,55	1,04	0,89	0,32
50	9,20	3,03	0,96	0,92	1,44	1,20	0,31
60	14,89	3,86	1,17	1,37	1,87	1,52	0,32
70	23,33	4,83	1,37	1,88	2,33	1,85	0,33
80	35,49	5,96	1,55	2,40	2,79	2,17	0,32
90	52,55	7,25	1,72	2,96	3,27	2,49	0,32
100	76,00	8,72	1,88	3,53	3,76	2,82	0,33
110	107,5	10,36	2,03	4,12	4,26	3,15	0,33

эту зависимость по правилам п. 9.4. После того как уравнение регрессии

$$\tilde{P} = \alpha + \beta t$$

будет найдено, мы сможем записать зависимость и между первоначальными переменными:

$$\varphi(P) = \alpha + \beta t.$$

Начиная с третьего, все столбцы таблицы 9.3 посвящены поискам преобразующей функции $\varphi(P)$. Поскольку P растет быстрее, чем линейно (разности между соседними значениями увеличиваются), характер этой функции должен быть замедляющим. Функция \sqrt{P} замедляет слишком слабо (разности между соседними значениями все еще растут), функция $\lg P$ — слишком сильно (разности начинают убывать). Пытаемся «исправить» функцию $\lg P$, возводя ее в квадрат, и снова рост становится слишком быстрым. Тогда начинаем добавлять к $\lg P$ медленно растущие множители $1+0,01t$ и $1+0,005t$. Из них наиболее удачным оказывается второй. Соответствующие ему разности между соседними значениями приведены в последнем столбце таблицы 9.3; видно, что они изменяются слабо.

Итак, производим замену

$$\tilde{P} = (1 + 0,005 t) \lg P$$

и ищем линейную зависимость между \bar{P} и t . По принципу наименьших квадратов получаем

$$\beta = 0,0318, \quad \alpha = -0,37, \quad r = 0,99,$$

откуда

$$\bar{P} = -0,37 + 0,0318 t.$$

Коэффициент корреляции r получился очень близким к единице; это значит, что найденная нами зависимость близка к строго функциональной.

Возвращаясь к первоначальным переменным, получаем

$$(1 + 0,005 t) \lg P = -0,37 + 0,0318 t,$$

или, после небольших преобразований,

$$P = 10^{\frac{8,36t - 74}{t + 200}}.$$

Отметим, что, несмотря на высокое значение коэффициента корреляции r , найденная зависимость является «удачной» лишь в вычислительном отношении, т. е. позволяет с высокой точностью находить давление P по заданной температуре t (в участке от 0° до 110°). Однако эта формула может не выражать никаких теоретических закономерностей. Многими исследователями найдены и другие формулы, выражающие зависимость давления водяных паров от температуры, которые, являясь не менее удобными для вычислений, содержат совсем другие типы функций. Так, например, Био рассматривал эту зависимость в виде

$$\ln P = \alpha + \beta \lambda^t - \gamma \mu^t,$$

Магнус — в виде

$$P = \alpha \beta^{t/(t+r)},$$

а в технике широко используют формулу вида

$$P = \alpha t^5 + \beta.$$

9.6. Параболическая регрессия. Перейдем теперь к методам отыскания *параболической регрессии*, т. е. регрессии, выражающейся уравнением

$$y = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_n x^n.$$

В случае, когда степень регрессии выбрана заранее, коэффициенты $\alpha_0, \alpha_1, \dots, \alpha_k$ отыскиваются непосредственно по принципу наименьших квадратов, и здесь не возникает никаких трудностей. Поэтому мы будем рассматривать лишь такой случай, когда степень многочлена заранее не известна и определяется, наряду с коэффициентами, по заданной выборке $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$; при этом будем считать, что все значения x_i равно отстоят друг от друга. При такой постановке задачи уравнение регрессии разыскивается путем последовательных уточнений. В качестве критерия прекращения вычислений рассматривается, как обычно, дисперсия

$$D_k = \frac{1}{m-k-1} \sum_{i=1}^m (y_i - \alpha_0 - \alpha_1 x_i - \dots - \alpha_k x_i^k)^2.$$

Как только D_{k+1} перестанет быть значимо меньше D_k , увеличение степени k нужно прекратить. Значимость различия между D_{k+1} и D_k проверяется по критерию Фишера (см. п. 9.3).

Итак, наша задача — найти методику последовательного вычисления коэффициентов $\alpha_0, \alpha_1, \dots, \alpha_k$ и дисперсий D_k . Казалось бы, здесь нет никакой проблемы — с помощью принципа наименьших квадратов можно вычислить коэффициенты многочлена любой наперед заданной степени. Будем последовательно определять регрессию первого, второго, третьего и т. д. порядков, попутно вычисляя дисперсии D_k , которые нам и покажут, когда следует остановиться.

Разумеется, такой путь является вполне законным. Но у него есть серьезный недостаток: при каждом увеличении степени все коэффициенты $\alpha_0, \alpha_1, \dots, \alpha_k$ нужно вычислять заново, а те, что были вычислены ранее, для меньших степеней, оказываются теперь абсолютно бесполезными. Иными словами, каждое уточнение уравнения регрессии затрагивает сразу все степени x .

Таким образом, все трудности заключаются в том, что при переходе от k -й степени к $(k+1)$ -й в правой части уравнения регрессии добавляется не просто одно слагаемое вида $\alpha_{k+1} x^{k+1}$, а целый многочлен $(k+1)$ -й степени, в котором содержатся $k+2$ новых неизвестных коэффициента. Но оказывается, в случае, когда все x_i равно отстоят друг от

друга, не все $k+2$ коэффициента в этой добавке могут быть произвольными. Достаточно вычислить один коэффициент, как все остальные определятся по нему совершенно однозначно. Происходит это потому, что добавляемый многочлен всегда имеет вид $\beta_{k+1}P_{k+1}(x)$, где лишь коэффициент β_{k+1} зависит от выборки, а многочлен $P_{k+1}(x)$ зависит только от объема выборки m . Зная многочлены-добавки $P_{k+1}(x)$, мы при каждом увеличении степени регрессии должны будем вычислять только один коэффициент β_{k+1} .

Многочлены $P_{k+1}(x)$ называются *многочленами Чебышева*. Первые два из них имеют вид

$$P_0(x) = 1, \quad P_1(x) = x - \frac{m+1}{2},$$

а остальные можно определять по формуле

$$P_{k+1}(x) = P_1(x)P_k(x) - \frac{k^2(m^2 - k^2)}{4(4k^2 - 1)}P_{k-1}(x).$$

Например,

$$P_2(x) = x^2 - (m+1)x + \frac{(m+1)(m+2)}{6},$$

$$P_3(x) = x^3 - \frac{3(m+1)}{2}x^2 + \frac{6m^2 + 15m + 11}{10}x - \frac{(m+1)(m+2)(m+3)}{20},$$

$$P_4(x) = x^4 - 2(m+1)x^3 + \frac{9m^2 + 21m + 4}{7}x^2 - \frac{(m+1)(2m^2 + 7m + 10)}{7}x + \frac{(m+1)(m+2)(m+3)(m+4)}{80}.$$

Общий вид уравнения регрессии k -го порядка будет теперь такой:

$$y = \beta_0 P_0(x) + \beta_1 P_1(x) + \dots + \beta_k P_k(x).$$

Коэффициенты $\beta_0, \beta_1, \dots, \beta_k$ можно найти с помощью метода наименьших квадратов. Опуская громоздкие промежуточные выкладки, сразу приведем окончательный результат:

$$\beta_0 = \frac{\sum y_i}{m}, \quad \beta_1 = \frac{\sum y_i P_1(x_i)}{\sum P_1^2(x_i)}, \quad \dots, \quad \beta_k = \frac{\sum y_i P_k(x_i)}{\sum P_k^2(x_i)}.$$

Хотя полученные формулы довольно громоздки (учитывая громоздкость выражений для $P_f(x)$) они обладают очень

важным качеством: вычисленные по ним коэффициенты β_j не зависят от того, каков будет порядок разыскиваемого уравнения регрессии. Это значит, что находя уравнение регрессии методом последовательных уточнений, мы используем все ранее найденные β_j , больше их не пересчитывая. Повышение порядка регрессии на 1 связано теперь с нахождением только одного коэффициента.

Очень удобными получаются формулы для нахождения дисперсий D_k . Для этого введем сокращенную запись

$$D_k = \frac{S_k}{m-k-1},$$

тогда суммы квадратов S_k можно последовательно искать по формуле

$$S_k = S_{k-1} - \beta_k^2 \sum P_k^2(x_i).$$

Здесь нужно только вначале вычислить S_0 :

$$S_0 = \sum [y_i - \beta_0 P_0(x)]^2 = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{m}.$$

Наиболее трудоемкой частью описанных вычислений является нахождение коэффициентов β_j . Эту работу можно облегчить, если полнее использовать тот факт, что значения x_i являются равноотстоящими. Обозначим разность между двумя соседними значениями x_i и x_{i+1} через h (шаг интерполяции). Тогда $x_2 = x_1 + h$, $x_3 = x_1 + 2h$, ..., $x_m = x_1 + (m-1)h$. Если сделать замену переменных

$$z = \frac{x - x_1}{h} + 1,$$

то каждое значение x_i заменится своим номером $z_i = i$. Уравнение регрессии запишется теперь в виде

$$y = \beta_0 P_0(z) + \beta_1 P_1(z) + \dots + \beta_k P_k(z),$$

где

$$\beta_0 = \frac{\sum y_i}{m}, \quad \beta_1 = \frac{\sum y_i P_1(i)}{\sum P_1^2(i)}, \quad \dots, \quad \beta_k = \frac{\sum y_i P_k(i)}{\sum P_k^2(i)}.$$

Числа $P_k(i)$ не зависят от элементов выборки и их можно рассчитать раз и навсегда для каждой пары значений k и

$m^*)$. Знаменатели дробей посчитать еще проще, используя сокращенную формулу $**$)

$$\sum P_k^2(i) = \frac{(k!)^2 m (m^2 - 1) (m^2 - 4) \dots (m^2 - k^2)}{[(2k - 1)!!]^2 2^{2k} (2k + 1)}.$$

В частности,

$$\sum P_1^2(i) = \frac{m(m^2 - 1)}{12}, \quad \sum P_2^2(i) = \frac{m(m^2 - 1)(m^2 - 4)}{180},$$

$$\sum P_3^2(i) = \frac{m(m^2 - 1)(m^2 - 4)(m^2 - 9)}{2800},$$

$$\sum P_4^2(i) = \frac{m(m^2 - 1)(m^2 - 4)(m^2 - 9)(m^2 - 16)}{44100}.$$

Эти же суммы будут нужны и при вычислении сумм S_k :

$$S_k = S_{k-1} - \beta_k^2 \sum P_k^2(i).$$

После того как уравнение регрессии будет найдено и исследовано, переменную z опять заменяют первоначальной переменной x .

Применим описанный в этом пункте метод Чебышева к исследованию зависимости степени диссоциации δ йодистого водорода от температуры $t^\circ \text{C}$, если известны следующие опытные данные:

t :	280	300	320	340	360	380	400	420	440	460	480
δ :	0,178	0,182	0,186	0,191	0,196	0,202	0,207	0,213	0,220	0,228	0,236.

Значения температуры отстоят друг от друга все время на 20° . Сделаем замену

$$z = \frac{t - 260}{20},$$

после которой каждое t_i заменится своим номером $z_i = i$. Одновременно для удобства вычислений заменим и δ :

$$y = 1000(\delta - 0,178).$$

*) Эти числа можно найти в специальных статистических таблицах, посвященных обработке зависимых случайных величин.

**) Выражение $(2k-1)!!$ означает произведение всех нечетных чисел от 1 до $2k-1$ включительно.

Т а б л и ц а 9.4

$z_i=i$	y_i	$P_1(i)$	$y_i P_1(i)$	$P_2(i)$	$y_i P_2(i)$	$P_3(i)$	$y_i P_3(i)$
1	0	-5	0	15	0	-36	0
2	4	-4	-16	6	24	7,2	28,8
3	8	-3	-24	-1	-8	26,4	211,2
4	13	-2	-26	-6	-78	27,6	358,8
5	18	-1	-18	-9	-162	16,8	302,4
6	24	0	0	-10	-240	0	0
7	29	1	29	-9	-261	-16,8	-487,2
8	35	2	70	-6	-210	-27,6	-966
9	42	3	126	-1	-42	-26,4	-1108,8
10	50	4	200	6	300	-7,2	-360
11	58	5	290	15	870	36	2088
Суммы	281		631		193		67,2

Полученные после замен значения z_i и y_i даны в первых двух столбцах таблицы 9.4; между этими величинами мы и будем искать зависимость.

В нашем примере $m=11$, поэтому $P_1(x)=x-6$. Подставляя $i=1, 2, \dots, 11$ вместо x , найдем все значения $P_1(i)$ — они представлены в третьем столбце таблицы 9.4.

В четвертом столбце даны $y_i P_1(i)$. Непосредственно по второму столбцу находим

$$\beta_0 = \frac{281}{11} = 25,6.$$

По сокращенной формуле

$$\Sigma P_1^2(i) = \frac{11(121-1)}{12} = 110,$$

после чего нетрудно найти β_1 :

$$\beta_1 = \frac{631}{110} = 5,73.$$

Найденные числа позволяют записать уравнение приближенной регрессии первого порядка:

$$y = 25,6 + 5,73(z - 6).$$

Определим теперь дисперсию D_1 . Для этого вычисляем

$$S_0 = \sum y_i^2 - \frac{(\sum y_i)^2}{11} = 10843 - \frac{78961}{11} = 3665,$$

$$S_1 = S_0 - \beta_1^2 \sum P_1^2(i) = 3665 - 5,73^2 \cdot 110 = 53,$$

откуда

$$D_1 = \frac{S_1}{m-2} = \frac{53}{9} = 5,9.$$

Переходим к вычислению регрессии второго порядка. В нашем случае $P_2(x) = x^2 - 12x + 26$, откуда находим числа $P_2(i)$, приведенные в пятом столбце таблицы 9.4; в шестом столбце приведены $y_i P_2(i)$. Вновь по сокращенной формуле находим

$$\sum P_2^2(i) = \frac{11(121-1)(121-4)}{180} = 858,$$

что позволяет определить

$$\beta_2 = \frac{193}{858} = 0,225.$$

В результате можно записать уравнение приближенной регрессии второго порядка:

$$y = 25,6 + 5,73(z - 6) + 0,225(z^2 - 12z + 26).$$

Вычислим дисперсию D_2 :

$$D_2 = \frac{S_2}{m-3} = \frac{S_1 - \beta_1^2 \sum P_2^2(i)}{m-3} = \frac{53 - 0,225^2 \cdot 858}{8} = \frac{9,6}{8} = 1,2.$$

Проверим, является ли уравнение второго порядка существенно лучшим приближением регрессии, чем уравнение первого порядка. Для этого сравним по критерию Фишера дисперсии D_1 , имеющую $f_1=9$, и D_2 , имеющую $f_2=8$ степеней свободы. По таблице VII Приложения находим $F_{0,95}(9,8)=3,4$. В нашем же случае

$$\frac{D_1}{D_2} = \frac{5,9}{1,2} = 4,9 > 3,4.$$

Значит, уравнение второго порядка является существенным уточнением первоначального уравнения первого порядка.

Начинаем определять регрессию третьего порядка. Подставляя $m=11$, находим $P_3(x)=x^3-18x^2+90,2x-191,1$. Соответствующие значения $P_3(i)$ и $y_i P_3(i)$ приведены в седьмом и восьмом столбцах таблицы 9.4. По сокращенной формуле находим

$$\sum P_3^2(i) = \frac{11(121-1)(121-4)(121-9)}{2800} = 6177,6,$$

откуда

$$\beta_3 = \frac{67,2}{6177,6} = 0,0109,$$

и, значит,

$$y = 25,6 + 5,73(z-6) + 0,225(z^2-12z+26) + \\ + 0,0109(z^3-18z+90,2z-191,1).$$

Проверим существенность перехода от регрессии второго к регрессии третьего порядка. Для этого находим

$$D_3 = \frac{S_3}{m-4} = \frac{S_2 - \beta_3^2 \sum P_3^2(i)}{m-4} = \frac{9,6 - 0,0109^2 \cdot 6177,6}{7} = \frac{8,9}{7} = 1,27.$$

Мы видим, что $D_3 > D_2$. Следовательно, переход к уравнению третьего порядка ухудшает приближение регрессии и нужно остановиться на регрессии второго порядка.

Делая обратную замену от z и y к t и δ , получаем соотношение

$$1000(\delta - 0,178) = 25,6 + 5,73 \left(\frac{t-260}{20} - 6 \right) + \\ + 0,225 \left[\left(\frac{t-260}{20} \right)^2 - 12 \frac{t-260}{20} + 26 \right]$$

или, после преобразований

$$\delta = 0,1731 - 0,000139t + 0,00000056t^2.$$

Это уравнение и выражает зависимость степени диссоциации йодистого водорода от температуры.

Найденное нами уравнение регрессии является окончательным в классе многочленов. Но это не значит, что оно вообще не нуждается в уточнениях. Еще раз напомним, что окончательность приближения регрессии определяется только сравнением дисперсии D и дисперсии случайных

наблюдений σ^2 (для отыскания которой нужны параллельные наблюдения). Может случиться так, что наше уравнение, не нуждающееся в добавочных членах третьего порядка, допускает уточняющую добавку типа $t^2 \ln t$.

В заключение оценим силу найденной связи с помощью корреляционного отношения. В нашем примере

$$\xi = \frac{S_2}{S_0} = \frac{9,6}{3665} = 0,00272,$$

откуда

$$\theta = \sqrt{1 - \xi} = \sqrt{1 - 0,00272} = 0,986.$$

Корреляционное отношение θ близко к 1, значит, найденная нами связь близка к строгой функциональной.

§ 10. НЕКОТОРЫЕ ВОПРОСЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

10.1. Способы отбора. Мы уже упоминали (п. 4.3), что применение методов математической статистики к обработке наблюдений основано на глубокой аналогии между производством наблюдений и отбором из некоторой генеральной совокупности. При этом в качестве генеральной совокупности рассматривается чисто гипотетическая совокупность всех возможных результатов наблюдений при данном комплексе условий испытания. Отбор данных из этой совокупности производится в процессе наблюдений независимо от нашей воли. Благодаря этому основным фактором отбора наблюдений всегда является случайность, что и позволяет применять для обработки наблюдений основные положения теории вероятностей.

Отбор данных, происходящий помимо нашей воли, можно назвать *естественным отбором*; именно таким является отбор данных при наблюдениях. Однако при производстве наблюдений и при дальнейшей их обработке часто возникает необходимость и в других, искусственных способах отбора. Например, при анализе вещества приходится делать пробы из разных участков, чтобы нейтрализовать возможную неоднородность материала; отбор проб при этом находится целиком в нашем распоряжении. При контроле за производством приходится выбирать образцы из общей продукции. При различных экономических и демографических исследованиях также нужен предварительный отбор объектов для изучения. Примеры таких случаев, когда исследователь вынужден делать сознательный выбор, можно еще долго продолжать; отметим лишь в заключение, что даже заготовленный уже цифровой материал может нуж-

даться в дополнительном отборе в целях сокращения объема, в целях удаления неподходящих данных и в целях проверки правильности и добросовестности полученных данных и подготовительных расчетов — подобный отбор также производится исследователем целиком по его воле.

Существует много способов искусственного отбора; выбор того или иного способа зависит от цели отбора, от поставленной задачи. В самом общем плане способы отбора делятся на две группы: *пристрастные* и *репрезентативные*.

Пристрастными называются такие способы отбора, при которых отбираются элементы по какому-либо заранее намеченному признаку; при этом проверке подлежат все элементы совокупности, из которой делается отбор. Например, из совокупности чисел отбираются n самых больших или отбираются все числа, не достигающие нужной величины. Пристрастный отбор применяют и для того, чтобы изъять все наблюдения с нарушенными условиями испытания. Так, при изучении жирности молока новой породы коров стараются отобрать и в дальнейшем не учитывать данные по больным и смешанно-породным особям. Пристрастный отбор является важной стадией эксперимента; его задача при этом состоит обычно в том, чтобы устранить все заметные нарушения условий испытания. С помощью пристрастного отбора нередко удается ликвидировать те или иные доминирующие (несимметричные) факторы, нарушающие нормальность распределения.

Пристрастный отбор всегда является сознательным с ясной характеристикой данных, подлежащих отбору, поэтому он редко вызывает затруднения. Сложнее обстоит дело со второй группой способов отбора.

Способы отбора называются *репрезентативными*, если отобранная группа элементов достаточно полно характеризует всю совокупность, из которой был сделан отбор. Разумеется, как бы ни был удачен отбор, в суждениях о всей совокупности будет элемент случайности. Более того, некоторые особенности всей совокупности вообще не отражатся на отобранной группе элементов, поэтому «репрезентативность» отбора, вообще говоря, является относительной и связана с конкретной числовой характеристикой совокупности, которая изучается с помощью отобранных элементов.

Фактически репрезентативный отбор применяется тогда, когда судить о характеристиках совокупности, используя все ее элементы, невозможно (либо слишком трудно) из-за того, что эта совокупность чересчур велика, а возможно, и не вся доступна анализу. Если объем N этой совокупности очень велик, то его практически можно считать бесконечным. В этом случае заданную совокупность можно рассматривать как генеральную, а отобранные элементы как выборку, применяя в дальнейшем все достижения общего выборочного метода. Если же N не очень велико по сравнению с количеством n отобранных элементов (скажем, $N < 10n$), то с числом N нужно считаться при интерпретации результатов отбора.

Количество отбираемых элементов можно определить, если известна дисперсия σ^2 изучаемой характеристики (по каким-либо предыдущим данным). Это количество будет зависеть от той точности, с которой мы хотим получить значение характеристики. Если эта точность задана допустимой дисперсией s^2 , то n определяется из формулы

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

которая при очень большом N перейдет в известную нам уже формулу

$$s^2 = \frac{\sigma^2}{n}.$$

Последнюю формулу легко объяснить: σ^2 есть как бы генеральная дисперсия, а s^2 есть дисперсия среднего выборки из n элементов, которая всегда в n раз меньше «одиначной» дисперсии.

Приведем пример. Допустим, что из предыдущих анализов известна дисперсия, связанная с неоднородностью вещества и равная $\sigma^2 = 0,20$. Допустим, что нам нужно знать содержание вещества с дисперсией s_1^2 , не превосходящей 0,40, причем на методику анализа приходится дисперсия $s_2^2 = 0,35$ (ошибка воспроизводимости). Это значит, что на долю неоднородности допускается оставить дисперсию $s^2 \leq \leq 0,05$. Если в нашем распоряжении $N = 50$ проб вещества,

то необходимое число проб n определится из равенства

$$0,05 = \frac{0,20}{n} \left(\frac{50-n}{50-1} \right),$$

т. е. равно пяти (точное решение уравнения равно $\frac{40}{9}$).

Выбор того или иного способа репрезентативного отбора зависит от степени наших знаний о всей совокупности. Если, например, нам известно, что элементы в совокупности расположены случайным образом, то можно применять *механический отбор* — отбирать каждый пятый или каждый десятый и т. п. элемент. Если же в последовательности элементов имеется некоторая ритмичность, то нужно применять *аритмичный отбор*, например, в первой десятке брать первый элемент, во второй десятке — второй и т. д. Скажем, анализируя раз в день качество продукции, мы не должны брать пробы в одно и то же время дня.

Иногда вся совокупность заведомо разбивается на отдельные части, которые желательно равномерно учесть. В этом случае применяют *типический отбор*, т. е. отдельно отбирают элементы из каждой части и лишь потом сводят их в общую группу. Например, анализируя производственную деятельность колхозов в области, мы должны вначале разбить эту область на зоны, сообразно природным условиям, и уже затем в каждой зоне отбирать для исследования необходимое число колхозов.

В тех случаях, когда о всей совокупности ничего не известно, единственной гарантией репрезентативности может служить *случайный отбор*. Для того чтобы отбор был случайным, нужно все элементы совокупности пронумеровать; номера отобранных элементов должны образовывать *случайную последовательность чисел*. Проверить случайность какой-нибудь последовательности чисел нетрудно хотя бы методом серий (см. п. 7.5). Но здесь речь идет о том, чтобы *создать* случайную последовательность, собственной волей имитировать случай.

Имитация случая является весьма трудной задачей. Как бы вы ни старались, выписывая числа, делать их случайными, в них обязательно проявится (особенно при больших объемах) какой-то бессознательно выбранный вами план. Можно провести следующий эксперимент. Предложите

кому-нибудь на выбор подбросить 100 раз монету, выписывая последовательность гербов и цифр в виде единиц и нулей, или же, не подбрасывая монету, самому произвольно выписать 100 единиц и нулей. С помощью метода серий (их здесь не должно быть меньше 40 и больше 61) вы почти всегда сможете узнать, как получена последовательность.

Итак, «подделать» случайность трудно. Тем не менее, в очень многих исследованиях случайный отбор или случайное перемешивание (рандомизация) данных настоятельно необходимы. Как же быть в таких ситуациях?

На помощь случаю может прийти... сам случай. Выпишите все номера элементов совокупности на отдельных билетах, положите их в какой-нибудь закрытый сосуд, перемешайте и тяните билетки — столько, сколько вам нужно. Случайность вытянутых номеров будет полностью обеспечена.

Можно воспользоваться и другим способом. Если взять, например, восьмизначные таблицы логарифмов (или синусов, тангенсов и т. п.), то последние три-четыре цифры в табличных числах будут абсолютно случайными. Берите подряд нужное количество двух-, трех- или четырехзначных чисел, составленных из этих последних цифр, и случайность выбора вам вновь будет гарантирована.

Конечно, найти такие таблицы не так-то легко. Поэтому для удобства пользования получаемые таким образом *случайные числа* выписывают в виде специальных таблиц (таблица XIII Приложения к настоящей книге).

Пользоваться таблицей случайных чисел очень просто. Допустим, вам нужно отобрать 10 элементов из совокупности, содержащей 100 элементов. Пронумеруйте все элементы от 00 до 99. Затем, начиная с любого места таблиц, выпишите две последние цифры десяти идущих подряд чисел. Получившиеся номера и покажут, какие элементы надо отобрать. Так, начиная с первого числа мы получим номера

82 49 18 48 09 50 17 10 37 51

(если числа будут повторяться, их надо опустить).

Выбранную последовательность случайных чисел в дальнейшем никоим образом нельзя изменять. Нужно помнить: именно случайность служит основой применения методов математической статистики, именно она обеспечивает репре-

зентативность отбора. Нарушение случайности, как правило, ведет к искажению результатов.

Аналогично отбору производится случайное перемешивание элементов. При этом нужно выписывать случайные номера до тех пор, пока они не охватят все заданные элементы.

10.2. Выбор числа наблюдений. Увеличение числа параллельных наблюдений n является основным способом повышения точности статистического анализа. Действительно, среднее выборки объема n имеет дисперсию в n раз меньше, чем одиночные наблюдения. Поэтому для определения необходимого числа наблюдений достаточно знать генеральную дисперсию наблюдений σ^2 и допустимую дисперсию результата s^2 ; при этом $n = \frac{\sigma^2}{s^2}$. Увеличивая число параллельных наблюдений, можно неограниченно повышать точность найденного результата, лишь бы только в процессе наблюдений не менялись условия испытания. Именно этим приемом пользуются, чтобы уменьшить доверительный интервал при неизменной доверительной вероятности.

Сложнее обстоит дело с выбором числа наблюдений при проверке гипотез. Это число должно одновременно служить двум целям: уменьшать вероятность α ошибки первого рода (отклонение правильной гипотезы) и вероятность β ошибки второго рода (принятие неверной гипотезы). Первая вероятность совпадает с принятым уровнем значимости p . Вторая вероятность зависит от многих факторов и в первую очередь от того, насколько неверна принимаемая гипотеза. Поэтому под β понимают обычно наибольшую возможную вероятность ошибки второго рода при любых отклонениях от правильной гипотезы.

Контроль за ошибками второго рода является трудной задачей, требуя специальных методов проверки гипотезы. Общий принцип такого контроля мы поясним на примере конкретной гипотезы.

Допустим, изучается нормально распределенная случайная величина ξ с дисперсией σ^2 . Высказывается гипотеза, что генеральное среднее этой величины $a \leq 0$ (односторонняя гипотеза). Требуется проверить эту гипотезу по одному наблюдению x . Если мы хотим, чтобы вероятность

ошибки первого рода была не больше α , мы должны считать гипотезу неверной только, если $x > u_{1-\alpha}\sigma$, где $u_{1-\alpha}$ — квантиль стандартного нормального распределения. Действительно, с доверительной вероятностью $1-\alpha$ справедлива односторонняя оценка

$$a \geq x - u_{1-\alpha}\sigma,$$

которая позволяет утверждать, что $a > 0$ лишь при $x > u_{1-\alpha}\sigma$.

Если окажется, что $x \leq u_{1-\alpha}\sigma$, то мы будем вынуждены рассматриваемую гипотезу признать справедливой, хотя, возможно, допустим при этом ошибку второго рода. Вероятность неравенства $x \leq u_{1-\alpha}\sigma$ зависит от того, каким на самом деле является a . Если обозначить через $F(x)$ функцию нормального распределения со средним a и дисперсией σ^2 , то

$$P\{x \leq u_{1-\alpha}\sigma\} = F(u_{1-\alpha}\sigma).$$

Используя функцию Лапласа (п. 3.2), получим, что

$$P\{x \leq u_{1-\alpha}\sigma\} = \frac{1}{2} + \Phi\left(\frac{u_{1-\alpha}\sigma - a}{\sigma}\right) = \frac{1}{2} + \Phi\left(u_{1-\alpha} - \frac{a}{\sigma}\right).$$

Если $a > 0$, то принимаемая гипотеза неверна. Однако при этом a может быть сколь угодно близким к нулю, а это значит, что $\Phi\left(u_{1-\alpha} - \frac{a}{\sigma}\right)$ будет сколь угодно близко к $\Phi(u_{1-\alpha}) = \frac{1}{2} - \alpha$. Иными словами, максимальная вероятность ошибки второго рода равна

$$\beta = \frac{1}{2} + \Phi(u_{1-\alpha}) = 1 - \alpha.$$

Например, при обычном уровне значимости $\alpha = 0,05$ будет $\beta = 0,95$, т. е. мы почти всегда ошибемся и примем неверную гипотезу.

Итак, желая застраховать себя от несправедливого отклонения верной гипотезы, мы впадаем в другую крайность, объявляя правильными почти сплошь неверные гипотезы. Точно такая же картина возникнет, если в качестве отправной точки взять фиксированную малую вероятность β . Для того чтобы вероятность принять неверную гипотезу не превышала β , мы должны будем отклонять гипотезу уже при

$x > -u_{1-\beta}\sigma$. А это значит, что большой станет вероятность α , по-прежнему равная $1-\beta$. Соотношение вероятностей α и β наглядно представлено на рис. 28, где эти вероятности выражаются площадями участков под нормальной кривой. Площадь, расположенная вправо от критического значения, равна α , расположенная влево равна β . Стараясь уменьшить α , мы сдвигаем это критическое значение вправо, стараясь уменьшить β — влево. Ясно, что с помощью единого критического значения невозможно даже сделать, чтобы обе эти вероятности были меньше $\frac{1}{2}$.

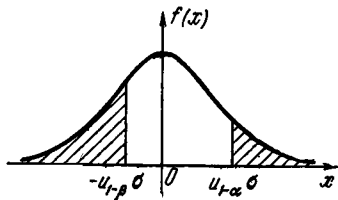


Рис. 28.

На практике обычно вероятности ошибок обоих родов задаются заранее и являются малыми числами (не больше 0,1). В этом случае при проверке гипотезы приходится пользоваться *одновременно* двумя критическими значениями: считать гипотезу верной, если $x < -u_{1-\beta}\sigma$, и неверной, если $x > u_{1-\alpha}\sigma$. Такой подход обеспечивает заданный уровень вероятностей ошибок первого и второго рода, однако при этом появляется *участок неопределенности* $[-u_{1-\beta}\sigma, u_{1-\alpha}\sigma]$. Длина этого участка

$$\delta = (u_{1-\beta} + u_{1-\alpha}) \sigma \quad (10.1)$$

называется *неопределенностью критерия*. Вопрос о том, какой вывод делать в случае, если x попадет на участок неопределенности, решается применительно к условиям задачи; чаще всего в этом случае гипотеза объявляется сомнительной и подвергается повторному анализу.

Неопределенность критерия сильно ухудшает эффект статистического анализа гипотез. Поэтому ее всячески пытаются, если не устранить вообще (мы уже видели, что это невозможно), то по крайней мере сделать как можно меньше. Непосредственно из формулы (10.1) видно, что это можно сделать и при неизменных заданных α и β , если уменьшать дисперсию σ^2 .

Уменьшать дисперсию можно, повышая точность методики. Но более универсальным и надежным средством является увеличение числа параллельных наблюдений.

Действительно, если вместо одиночного провести n параллельных наблюдений и рассмотреть их среднее \bar{x} , то его дисперсия $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$. Проверая гипотезу по значению \bar{x} , мы должны будем принять ее, если $\bar{x} < -u_{1-\beta} \frac{\sigma}{\sqrt{n}}$, и отвергнуть, если $\bar{x} > u_{1-\alpha} \frac{\sigma}{\sqrt{n}}$. Неопределенность станет теперь равной

$$\delta = (u_{1-\beta} + u_{1-\alpha}) \frac{\sigma}{\sqrt{n}}, \quad \text{т. е. уменьшится в } \sqrt{n} \text{ раз.}$$

В практических задачах допустимая неопределенность критерия задается обычно заранее, наряду с вероятностями α и β . Это позволяет сразу же определить необходимое число параллельных наблюдений:

$$n = (u_{1-\beta} + u_{1-\alpha})^2 \left(\frac{\sigma}{\delta} \right)^2.$$

Проиллюстрируем сказанное конкретным примером. Лаборатория химического завода проверяет выпускаемую продукцию, чтобы содержание вредной примеси не превышало 0,04%. Ошибка воспроизводимости одиночного анализа $\sigma = 0,002\%$. Требуется, чтобы вероятность пропустить бракованную продукцию не превышала $\beta = 0,02$. Вместе с тем анализ не должен браковать хорошую продукцию. Разумеется, если содержание примеси в продукции «на пределе», какие-либо гарантии дать трудно. Но мы можем, например, потребовать, чтобы вероятность забраковать продукцию с содержанием примеси ниже 0,036% (достаточно хорошая продукция) не превышала $\alpha = 0,05$. Нужно определить число параллельных анализов n , которое позволило бы удовлетворить всем поставленным требованиям.

Чтобы использовать проведенные ранее теоретические рассуждения, будем в качестве случайной величины ξ рассматривать отклонение результата анализа примеси от заданного предела 0,04. Генеральное среднее a величины ξ даст истинное значение этого отклонения. Предположение «продукция годна» совпадает с гипотезой $a \leq 0$. Вероятности α и β заданы в явном виде, допустимую неопределенность критерия легко найти:

$$\delta = 0,04 - 0,036 = 0,004.$$

Отсюда

$$n = (u_{0,95} + u_{0,98})^2 \left(\frac{\sigma}{\delta} \right)^2 = (1,64 + 2,05)^2 \left(\frac{0,002}{0,004} \right)^2 = 3,42,$$

т. е. для выполнения всех требований задачи нужно делать четыре параллельных анализа. Продукция будет при этом признана безоговорочно годной, если среднее отклонение

$$\bar{x} < -2,05 \frac{0,002}{\sqrt{4}} \approx -0,002,$$

т. е. если по результатам четырех анализов содержание примеси (в среднем) окажется меньше 0,038%.

Если результат анализа будет больше, чем 0,038%, то продукция является или сомнительной (в пределах участка неопределенности), или бракованной. Обычно участок неопределенности присоединяют к участку непригодной продукции, считая продукцию бракованной всегда, как только результат анализа окажется больше 0,038%. Связано это с тем, что завод при выпуске продукции никогда не работает «на пределе», поэтому вероятность продукции с содержанием примеси, близким к 0,04%, достаточно мала.

Описанная схема анализа предполагает, что число параллельных наблюдений n выбирается до опыта и в дальнейшем не меняется. Однако если бы первый же анализ продукции химзавода показал содержание примеси 0,02%, вряд ли стоило бы продолжать анализы. Иными словами, выбранное заранее n всегда достаточно для получения заданных вероятностей α и β , но отнюдь не всегда необходимо.

Излишних наблюдений можно избежать, поступая, например, следующим образом. Вначале делают одно наблюдение x_1 и сравнивают его с пределами $-u_{1-\beta}\sigma$ и $u_{1-\alpha}\sigma$. Может быть, это наблюдение будет меньше $-u_{1-\beta}\sigma$, тогда гипотезу $a \leq 0$ можно сразу же принимать. Если $x_1 > u_{1-\alpha}\sigma$, то гипотезу можно сразу отвергнуть. И только если x_1 попадает на участок неопределенности $[-u_{1-\beta}\sigma, u_{1-\alpha}\sigma]$, нужно проводить второе наблюдение.

Сделав второе наблюдение x_2 , находим среднее:

$$\bar{x}_2 = \frac{x_1 + x_2}{2}.$$

Это среднее сравниваем теперь с пределами $-u_{1-\beta} \frac{\sigma}{\sqrt{2}}$ и $u_{1-\alpha} \frac{\sigma}{\sqrt{2}}$. Гипотеза может быть снова отвергнута или принята, и тогда наблюдения кончаются. Если же \bar{x}_2 попадает на новый участок неопределенности $\left[-u_{1-\beta} \frac{\sigma}{\sqrt{2}}, u_{1-\alpha} \frac{\sigma}{\sqrt{2}} \right]$, то делают третье наблюдение x_3 , находят среднее

$$\bar{x}_3 = \frac{x_1 + x_2 + x_3}{3}$$

и начинают новую проверку по пределам $-u_{1-\beta} \frac{\sigma}{\sqrt{3}}$ и $u_{1-\alpha} \frac{\sigma}{\sqrt{3}}$. Процесс продолжается до тех пор, пока количество наблюдений не достигнет выбранного заранее по формуле числа n (в примере с химзаводом максимальное число анализов $n=4$). Ясно, что при таком способе проверки гипотезы число необходимых наблюдений будет в большинстве случаев значительно меньше первоначально выбранного n .

Разумеется, последовательное увеличение числа наблюдений неприемлемо там, где наблюдения проводятся заранее намеченными сериями и добавление новых наблюдений сопряжено с большими трудностями. Например, при изучении некоторого процесса фотоспособом число снимков нужно определять заранее, иначе для каждого нового снимка придется повторять весь процесс.

До сих пор во всех наших рассуждениях проверялась гипотеза $a \leq 0$. Гипотеза $a \geq 0$ проверяется аналогично, только теперь ее нужно принимать, если

$$\bar{x} > u_{1-\beta} \frac{\sigma}{\sqrt{n}}.$$

Обе эти гипотезы односторонние и проверяются с помощью односторонних критериев.

Если же проверяется двусторонняя гипотеза $a=0$ (ни больше, ни меньше), то для проверки нужно применить двусторонний критерий. Пусть, по-прежнему, вероятность отклонить верную гипотезу равна α , а вероятность принять гипотезу, в то время как на самом деле $|a| \geq \delta$, равна β .

Тогда число наблюдений определится формулой

$$n = \left(u_{1-\frac{\alpha}{2}} + u_{1-\beta} \right)^2 \left(\frac{\sigma}{\delta} \right)^2,$$

и гипотеза признается справедливой только, если среднее \bar{x} одновременно удовлетворяет двум неравенствам:

$$-u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

Мы рассмотрели вопрос о выборе числа параллельных наблюдений, когда дисперсия наблюдений σ^2 известна заранее. Если же σ^2 заранее неизвестна, то для оценок нужно применять критерий Стьюдента. Выбор числа наблюдений n при этом более сложен и в настоящей книге рассматриваться не будет.

10.3. Последовательный анализ. В предыдущем пункте приводились соображения, показывающие, что число наблюдений можно сократить, если по ходу анализа учитывать уже сделанные наблюдения. Обобщением этой идеи служит разработанный Вальдом метод *последовательного анализа*. При этом методе после каждого нового наблюдения решают, принять гипотезу, отклонить или продолжать испытания. Последовательный анализ позволяет сокращать число необходимых наблюдений в среднем в два раза по сравнению с обычными методами, фиксирующими число наблюдений заранее.

Последовательный анализ разросся сейчас в обширную теорию. Мы рассмотрим его применение только к анализу генерального среднего a наблюдаемой случайной величины. Допустим, что нам нужно сделать выбор между гипотезами $a \leq a_1$ и $a \geq a_2$ (предполагается, что $a_1 < a_2$). Вероятность принять гипотезу $a \geq a_2$, когда в действительности $a \leq a_1$, обозначим через α . Вероятность противоположной ошибки, т. е. принятия гипотезы $a \leq a_1$, когда в действительности $a \geq a_2$, обозначим через β . Числа α и β обычно малы и задаются заранее.

Основная идея последовательного анализа заключается в следующем. При каждой совокупности наблюдений x_1, x_2, \dots, x_n мы можем найти вероятность p'_n того, что эти

наблюдения получены из совокупности с генеральным средним a_1 , и вероятность p_n'' того, что они получены из совокупности с генеральным средним a_2 . Согласно принципу максимума правдоподобия (см. п. 7.3) на практике осуществляются события с максимальной вероятностью. Это значит, что при $p_n' > p_n''$ нужно считать более правдоподобным значение $a = a_1$ (а с ним и всю гипотезу $a \leq a_1$). Если же $p_n'' > p_n'$, то предпочтение нужно отдать второй гипотезе $a \geq a_2$.

Итак, все решается *отношением правдоподобия* $\frac{p_n''}{p_n'}$ — будет ли оно больше или меньше единицы. Ясно, однако, что в случае, когда отношение правдоподобия лишь немного отличается от единицы, предпочтение соответствующей гипотезе будет весьма сомнительным и лучше всего продолжить испытания. Точные показатели, насколько должно отношение правдоподобия отличаться от единицы, чтобы между гипотезами можно было сделать уверенный выбор, определяются заданными вероятностями α и β . Вальд показал, что гипотезу $a \leq a_1$ можно принять, если

$$\frac{p_n''}{p_n'} \leq \frac{\beta}{1-\alpha},$$

и гипотезу $a \geq a_2$ можно принять, если

$$\frac{p_n''}{p_n'} \geq \frac{1-\beta}{\alpha}.$$

Если же

$$\frac{\beta}{1-\alpha} < \frac{p_n''}{p_n'} < \frac{1-\beta}{\alpha},$$

то испытания надо продолжать.

При каждом новом наблюдении границы для отношения правдоподобия не меняются, меняется лишь само отношение. Это облегчает применение последовательного анализа, позволяет его свести к простым алгоритмам.

Если наблюдаемая случайная величина имеет нормальное распределение с заранее известной дисперсией σ^2 , то условие продолжения испытаний можно преобразовать к виду

$$A_1 + bn < \sum x_i < A_2 + bn,$$

где

$$A_1 = 2,3 \frac{\sigma^2}{a_2 - a_1} \lg \frac{\beta}{1 - \alpha}, \quad A_2 = 2,3 \frac{\sigma^2}{a_2 - a_1} \lg \frac{1 - \beta}{\alpha}, \quad b = \frac{a_1 + a_2}{2}.$$

Числа A_1 , A_2 и b подсчитываются сразу же по исходным данным, благодаря чему дальнейшая проверка ведется только по сумме проделанных наблюдений и их числу.

Последовательный анализ нормально распределенной случайной величины особенно удобно проводить геометрически. Для этого после каждого наблюдения строят точку на координатной плоскости,

откладывая по оси абсцисс число проделанных наблюдений n , а по оси ординат их сумму $\sum x_i$. Испытания нужно продолжать, пока эти точки будут находиться в полосе между прямыми $y = A_1 + bn$ и $y = A_2 + bn$ (рис. 29). Если же хоть одна точка окажется ниже этой полосы, анализ

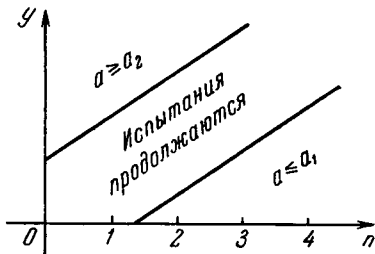


Рис. 29.

нужно прекращать и принимать гипотезу $a \leq a_1$. Точно так же гипотеза $a \geq a_2$ принимается, как только хоть одна точка окажется выше указанной полосы.

Рассмотрим следующий пример. На хлебоприемном пункте принимают зерно засоренностью не свыше 6%. Зерно поступает партиями, из которых берутся пробы на анализ. В связи с ошибками анализа и неоднородностью партий возникает средняя квадратичная ошибка результатов, известная по большому числу предыдущих анализов и равная $\sigma = 0,24\%$. Чтобы гарантировать доброкачественность, требуется браковать зерно, засоренное больше, чем на 6%, с вероятностью, не меньшей 0,95. Нежелательно впадать и в другую крайность — браковать хорошее зерно. Поэтому ставится еще одно требование — зерно, засоренное не больше, чем на 5,8%, принимать с вероятностью 0,90.

Покажем, как для контроля качества зерна применять последовательный анализ. В поставленных условиях $a_1 = 5,8$, $a_2 = 6,0$. Вероятности ошибок допускаются

$\alpha = 1 - 0,90 = 0,10$ и $\beta = 1 - 0,95 = 0,05$. Непосредственно вычисляем:

$$A_1 = 2,3 \frac{0,24^2}{0,20} \lg \frac{0,05}{0,90} = -8,31;$$

$$A_2 = 2,3 \frac{0,24^2}{0,20} \lg \frac{0,95}{0,10} = 13,10;$$

$$b = \frac{5,8 + 6,0}{2} = 5,9.$$

Область продолжения анализов будет ограничена прямыми

$$y = -8,31 + 5,9n, \quad y = 13,10 + 5,9n,$$

которые построены на рис. 30.

Анализ первой пробы из партии зерна дал результат 4,2%. Ему соответствует на рис. 30 точка M_1 с абсциссой $n=1$ и ординатой $y=4,2$. Точка M_1 находится в «полосе продолжения испытаний», поэтому берем вторую пробу. Ее

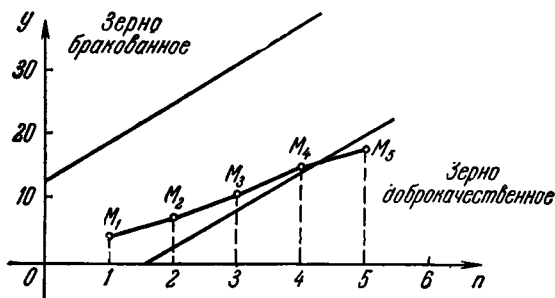


Рис. 30

анализ дал результат 3,9%, что соответствует точке M_2 с абсциссой $n=2$ и ординатой $y=4,2+3,9=8,1$. Испытания снова нужно продолжать. Третий анализ дал значение 3,6%; ему соответствует точка M_3 с абсциссой $n=3$ и ординатой $y=8,1+3,6=11,7$. Четвертому анализу с результатом 4,0% соответствует точка M_4 ($n=4$, $y=15,7$). Наконец, пятая точка M_5 , соответствующая еще одному анализу с результатом 4,4%, получает координаты $n=5$, $y=15,7+4,4=20,1$, и выходит за пределы «полосы продолжения испытаний» вниз. Партию зерна нужно считать доброкачественной.

Может показаться, что доброкачественность зерна видна уже из первого анализа. К сожалению, такой вывод поспешен, так как из-за высокого значения σ мы не можем гарантировать заданные вероятности ошибок α и β .

Для решения поставленной задачи с помощью метода последовательного анализа нам понадобилось пять наблюдений. Если же мы захотим определить число наблюдений n заранее по формулам предыдущего пункта, то мы получим

$$n = (u_{0,90} + u_{0,95})^2 \left(\frac{\sigma}{a_2 - a_1} \right)^2 = (1,28 + 1,65)^2 \left(\frac{0,24}{0,20} \right)^2 \approx 13,$$

т. е. классический метод потребовал бы 13 анализов! Преимущества последовательного анализа неоспоримы.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. Г. К р а м е р, Математические методы статистики, ИЛ, М. (1948): Полное и систематическое изложение математической статистики. Все утверждения строго доказываются, для используемых распределений выводятся все необходимые формулы. В полном объеме книга доступна лишь лицам, имеющим специальное математическое образование, однако написана она очень простым и ясным языком, благодаря чему отдельные разделы книги доступны даже начинающим. Особую ценность представляет четкая постановка основных задач математической статистики.
2. Б. В. Г н е д е н к о, Курс теории вероятностей, Гостехиздат, М. (1954).
Учебник по теории вероятностей для математических факультетов университетов. Содержит систематическое и ясное изложение основ этой теории, снабженное большим количеством примеров. В конце книги излагаются элементы статистики, главным образом, с вероятностной точки зрения. Отдельные разделы книги доступны начинающим.
3. Н. В. С м и р н о в и И. В. Д у н и н-Б а р к о в с к и й, Краткий курс математической статистики для технических приложений, Физматгиз, М. (1959).
Учебное пособие для студентов вузов; доступно лицам, владеющим математическим анализом в объеме программы вуза. Книга содержит большую теоретическую часть, а также изложение основных принципов математической статистики. Приведено много примеров, большое внимание уделено методике соответствующих расчетов. Особую ценность представляет глава, посвященная вероятностно-статистическим методам расчета, анализа и контроля точности производственных процессов.
4. А. К. М и т р о п о л ь с к и й, Техника статистических вычислений, Физматгиз, М. (1961).
Развернутое справочное пособие по математической обработке наблюдений. Изложение большей частью рецептурное, но на высоком математическом уровне. Большое внимание уделено применению таблиц; некоторые из таблиц, помещенные в книге, уникальны. Особую ценность представляют глава, посвященная исследованию различных типов распределений, а также не имеющая себе равных в отечественной литературе глава по отысканию корреляционных уравнений.
5. К. А. Б р а у н л и, Статистические исследования в производстве, ИЛ, М. (1949).

Небольшая и весьма доступная широкому кругу читателей книга. Изложение ведется исключительно в рецептурной форме. Тонкие и сложные вопросы математической статистики раскрываются в основном на удачно подобранных примерах, которых в книге очень много. Особую ценность составляют разделы, посвященные планированию эксперимента, а разделы, посвященные дисперсионному анализу, хотя и изложены очень нестрого, но по полноте почти не имеют себе равных. К недостаткам книги относятся догматизм изложения и отсутствие общих постановок задач, что несколько затрудняет дальнейшую самостоятельную работу читателя.

6. В. В. Н а л и м о в, Применение математической статистики при анализе веществ, Физматгиз, М. (1960).

Несмотря на узкую направленность, книга содержит подробное и доступное изложение общих основ математической статистики. Для тех же, кто работает в аналитических лабораториях, она просто незаменима как настольное пособие. Особую ценность представляют большие разделы, посвященные дисперсионному анализу и линейной регрессии.

7. Я. Б. Ш о р, Статистические методы анализа и контроля качества и надежности, «Советское радио», М. (1962).

Книга посвящена общим вопросам статистической обработки опытных данных, а также методам исследования качества и надежности продукции. Изложение весьма лаконично, что, однако, искупается его необычайной ясностью и обилием удачных примеров. Наряду с нормальным, в книге изучается большое количество других специальных распределений. Особую ценность представляет раздел, посвященный обработке опытных данных в случае нестабильных условий испытаний. Большое количество рецептов и таблиц делает книгу ценным справочным пособием.

8. Б. М. Щ и г о л е в, Математическая обработка наблюдений, Физматгиз, М. (1962).

Книга рассчитана в основном на представителей «точных наук» — физиков, астрономов. Поэтому большое место занимают в ней вопросы нестатистической обработки данных — точечное интерполирование, действие с приближенными числами. Подробно освещены также способы обработки неравноточных измерений и методика подбора и анализа эмпирических формул.

9. Я. Я н к о, Математико-статистические таблицы, Госстатиздат, М. (1961).

В книге содержатся почти все используемые в математической статистике таблицы, а также графики и номограммы, составленные с высокой степенью точности. Особую ценность представляет обширное введение, где в сжатой форме излагается большинство современных статистических методов обработки наблюдений. Изложение сопровождается большим количеством примеров, а также подробными списками литературы.

ПРИЛОЖЕНИЕ

Таблица I

Значения функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,23	0,0910	0,46	0,1772	0,69	0,2549
0,01	0,0040	0,24	0,0948	0,47	0,1808	0,70	0,2580
0,02	0,0080	0,25	0,0987	0,48	0,1844	0,71	0,2611
0,03	0,0120	0,26	0,1026	0,49	0,1879	0,72	0,2642
0,04	0,0160	0,27	0,1064	0,50	0,1915	0,73	0,2673
0,05	0,0199	0,28	0,1103	0,51	0,1950	0,74	0,2703
0,06	0,0239	0,29	0,1141	0,52	0,1985	0,75	0,2734
0,07	0,0279	0,30	0,1179	0,53	0,2019	0,76	0,2764
0,08	0,0319	0,31	0,1217	0,54	0,2054	0,77	0,2794
0,09	0,0359	0,32	0,1255	0,55	0,2088	0,78	0,2823
0,10	0,0398	0,33	0,1293	0,56	0,2123	0,79	0,2852
0,11	0,0438	0,34	0,1331	0,57	0,2157	0,80	0,2881
0,12	0,0478	0,35	0,1368	0,58	0,2190	0,81	0,2910
0,13	0,0517	0,36	0,1406	0,59	0,2224	0,82	0,2939
0,14	0,0557	0,37	0,1443	0,60	0,2257	0,83	0,2967
0,15	0,0596	0,38	0,1480	0,61	0,2291	0,84	0,2995
0,16	0,0636	0,39	0,1517	0,62	0,2324	0,85	0,3023
0,17	0,0675	0,40	0,1554	0,63	0,2357	0,86	0,3051
0,18	0,0714	0,41	0,1591	0,64	0,2389	0,87	0,3078
0,19	0,0753	0,42	0,1628	0,65	0,2422	0,88	0,3106
0,20	0,0793	0,43	0,1664	0,66	0,2454	0,89	0,3133
0,21	0,0832	0,44	0,1700	0,67	0,2486	0,90	0,3159
0,22	0,0871	0,45	0,1736	0,68	0,2517	0,91	0,3186

Таблица I (продолжение)

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,92	0,3212	1,34	0,4099	1,76	0,4608	2,34	0,4904
0,93	0,3238	1,35	0,4115	1,77	0,4616	2,36	0,4909
0,94	0,3264	1,36	0,4131	1,78	0,4625	2,38	0,4913
0,95	0,3289	1,37	0,4147	1,79	0,4633	2,40	0,4918
0,96	0,3315	1,38	0,4162	1,80	0,4641	2,42	0,4922
0,97	0,3340	1,39	0,4177	1,81	0,4649	2,44	0,4927
0,98	0,3365	1,40	0,4192	1,82	0,4656	2,46	0,4931
0,99	0,3389	1,41	0,4207	1,83	0,4664	2,48	0,4934
1,00	0,3413	1,42	0,4222	1,84	0,4671	2,50	0,4938
1,01	0,3438	1,43	0,4236	1,85	0,4678	2,52	0,4941
1,02	0,3461	1,44	0,4251	1,86	0,4686	2,54	0,4945
1,03	0,3485	1,45	0,4265	1,87	0,4693	2,56	0,4948
1,04	0,3508	1,46	0,4279	1,88	0,4699	2,58	0,4951
1,05	0,3531	1,47	0,4292	1,89	0,4706	2,60	0,4953
1,06	0,3554	1,48	0,4306	1,90	0,4713	2,62	0,4956
1,07	0,3577	1,49	0,4319	1,91	0,4719	2,64	0,4959
1,08	0,3599	1,50	0,4332	1,92	0,4726	2,66	0,4961
1,09	0,3621	1,51	0,4345	1,93	0,4732	2,68	0,4963
1,10	0,3643	1,52	0,4357	1,94	0,4738	2,70	0,4965
1,11	0,3665	1,53	0,4370	1,95	0,4744	2,72	0,4967
1,12	0,3686	1,54	0,4382	1,96	0,4750	2,74	0,4969
1,13	0,3708	1,55	0,4394	1,97	0,4756	2,76	0,4971
1,14	0,3729	1,56	0,4406	1,98	0,4761	2,78	0,4973
1,15	0,3749	1,57	0,4418	1,99	0,4767	2,80	0,4974
1,16	0,3770	1,58	0,4429	2,00	0,4772	2,82	0,4976
1,17	0,3790	1,59	0,4441	2,02	0,4783	2,84	0,4977
1,18	0,3810	1,60	0,4452	2,04	0,4793	2,86	0,4979
1,19	0,3830	1,61	0,4463	2,06	0,4803	2,88	0,4980
1,20	0,3849	1,62	0,4474	2,08	0,4812	2,90	0,4981
1,21	0,3869	1,63	0,4484	2,10	0,4821	2,92	0,4982
1,22	0,3883	1,64	0,4495	2,12	0,4830	2,94	0,4984
1,23	0,3907	1,65	0,4505	2,14	0,4838	2,96	0,49846
1,24	0,3925	1,66	0,4515	2,16	0,4846	2,98	0,49856
1,25	0,3944	1,67	0,4525	2,18	0,4854	3,00	0,49865
1,26	0,3962	1,68	0,4535	2,20	0,4861	3,20	0,49931
1,27	0,3980	1,69	0,4545	2,22	0,4868	3,40	0,49966
1,28	0,3997	1,70	0,4554	2,24	0,4875	3,60	0,49984
1,29	0,4015	1,71	0,4564	2,26	0,4881	3,80	0,499928
1,30	0,4032	1,72	0,4573	2,28	0,4887	4,00	0,499968
1,31	0,4049	1,73	0,4582	2,30	0,4893	5,00	0,499997
1,32	0,4066	1,74	0,4591	2,32	0,4898		
1,33	0,4082	1,75	0,4599				

Таблица II

Квантили нормального распределения u
 $1 - \frac{p}{2}$

p	$1 - \frac{p}{2}$	$u_{1 - \frac{p}{2}}$	p	$1 - \frac{p}{2}$	$u_{1 - \frac{p}{2}}$	p	$1 - \frac{p}{2}$	$u_{1 - \frac{p}{2}}$
0,80	0,60	0,25	0,15	0,925	1,44	0,01	0,995	2,58
0,50	0,75	0,67	0,10	0,95	1,64	0,005	0,9975	2,81
0,40	0,80	0,84	0,05	0,975	1,96	0,002	0,999	3,09
0,30	0,85	1,04	0,04	0,980	2,05	0,001	0,9995	3,29
0,25	0,875	1,15	0,02	0,990	2,33	0,0001	0,99995	3,89
0,20	0,90	1,28						

Таблица III

Квантили распределения Стьюдента t
 $1 - \frac{p}{2}$

Число степеней свободы, f	Уровни значимости p						
	0,20	0,10	0,05	0,02	0,01	0,005	0,001
1	3,08	6,31	12,71	31,82	63,66	127,32	636,62
2	1,89	2,92	4,30	6,97	9,93	14,09	31,60
3	1,64	2,35	3,18	4,54	5,84	7,45	12,94
4	1,53	2,13	2,78	3,75	4,60	5,60	8,61
5	1,48	2,02	2,57	3,37	4,03	4,77	6,86
6	1,44	1,94	2,45	3,14	3,71	4,32	5,96
7	1,42	1,90	2,37	3,00	3,50	4,03	5,41
8	1,40	1,86	2,31	2,90	3,36	3,83	5,04
9	1,38	1,83	2,26	2,82	3,25	3,69	4,78
10	1,37	1,81	2,23	2,76	3,17	3,58	4,59
11	1,36	1,80	2,20	2,72	3,11	3,50	4,44
12	1,36	1,78	2,18	2,68	3,06	3,43	4,32
13	1,35	1,77	2,16	2,65	3,01	3,37	4,22
14	1,34	1,76	2,15	2,62	2,98	3,33	4,14
15	1,34	1,75	2,13	2,60	2,95	3,29	4,07
16	1,34	1,75	2,12	2,58	2,92	3,25	4,02
17	1,33	1,74	2,11	2,57	2,90	3,22	3,97
18	1,33	1,73	2,10	2,55	2,88	3,20	3,92
19	1,33	1,73	2,09	2,54	2,86	3,17	3,88
20	1,33	1,73	2,09	2,53	2,85	3,15	3,85
21	1,32	1,72	2,08	2,52	2,83	3,14	3,82
22	1,32	1,72	2,07	2,51	2,82	3,12	3,79
23	1,32	1,71	2,07	2,50	2,81	3,10	3,77
24	1,32	1,71	2,06	2,49	2,80	3,09	3,75
25	1,32	1,71	2,06	2,48	2,79	3,08	3,73
26	1,32	1,71	2,06	2,48	2,78	3,07	3,71

Таблица III (продолжение)

Число степеней свободы f	Уровни значимости p						
	0,20	0,10	0,05	0,02	0,01	0,005	0,001
27	1,31	1,70	2,05	2,47	2,77	3,06	3,69
28	1,31	1,70	2,05	2,47	2,76	3,05	3,67
29	1,31	1,70	2,04	2,46	2,76	3,04	3,66
30	1,31	1,70	2,04	2,46	2,75	3,03	3,65
40	1,30	1,68	2,02	2,42	2,70	2,97	3,55
60	1,30	1,67	2,00	2,39	2,66	2,91	3,46
120	1,29	1,66	1,98	2,36	2,62	2,86	3,37
∞	1,28	1,64	1,96	2,33	2,58	2,81	3,29

Таблица IV*)

Квантили распределения Пирсона χ^2_{1-p}

Число степеней свободы f	Уровни значимости p							
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30
1	0,00016	0,0006	0,0039	0,016	0,064	0,148	0,455	1,07
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386	2,41
3	0,115	0,185	0,352	0,584	1,005	1,424	2,366	3,66
4	0,30	0,43	0,71	1,06	1,65	2,19	3,36	4,9
5	0,55	0,75	1,14	1,61	2,34	3,00	4,35	6,1
6	0,87	1,13	1,63	2,20	3,07	3,83	5,35	7,2
7	1,24	1,56	2,17	2,83	3,82	4,67	6,35	8,4
8	1,65	2,03	2,73	3,49	4,59	5,53	7,34	9,5
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,7
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,8
11	3,1	3,6	4,6	5,6	7,0	8,1	10,3	12,9
12	3,6	4,2	5,2	6,3	7,8	9,0	11,3	14,0
13	4,1	4,8	5,9	7,0	8,6	9,9	12,3	15,1
14	4,7	5,4	6,6	7,8	9,5	10,8	13,3	16,2
15	5,2	6,0	7,3	8,5	10,3	11,7	14,3	17,3
16	5,8	6,6	8,0	9,3	11,2	12,6	15,3	18,4
17	6,4	7,3	8,7	10,1	12,0	13,5	16,3	19,5
18	7,0	7,9	9,4	10,9	12,9	14,4	17,3	20,6
19	7,6	8,6	10,1	11,7	13,7	15,4	18,3	21,7
20	8,3	9,2	10,9	12,4	14,6	16,3	19,3	22,8
21	8,9	9,9	11,6	13,2	15,4	17,2	20,3	23,9
22	9,5	10,6	12,3	14,0	16,3	18,1	21,3	24,9

*) При $f > 30$ значения χ^2_{1-p} можно вычислять по приближенной формуле

$$\chi^2_{1-p} = \frac{1}{2} (\sqrt{2f-1} + u_{1-p})^2,$$

где u_{1-p} — квантили нормального распределения, приведенные в таблице II Приложения.

Таблица IV (продолжение)

Число степеней свободы f	Уровни значимости p							
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30
23	10,2	11,3	13,1	14,8	17,2	19,0	22,3	26,0
24	10,9	12,0	13,8	15,7	18,1	19,9	23,3	27,1
25	11,5	12,7	14,6	16,5	18,9	20,9	24,3	28,2
26	12,2	13,4	15,4	17,3	19,8	21,8	25,3	29,3
27	12,9	14,1	16,2	18,1	20,7	22,7	26,3	30,3
28	13,6	14,8	16,9	18,9	21,6	23,6	27,3	31,4
29	14,3	15,6	17,7	19,8	22,4	24,6	28,3	32,5
30	15,0	16,3	18,5	20,6	23,4	25,5	29,3	33,5

Число степеней свободы f	Уровни значимости p							
	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
1	1,64	2,7	3,8	5,4	6,6	7,9	9,5	10,8
2	3,22	4,6	6,0	7,8	9,2	10,6	12,4	13,8
3	4,64	6,3	7,8	9,8	11,3	12,8	14,8	16,3
4	6,0	7,8	9,5	11,7	13,3	14,9	16,9	18,5
5	7,3	9,2	11,1	13,4	15,1	16,3	18,9	20,5
6	8,6	10,6	12,6	15,0	16,8	18,6	20,7	22,5
7	9,8	12,0	14,1	16,6	18,5	20,3	22,6	24,3
8	11,0	13,4	15,5	18,2	20,1	21,9	24,3	26,1
9	12,2	14,7	16,9	19,7	21,7	23,6	26,1	27,9
10	13,4	16,0	18,3	21,2	23,2	25,2	27,7	29,6
11	14,6	17,3	19,7	22,6	24,7	26,8	29,4	31,3
12	15,8	18,5	21,0	24,1	26,2	28,3	31	32,9
13	17,0	19,8	22,4	25,5	27,7	29,8	32,5	34,5
14	18,2	21,1	23,7	26,9	29,1	31,3	34	36,1
15	19,3	22,3	25,0	28,3	30,6	32,8	35,5	37,7
16	20,5	23,5	26,3	29,6	32,0	34,3	37	39,2
17	21,6	24,8	27,6	31,0	33,4	35,7	38,5	40,8
18	22,8	26,0	28,9	32,3	34,8	37,2	40	42,3
19	23,9	27,2	30,1	33,7	36,2	38,6	41,5	43,8
20	25,0	28,4	31,4	35,0	37,6	40,0	43	45,3
21	26,2	29,6	32,7	36,3	38,9	41,4	44,5	46,8
22	27,3	30,8	33,9	37,7	40,3	42,8	46	48,3
23	28,4	32,0	35,2	39,0	41,6	44,2	47,5	49,7
24	29,6	33,2	36,4	40,3	43,0	45,6	48,5	51,2
25	30,7	34,4	37,7	41,6	44,3	46,9	50	52,6
26	31,8	35,6	38,9	42,9	45,6	48,3	51,5	54,1
27	32,9	36,7	40,1	44,1	47,0	49,6	53	55,5
28	34,0	37,9	41,3	45,4	48,3	51,0	54,5	56,9
29	35,1	39,1	42,6	46,7	49,6	52,3	56	58,3
30	36,3	40,3	43,8	48,0	50,9	53,7	57,5	59,7

Таблица V

Значения величины $v_p = \sqrt{\frac{f}{\chi^2_{1-p}}}$

Число степеней свободы. f	Уровни значимости p							
	0,995	0,99	0,98	0,95	0,90	0,80	0,70	0,30
1	160	79	41	16	7,9	3,9	2,6	0,97
2	14	10	7,1	4,4	3,1	2,1	1,7	0,91
3	6,5	5,1	4,0	2,9	2,3	1,7	1,4	0,91
4	4,4	3,7	3,1	2,4	1,9	1,6	1,4	0,90
5	3,5	3,0	2,6	2,1	1,8	1,5	1,3	0,91
6	3,0	2,6	2,3	1,9	1,6	1,4	1,3	0,91
7	2,7	2,4	2,1	1,8	1,6	1,4	1,2	0,91
8	2,4	2,2	2,0	1,7	1,5	1,3	1,2	0,92
9	2,3	2,1	1,9	1,6	1,5	1,3	1,2	0,92
10	2,2	2,0	1,8	1,6	1,4	1,3	1,2	0,92
11	2,1	1,9	1,7	1,5	1,4	1,3	1,2	0,92
12	2,0	1,8	1,7	1,5	1,4	1,2	1,2	0,92
13	1,9	1,8	1,6	1,5	1,4	1,2	1,1	0,93
14	1,9	1,7	1,6	1,5	1,3	1,2	1,1	0,93
15	1,8	1,7	1,6	1,4	1,3	1,2	1,1	0,93
16	1,8	1,7	1,6	1,4	1,3	1,2	1,1	0,93
17	1,7	1,6	1,5	1,4	1,3	1,2	1,1	0,93
18	1,7	1,6	1,5	1,4	1,3	1,2	1,1	0,94
19	1,7	1,6	1,5	1,4	1,3	1,2	1,1	0,94
20	1,6	1,6	1,5	1,4	1,3	1,2	1,1	0,94
21	1,6	1,5	1,5	1,3	1,3	1,2	1,1	0,94
22	1,6	1,5	1,4	1,3	1,3	1,2	1,1	0,94
23	1,6	1,5	1,4	1,3	1,2	1,2	1,1	0,94
24	1,6	1,5	1,4	1,3	1,2	1,2	1,1	0,94
25	1,5	1,5	1,4	1,3	1,2	1,1	1,1	0,94
26	1,5	1,5	1,4	1,3	1,2	1,1	1,1	0,94
27	1,5	1,4	1,4	1,3	1,2	1,1	1,1	0,94
28	1,5	1,4	1,4	1,3	1,2	1,1	1,1	0,94
29	1,5	1,4	1,4	1,3	1,2	1,1	1,1	0,94
30	1,5	1,4	1,4	1,3	1,2	1,1	1,1	0,94

Таблица V (продолжение)

Число степеней свободы, f	Уровни значимости p							
	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
1	0,78	0,61	0,51	0,43	0,39	0,36	0,32	0,30
2	0,79	0,66	0,58	0,51	0,47	0,43	0,40	0,38
3	0,80	0,69	0,62	0,55	0,51	0,48	0,45	0,43
4	0,82	0,72	0,64	0,59	0,56	0,52	0,49	0,46
5	0,83	0,74	0,67	0,61	0,58	0,55	0,52	0,49
6	0,84	0,75	0,69	0,63	0,60	0,57	0,54	0,52
7	0,84	0,76	0,71	0,65	0,62	0,59	0,56	0,54
8	0,85	0,77	0,72	0,66	0,63	0,60	0,57	0,55
9	0,86	0,78	0,73	0,68	0,64	0,62	0,59	0,57
10	0,86	0,79	0,74	0,69	0,66	0,63	0,60	0,58
11	0,87	0,80	0,75	0,70	0,67	0,64	0,61	0,59
12	0,87	0,81	0,76	0,71	0,68	0,65	0,62	0,60
13	0,87	0,81	0,76	0,71	0,68	0,66	0,63	0,61
14	0,88	0,81	0,77	0,72	0,69	0,67	0,64	0,62
15	0,88	0,82	0,77	0,73	0,70	0,68	0,65	0,63
16	0,88	0,82	0,78	0,74	0,71	0,68	0,66	0,64
17	0,89	0,83	0,79	0,74	0,71	0,69	0,66	0,65
18	0,89	0,83	0,79	0,75	0,72	0,69	0,67	0,65
19	0,89	0,84	0,80	0,75	0,72	0,70	0,68	0,66
20	0,89	0,84	0,80	0,76	0,73	0,71	0,68	0,66
21	0,89	0,84	0,80	0,76	0,74	0,71	0,69	0,67
22	0,90	0,84	0,81	0,76	0,74	0,72	0,69	0,67
23	0,90	0,85	0,81	0,77	0,74	0,72	0,69	0,68
24	0,90	0,85	0,81	0,77	0,75	0,73	0,70	0,68
25	0,90	0,85	0,81	0,78	0,75	0,73	0,71	0,69
26	0,91	0,85	0,82	0,78	0,76	0,73	0,71	0,69
27	0,91	0,86	0,82	0,78	0,76	0,74	0,71	0,70
28	0,91	0,86	0,82	0,79	0,76	0,74	0,72	0,70
29	0,91	0,86	0,82	0,79	0,76	0,75	0,72	0,71
30	0,91	0,86	0,83	0,79	0,77	0,75	0,72	0,71

Таблица VI

Некоторые параметры распределения размаха выборки

n	α_n	β_n	γ_n	n	α_n	β_n	γ_n
2	0,8862	0,7555	1,75	12	0,3069	0,2389	17,5
3	0,5908	0,5249	3,63	13	0,2998	0,2310	18,8
4	0,4857	0,4273	5,48	14	0,2935	0,2239	19,9
5	0,4299	0,3715	7,25	15	0,2880	0,2188	21,1
6	0,3946	0,3346	8,93	16	0,2831	0,2123	22,2
7	0,3698	0,3081	10,53	17	0,2787	0,2074	23,3
8	0,3512	0,2879	12,06	18	0,2747	0,2029	24,3
9	0,3367	0,2720	13,52	19	0,2711	0,1989	25,3
10	0,3249	0,2590	14,91	20	0,2677	0,1951	26,3
11	0,3152	0,2482	16,2				

Таблица VII

Квантили распределения Фишера F_{1-p}

$f_1 \backslash f_2$	Уровень значимости 0,20								
	1	2	3	4	5	6	12	24	∞
1	9,5	12,0	13,1	13,7	14,0	14,3	14,9	15,2	15,6
2	3,6	4,0	4,2	4,2	4,3	4,3	4,4	4,4	4,5
3	2,7	2,9	2,9	3,0	3,0	3,0	3,0	3,0	3,0
4	2,4	2,5	2,5	2,5	2,5	2,5	2,5	2,4	2,4
5	2,2	2,3	2,3	2,2	2,2	2,2	2,2	2,2	2,1
6	2,1	2,1	2,1	2,1	2,1	2,1	2,0	2,0	2,0
7	2,0	2,0	2,0	2,0	2,0	2,0	1,9	1,9	1,8
8	2,0	2,0	2,0	1,9	1,9	1,9	1,8	1,8	1,7
9	1,9	1,9	1,9	1,9	1,9	1,8	1,8	1,7	1,7
10	1,9	1,9	1,9	1,8	1,8	1,8	1,7	1,7	1,6
11	1,9	1,9	1,8	1,8	1,8	1,8	1,7	1,6	1,6
12	1,8	1,8	1,8	1,8	1,7	1,7	1,7	1,6	1,5
13	1,8	1,8	1,8	1,8	1,7	1,7	1,6	1,6	1,5
14	1,8	1,8	1,8	1,7	1,7	1,7	1,6	1,6	1,5
15	1,8	1,8	1,8	1,7	1,7	1,7	1,6	1,5	1,5
16	1,8	1,8	1,7	1,7	1,7	1,6	1,6	1,5	1,4
17	1,8	1,8	1,7	1,7	1,7	1,6	1,6	1,5	1,4
18	1,8	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4
19	1,8	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4
20	1,8	1,8	1,7	1,7	1,6	1,6	1,5	1,5	1,4
22	1,8	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,4
24	1,7	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,3

Таблица VII (продолжение)

$f_1 \backslash f_2$	Уровень значимости 0,20								
	1	2	3	4	5	6	12	24	∞
26	1,7	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,3
28	1,7	1,7	1,7	1,6	1,6	1,6	1,5	1,4	1,3
30	1,7	1,7	1,6	1,6	1,6	1,5	1,5	1,4	1,3
40	1,7	1,7	1,6	1,6	1,5	1,5	1,4	1,4	1,2
60	1,7	1,7	1,6	1,6	1,5	1,5	1,4	1,3	1,2
120	1,7	1,6	1,6	1,5	1,5	1,5	1,4	1,3	1,1
∞	1,6	1,6	1,6	1,5	1,5	1,4	1,3	1,2	1,0

$f_1 \backslash f_2$	Уровень значимости 0,05								
	1	2	3	4	5	6	12	24	∞
1	164,4	199,5	215,7	224,6	230,2	234,0	244,9	249,0	254,3
2	18,5	19,2	19,2	19,3	19,3	19,3	19,4	19,5	19,5
3	10,1	9,6	9,3	9,1	9,0	8,9	8,7	8,6	8,5
4	7,7	6,9	6,6	6,4	6,3	6,2	5,9	5,8	5,6
5	6,6	5,8	5,4	5,2	5,1	5,0	4,7	4,5	4,4
6	6,0	5,1	4,8	4,5	4,4	4,3	4,0	3,8	3,7
7	5,6	4,7	4,4	4,1	4,0	3,9	3,6	3,4	3,2
8	5,3	4,5	4,1	3,8	3,7	3,6	3,3	3,1	2,9
9	5,1	4,3	3,9	3,6	3,5	3,4	3,1	2,9	2,7
10	5,0	4,1	3,7	3,5	3,3	3,2	2,9	2,7	2,5
11	4,8	4,0	3,6	3,4	3,2	3,1	2,8	2,6	2,4
12	4,8	3,9	3,5	3,3	3,1	3,0	2,7	2,5	2,3
13	4,7	3,8	3,4	3,2	3,0	2,9	2,6	2,4	2,2
14	4,6	3,7	3,3	3,1	3,0	2,9	2,5	2,3	2,1
15	4,5	3,7	3,3	3,1	2,9	2,8	2,5	2,3	2,1
16	4,5	3,6	3,2	3,0	2,9	2,7	2,4	2,2	2,0
17	4,5	3,6	3,2	3,0	2,8	2,7	2,4	2,2	2,0
18	4,4	3,6	3,2	2,9	2,8	2,7	2,3	2,1	1,9
19	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,8
20	4,4	3,5	3,1	2,9	2,7	2,6	2,3	2,1	1,8
22	4,3	3,4	3,1	2,8	2,7	2,6	2,2	2,0	1,8
24	4,3	3,4	3,0	2,8	2,6	2,5	2,2	2,0	1,7
26	4,2	3,4	3,0	2,7	2,6	2,4	2,1	1,9	1,7
28	4,2	3,3	2,9	2,7	2,6	2,4	2,1	1,9	1,6
30	4,2	3,3	2,9	2,7	2,5	2,4	2,1	1,9	1,6
40	4,1	3,2	2,9	2,6	2,5	2,3	2,0	1,8	1,5
60	4,0	3,2	2,8	2,5	2,4	2,3	1,9	1,7	1,4
120	3,9	3,1	2,7	2,5	2,3	2,2	1,8	1,6	1,3
∞	3,8	3,0	2,6	2,4	2,2	2,1	1,8	1,5	1,0

Таблица VII (продолжение)

$f_1 \backslash f_2$	Уровень значимости 0,01									
	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98,5	99,0	99,2	99,3	99,3	99,4	99,3	99,4	99,5	99,5
3	34,1	30,8	29,5	28,7	28,2	27,9	27,5	27,1	26,6	26,1
4	21,2	18,0	16,7	16,0	15,5	15,2	14,8	14,4	13,9	13,5
5	16,3	13,3	12,1	11,4	11,0	10,7	10,3	9,9	9,5	9,0
6	13,7	10,9	9,8	9,2	8,8	8,5	8,1	7,7	7,3	6,9
7	12,3	9,6	8,5	7,9	7,5	7,2	6,8	6,5	6,1	5,7
8	11,3	8,7	7,6	7,0	6,6	6,4	6,0	5,7	5,3	4,9
9	10,6	8,0	7,0	6,4	6,1	5,8	5,5	5,1	4,7	4,3
10	10,0	7,6	6,6	6,0	5,6	5,4	5,1	4,7	4,3	3,9
11	9,7	7,2	6,2	5,7	5,3	5,1	4,7	4,4	4,0	3,6
12	9,3	6,9	6,0	5,4	5,1	4,8	4,5	4,2	3,8	3,4
13	9,1	6,7	5,7	5,2	4,9	4,6	4,3	4,0	3,6	3,2
14	8,9	6,5	5,6	5,0	4,7	4,5	4,1	3,8	3,4	3,0
15	8,7	6,4	5,4	4,9	4,6	4,3	4,0	3,7	3,3	2,9
16	8,5	6,2	5,3	4,8	4,4	4,2	3,9	3,6	3,2	2,8
17	8,4	6,1	5,2	4,7	4,3	4,1	3,8	3,5	3,1	2,7
18	8,3	6,0	5,1	4,6	4,3	4,0	3,7	3,4	3,0	2,6
19	8,2	5,9	5,0	4,5	4,2	3,9	3,6	3,3	2,9	2,4
20	8,1	5,9	4,9	4,4	4,1	3,9	3,6	3,2	2,9	2,4
22	7,9	5,7	4,8	4,3	4,0	3,8	3,5	3,1	2,8	2,3
24	7,8	5,6	4,7	4,2	3,9	3,7	3,3	3,0	2,7	2,2
26	7,7	5,5	4,6	4,1	3,8	3,6	3,3	3,0	2,6	2,1
28	7,6	5,5	4,6	4,1	3,8	3,5	3,2	2,9	2,5	2,1
30	7,6	5,4	4,5	4,0	3,7	3,5	3,2	2,8	2,5	2,0
40	7,3	5,2	4,3	3,8	3,5	3,3	3,0	2,7	2,3	1,8
60	7,1	5,0	4,1	3,7	3,3	3,1	2,8	2,5	2,1	1,6
120	6,9	4,8	4,0	3,5	3,2	3,0	2,7	2,3	2,0	1,4
∞	6,6	4,6	3,8	3,3	3,0	2,8	2,5	2,2	1,8	1,0

Таблица VII (продолжение)

$f_1 \backslash f_2$	Уровень значимости 0,001									
	1	2	3	4	5	6	8	12	24	∞
1	Изменяется от 400 000 до 600 000									
2	998	999	999	999	999	999	999	999	999	999
3	167	148	141	137	135	133	131	128	126	123
4	74,1	61,3	56,2	53,4	51,7	50,5	49,0	47,4	45,8	44,1
5	47,0	36,6	33,2	31,1	29,8	28,8	27,6	26,4	25,1	23,8
6	35,5	27,0	23,7	21,9	20,8	20,0	19,0	18,0	16,9	15,8
7	29,2	21,7	18,8	17,2	16,2	15,5	14,6	13,7	12,7	11,7
8	25,4	18,5	15,8	14,4	13,5	12,9	12,0	11,2	10,3	9,3
9	22,9	16,4	13,9	12,6	11,7	11,1	10,4	9,6	8,7	7,8
10	21,0	14,9	12,6	11,3	10,5	9,9	9,2	8,5	7,6	6,8
11	19,7	13,8	11,6	10,4	9,6	9,1	8,3	7,6	6,9	6,0
12	18,6	13,0	10,8	9,6	8,9	8,4	7,7	7,0	6,3	5,4
13	17,8	12,3	10,2	9,1	8,4	7,9	7,2	6,5	5,8	5,0
14	17,1	11,8	9,7	8,6	7,9	7,4	6,8	6,1	5,4	4,6
15	16,6	11,3	9,3	8,3	7,6	7,1	6,5	5,8	5,1	4,3
16	16,1	11,0	9,0	7,9	7,3	6,8	6,2	5,6	4,9	4,1
17	15,7	10,7	8,7	7,7	7,0	6,6	6,0	5,3	4,6	3,9
18	15,4	10,4	8,5	7,5	6,8	6,4	5,8	5,1	4,5	3,7
19	15,1	10,2	8,3	7,3	6,6	6,2	5,6	5,0	4,3	3,5
20	14,8	10,0	8,1	7,1	6,5	6,0	5,4	4,8	4,2	3,4
22	14,4	9,6	7,8	6,8	6,2	5,8	5,2	4,6	3,9	3,2
24	14,0	9,3	7,6	6,6	6,0	5,6	5,0	4,4	3,7	3,0
26	13,7	9,1	7,4	6,4	5,8	5,4	4,8	4,2	3,6	2,8
28	13,5	8,9	7,2	6,3	5,7	5,2	4,7	4,1	3,5	2,7
30	13,3	8,8	7,1	6,1	5,5	5,1	4,6	4,0	3,4	2,6
40	12,6	8,2	6,6	5,7	5,1	4,7	4,2	3,6	3,0	2,2
60	12,0	7,8	6,2	5,3	4,8	4,4	3,9	3,3	2,7	1,9
120	11,4	7,3	5,8	5,0	4,4	4,0	3,5	3,0	2,4	1,6
∞	10,8	6,9	5,4	4,6	4,1	3,7	3,3	2,7	2,1	1,0

Таблица VIII*)

Квантили распределения Кохрана g_{1-p}

Уровень значимости 0,05															
$k \backslash f$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞	
2	9985	9750	9392	9057	8772	8534	8332	8159	8010	7880	7341	6602	5813	5000	
3	9669	8709	7977	7457	7071	6771	6530	6333	6167	6025	5466	4748	4031	3333	
4	9065	7679	6841	6287	5895	5598	5365	5175	5017	4884	4366	3720	3093	2500	
5	8412	6838	5981	5441	5065	4783	4564	4387	4241	4118	3645	3066	2513	2000	
6	7808	6161	5321	4803	4447	4184	3980	3817	3682	3568	3135	2612	2119	1667	
7	7271	5612	4800	4307	3974	3726	3535	3384	3259	3154	2756	2278	1833	1429	
8	6798	5157	4377	3910	3595	3362	3185	3043	2926	2829	2462	2022	1616	1250	
9	6385	4775	4027	3584	3286	3067	2901	2768	2659	2568	2226	1820	1446	1111	
10	6020	4450	3733	3311	3029	2823	2666	2541	2439	2353	2032	1655	1308	1000	
12	5410	3924	3264	2880	2624	2439	2299	2187	2098	2020	1737	1403	1100	8833	
15	4709	3346	2758	2419	2195	2034	1911	1815	1736	1671	1429	1144	889	6667	
20	3894	2705	2205	1921	1735	1602	1501	1422	1357	1303	1108	879	675	5000	
24	3434	2354	1907	1656	1493	1374	1286	1216	1160	1113	942	743	567	4417	
30	2929	1980	1593	1377	1237	1137	1061	1002	958	921	771	604	457	3333	
40	2370	1576	1259	1082	968	887	827	780	745	713	595	462	347	2250	
60	1737	1131	895	765	682	623	583	552	520	497	411	316	234	167	
120	0998	0632	0495	0419	0371	0337	0312	0292	0279	0266	0218	0165	0120	0083	
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	

*) Все квантили g_{1-p} меньше единицы, поэтому в таблице приведены лишь десятичные знаки, следующие после запятой, перед которой при пользовании таблицей нужно ставить нуль целых. Например, при $k=6$, $f=3$ имеем $g_{0.95}=0.5321$.

[illegible]

Таблица IX

Квантили распределения, максимального относительного отклонения τ_{1-p}

n	Уровни значимости p				n	Уровни значимости p			
	0,10	0,05	0,025	0,01		0,10	0,05	0,025	0,01
3	1,41	1,41	1,41	1,41	15	2,33	2,49	2,64	2,80
4	1,65	1,69	1,71	1,72	16	2,35	2,52	2,67	2,84
5	1,79	1,87	1,92	1,96	17	2,38	2,55	2,70	2,87
6	1,89	2,00	2,07	2,13	18	2,40	2,58	2,73	2,90
7	1,97	2,09	2,18	2,27	19	2,43	2,60	2,75	2,93
8	2,04	2,17	2,27	2,37	20	2,45	2,62	2,78	2,96
9	2,10	2,24	2,35	2,46	21	2,47	2,64	2,80	2,98
10	2,15	2,29	2,41	2,54	22	2,49	2,66	2,82	3,01
11	2,19	2,34	2,47	2,61	23	2,50	2,68	2,84	3,03
12	2,23	2,39	2,52	2,66	24	2,52	2,70	2,86	3,05
13	2,26	2,43	2,56	2,71	25	2,54	2,72	2,88	3,07
14	2,30	2,46	2,60	2,76					

Таблица X

Квантили распределения Колмогорова λ_{1-p}

p	λ_{1-p}	p	λ_{1-p}	p	λ_{1-p}
0,99	0,44	0,50	0,83	0,15	1,14
0,90	0,57	0,40	0,89	0,10	1,22
0,80	0,64	0,30	0,97	0,05	1,36
0,70	0,71	0,25	1,02	0,02	1,52
0,60	0,77	0,20	1,07	0,01	1,63

Таблица XI

Значения плотности стандартного нормального распределения

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \text{ и ее производных}$$

u	$\varphi(u)$	$\varphi'(u)$	$\varphi(u)$	$\varphi''(u)$	$\varphi^{IV}(u)$
0,0	0,3989	-0,0000	-0,3989	+0,0000	+1,1968
0,1	3970	0397	3930	1187	1671
0,2	3910	0782	3754	2315	1,0799
0,3	3814	1144	3471	3330	0,9413
0,4	3683	1473	3094	4184	7607
0,5	3521	1760	2641	4841	5501
0,6	3332	1999	2133	5278	3231
0,7	3123	2186	1593	5486	+0,0937
0,8	2897	2318	1043	5469	-0,1247
0,9	2661	2395	0506	5245	3203
1,0	2420	2420	-0,0000	4839	4839
1,1	2179	2396	+0,0458	4290	6091
1,2	1942	2330	0854	3635	6926
1,3	1714	2228	1182	2918	7341
1,4	1497	2096	1437	2180	7364
1,5	1295	1943	1619	1457	7043
1,6	1109	1775	1730	0781	6441
1,7	0941	1599	1778	+0,0176	5632
1,8	0790	1421	1769	-0,0341	4692
1,9	0656	1247	1713	0761	3693
2,0	0540	1080	1620	1080	2700
2,1	0440	0924	1500	1302	1765
2,2	0355	0780	1362	1436	0927
2,3	0283	0652	1215	1492	-0,0214
2,4	0224	0538	1066	1483	+0,0362
2,5	0175	0438	0920	1424	0800
2,6	0136	0353	0782	1328	1105
2,7	0104	0281	0656	1207	1293
2,8	0079	0222	0541	1073	1379
2,9	0060	0173	0441	0934	1385
3,0	0044	0133	0355	0793	1330
3,1	0033	0101	0281	0669	1231
3,2	0024	0076	0220	0552	1107
3,3	0017	0057	0170	0449	0969
3,4	0012	0042	0130	0359	0829
3,5	0009	0031	0098	0283	0694
3,6	0006	0022	0073	0219	0570
3,7	0004	0016	0054	0168	0460
3,8	0003	0011	0039	0127	0365
3,9	0002	0008	0028	0095	0284
4,0	0,0001	-0,0005	+0,0020	-0,0070	+0,0218

Таблица XII

Квантили распределения выборочного коэффициента
корреляции $r_{1-\frac{p}{2}}$

Число степеней свободы f	Уровни значимости p				
	0,10	0,05	0,02	0,01	0,001
1	0,988	0,997	0,999	1,000	1,000
2	0,900	0,950	0,980	0,990	0,999
3	0,805	0,878	0,934	0,959	0,992
4	0,729	0,811	0,882	0,917	0,974
5	0,669	0,754	0,833	0,874	0,951
6	0,621	0,707	0,789	0,834	0,925
7	0,582	0,666	0,750	0,798	0,898
8	0,549	0,632	0,716	0,765	0,872
9	0,521	0,602	0,685	0,735	0,847
10	0,497	0,576	0,658	0,708	0,823
11	0,476	0,553	0,634	0,684	0,801
12	0,457	0,532	0,612	0,661	0,780
13	0,441	0,514	0,592	0,641	0,760
14	0,426	0,497	0,574	0,623	0,742
15	0,412	0,482	0,558	0,606	0,725
16	0,400	0,468	0,543	0,590	0,708
17	0,389	0,456	0,528	0,575	0,693
18	0,378	0,444	0,516	0,561	0,679
19	0,369	0,433	0,503	0,549	0,665
20	0,360	0,423	0,492	0,537	0,652
25	0,323	0,381	0,445	0,487	0,597
30	0,296	0,349	0,409	0,449	0,554
35	0,275	0,325	0,381	0,418	0,519
40	0,257	0,304	0,358	0,393	0,490
45	0,243	0,287	0,338	0,372	0,465
50	0,231	0,273	0,322	0,354	0,443
60	0,211	0,250	0,295	0,325	0,408
70	0,195	0,232	0,274	0,302	0,380
80	0,183	0,217	0,256	0,283	0,357
90	0,173	0,205	0,242	0,267	0,337
100	0,164	0,195	0,230	0,254	0,321

Таблица XIII

Случайные числа

2182	1666	7373	4982	2368	2613	9025	2836	8493	8207
1549	8441	3351	3079	0026	4161	6224	4184	2633	2736
5118	4796	7035	2010	3449	7061	3847	9508	4528	1226

Таблица XIII (продолжение)

6848	3420	6583	7520	4809	1575	3209	9070	0644	3614
3309	4853	4021	8644	3980	5318	1959	4783	1810	2020
2050	3603	1812	4020	6573	0312	7739	9374	4788	4350
6817	6736	4591	9037	2949	7406	4238	4279	6206	1699
8310	3044	0433	1322	7664	3310	2487	3926	2233	8260
7050	7670	1848	5173	2146	1289	8504	0911	2001	5804
5637	5325	9367	5939	3191	5930	3361	6743	5995	4194
8251	3537	5139	5050	1516	8792	5513	5583	6103	9872
6129	7391	0429	2836	5284	5485	2643	4035	3089	4991
1311	4847	2317	3561	4533	6655	7354	1903	6909	9776
5151	2477	3250	7859	1764	1590	3309	2555	9334	3869
4245	4289	9301	4788	3937	9122	9301	4741	0683	0253
3395	7157	1084	0561	8388	7575	7513	0431	9659	8204
6542	9575	7896	7029	4986	6632	8145	7080	3287	3336
8068	1543	2019	3678	9248	2452	1224	0260	4158	3735
3370	3763	4713	4726	7010	3736	2412	9066	8663	5408
9306	0691	3221	3010	5948	5659	4145	2452	3340	5540
2166	2583	6878	3080	6318	6494	8057	7173	5753	3592
6844	3518	1638	7438	6106	2268	6361	7008	7995	8010
1479	7233	3827	9134	1023	6356	5108	3033	3546	0827
6342	6329	1364	1015	3892	5611	7390	2421	9873	9796
0485	7584	6030	2040	8568	3539	6208	9450	7613	4591
3603	9141	3778	1939	2596	5841	0528	4948	3849	4894
6736	5786	4356	9390	7245	7645	6476	1022	3393	5062
3044	9355	5064	6508	6771	2173	0511	7104	6753	6486
7670	6721	2387	4220	9357	2070	8506	9126	4729	1798
5325	7972	3664	8187	2040	3668	4930	9203	0248	8489
3537	8129	6485	3472	3915	2040	4353	8415	4506	3644
7391	3452	6919	3520	1417	7784	0734	7174	2648	6464
2066	3343	8506	9336	8366	9748	9925	4961	6179	0922
5120	2905	3163	6449	2833	2971	3390	4915	8231	6403
4058	0542	7260	9652	0310	5901	5420	9443	8654	5492
0079	8768	8409	2949	4751	5945	6473	7475	9696	2861
5609	9172	4284	2793	1333	2760	8276	6593	2346	7244
7070	3543	3812	8478	3696	8015	4970	1574	1636	6966
4748	0896	3758	6368	9968	6245	3108	9344	4513	0218
1093	8312	0310	1512	0178	6667	2966	0873	0582	7541
9818	4719	8187	6589	8807	2195	1160	4756	6503	1341
5690	7934	8875	9796	0520	7012	6355	5557	2070	5013
0943	8798	6149	5385	2884	1816	5453	8893	3695	4625
2166	5084	9117	0199	7127	4378	2940	2861	4690	5711
6702	2902	6365	7014	6121	5108	7369	7804	6732	3310
3208	4252	1361	8838	6770	9128	7183	8966	8292	9768
9692	1021	2415	6337	6060	5803	8205	2398	5607	0046
1211	1918	0260	7193	7603	8314	2891	9541	2529	6572
3230	7073	1890	4899	6533	4839	2512	0938	0233	7302
5969	3070	6557	9925	1969	4212	3052	4238	4210	6860

Таблица XIV

Значения некоторых функций, часто встречающихся
в статистических вычислениях

x	x^2	x^3	x^4	\sqrt{x}	$\sqrt{10x}$	$\frac{1}{x}$	$\lg x$
1,0	1,00	1,000	1,000	1,000	3,162	1,0000	0,0000
1,1	1,21	1,331	1,464	1,049	3,317	0,9091	0414
1,2	1,44	1,728	2,074	1,095	3,464	8333	0792
1,3	1,69	2,197	2,856	1,140	3,606	7692	1139
1,4	1,96	2,744	3,842	1,183	3,742	7143	1461
1,5	2,25	3,375	5,063	1,225	3,873	6667	1761
1,6	2,56	4,096	6,554	1,265	4,000	6250	2041
1,7	2,89	4,913	8,352	1,304	4,123	5882	2304
1,8	3,24	5,832	10,50	1,342	4,243	5556	2553
1,9	3,61	6,859	13,03	1,378	4,359	5263	2788
2,0	4,00	8,000	16,00	1,414	4,472	5000	3010
2,1	4,41	9,261	19,45	1,449	4,583	4762	3222
2,2	4,84	10,65	23,43	1,483	4,690	4546	3424
2,3	5,29	12,17	27,98	1,517	4,796	4348	3617
2,4	5,76	13,82	33,18	1,549	4,899	4167	3802
2,5	6,25	15,63	39,06	1,581	5,000	4000	3979
2,6	6,76	17,58	45,70	1,612	5,099	3846	4150
2,7	7,29	19,68	53,14	1,643	5,196	3704	4314
2,8	7,84	21,95	61,47	1,673	5,292	3571	4472
2,9	8,41	24,39	70,73	1,703	5,385	3448	4624
3,0	9,00	27,00	81,00	1,732	5,477	3333	4771
3,1	9,61	29,79	92,35	1,761	5,568	3226	4914
3,2	10,24	32,77	104,9	1,789	5,657	3125	5051
3,3	10,89	35,94	118,6	1,817	5,745	3030	5185
3,4	11,56	39,30	133,6	1,844	5,831	2941	5315
3,5	12,25	42,88	150,1	1,871	5,916	2857	5441
3,6	12,96	46,66	168,0	1,897	6,000	2778	5563
3,7	13,69	50,65	187,4	1,924	6,083	2703	5682
3,8	14,44	54,87	208,5	1,949	6,164	2632	5798
3,9	15,21	59,32	231,3	1,975	6,245	2564	5911
4,0	16,00	64,00	256,0	2,000	6,325	2500	6021
4,1	16,81	68,92	282,6	2,025	6,403	2439	6128
4,2	17,64	74,09	311,2	2,049	6,481	2381	6232
4,3	18,49	79,51	341,9	2,074	6,557	2326	6335
4,4	19,36	85,18	374,8	2,098	6,633	2273	6435
4,5	20,25	91,13	410,1	2,121	6,708	2222	6532
4,6	21,16	97,34	447,7	2,145	6,782	2174	6628
4,7	22,09	103,8	488,0	2,168	6,856	2128	6721
4,8	23,04	110,6	530,8	2,191	6,928	2083	6812
4,9	24,01	117,6	576,5	2,214	7,000	2041	6902
5,0	25,00	125,0	625,0	2,236	7,071	2000	6990
5,1	26,01	132,7	676,5	2,258	7,141	1961	7076
5,2	27,04	140,6	731,2	2,280	7,211	0,1923	7160

Таблица XIV (продолжение)

x	x^2	x^3	x^4	\sqrt{x}	$\sqrt{10x}$	$\frac{1}{x}$	$\lg x$
5,3	28,09	148,9	789,0	2,302	7,280	0,1887	0,7243
5,4	29,16	157,5	850,3	2,324	7,348	1852	7324
5,5	30,25	166,4	915,1	2,345	7,416	1818	7404
5,6	31,36	175,6	983,4	2,366	7,483	1786	7482
5,7	32,49	185,2	1056	2,387	7,550	1754	7559
5,8	33,64	195,1	1132	2,408	7,616	1724	7634
5,9	34,81	205,4	1212	2,429	7,681	1695	7709
6,0	36,00	216,0	1296	2,449	7,746	1667	7782
6,1	37,21	227,0	1385	2,470	7,810	1639	7853
6,2	38,44	238,3	1478	2,490	7,874	1613	7924
6,3	39,69	250,0	1575	2,510	7,937	1587	7993
6,4	40,96	262,1	1678	2,530	8,000	1563	8062
6,5	42,25	274,6	1785	2,550	8,062	1539	8129
6,6	43,56	287,5	1897	2,569	8,124	1515	8195
6,7	44,89	300,8	2015	2,588	8,185	1493	8261
6,8	46,24	314,4	2138	2,608	8,246	1471	8325
6,9	47,61	328,5	2267	2,627	8,307	1449	8388
7,0	49,00	343,0	2401	2,646	8,367	1429	8451
7,1	50,41	357,9	2541	2,665	8,426	1409	8513
7,2	51,84	373,2	2687	2,683	8,485	1389	8573
7,3	53,29	389,0	2840	2,702	8,544	1370	8633
7,4	54,76	405,2	2999	2,720	8,602	1351	8692
7,5	56,25	421,9	3164	2,739	8,660	1333	8751
7,6	57,76	439,0	3336	2,757	8,718	1316	8808
7,7	59,29	456,5	3515	2,775	8,775	1299	8865
7,8	60,84	474,6	3702	2,793	8,832	1282	8921
7,9	62,41	493,0	3895	2,811	8,888	1266	8976
8,0	64,00	512,0	4096	2,828	8,944	1250	9031
8,1	65,61	531,4	4305	2,846	9,000	1235	9085
8,2	67,24	551,4	4521	2,864	9,055	1220	9138
8,3	68,89	571,8	4746	2,881	9,110	1205	9191
8,4	70,56	592,7	4979	2,898	9,165	1191	9243
8,5	72,25	614,1	5220	2,915	9,220	1177	9294
8,6	73,96	636,1	5470	2,933	9,274	1163	9345
8,7	75,69	658,5	5729	2,950	9,327	1149	9395
8,8	77,44	681,5	5997	2,966	9,381	1136	9445
8,9	79,21	705,0	6274	2,983	9,434	1124	9494
9,0	81,00	729,0	6561	3,000	9,487	1111	9542
9,1	82,81	753,6	6857	3,017	9,539	1099	9590
9,2	84,64	778,7	7164	3,033	9,592	1087	9638
9,3	86,49	804,4	7481	3,050	9,644	1075	9685
9,4	88,36	830,6	7807	3,066	9,695	1064	9731
9,5	90,25	857,4	8145	3,082	9,747	1053	9777
9,6	92,16	884,7	8493	3,098	9,798	1042	9823
9,7	94,09	912,7	8853	3,114	9,849	1031	9868
9,8	96,04	941,2	9224	3,131	9,899	1020	9912
9,9	98,01	970,3	9606	3,146	9,950	0,1010	0,9956

