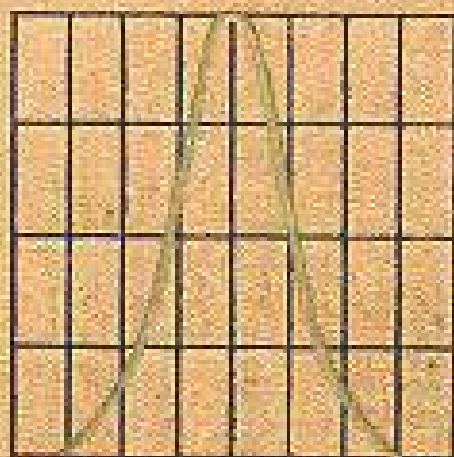


*М. Дж. Кендалл
А. Стьюарт*

ТЕОРИЯ РАСПРЕДЕЛЕНИЙ



THE ADVANCED THEORY OF STATISTICS

MAURICE G. KENDALL, M.A., Sc.D.

*Formerly Professor of Statistics in the University of London
President of the Royal Statistical Society, 1960-2*

and

ALAN STUART, B.Sc. (ECON.)

Reader in Statistics in the University of London

IN THREE VOLUMES

VOLUME 1

DISTRIBUTION THEORY

SECOND EDITION



CHARLES GRIFFIN & COMPANY LIMITED
LONDON

М. КЕНДАЛЛ, А. СТЫЮАРТ

ТЕОРИЯ РАСПРЕДЕЛЕНИЙ

Перевод с английского
В. В. САЗОНОВА, А. Н. ШИРЯЕВА

Под редакцией
А. Н. КОЛМОГОРОВА



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1966

517.8

К 35

УДК 519.240

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ	9
ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ	11
ВВОДНЫЕ ЗАМЕЧАНИЯ	12

Г Л А В Ы

1. ЧАСТОТНЫЕ РАСПРЕДЕЛЕНИЯ	15
2. МЕРЫ РАСПОЛОЖЕНИЯ И РАССЕЯНИЯ	54
3. МОМЕНТЫ И СЕМИИНВАРИАНТЫ	83
4. ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ	136
5. СТАНДАРТНЫЕ РАСПРЕДЕЛЕНИЯ — (1)	171
6. СТАНДАРТНЫЕ РАСПРЕДЕЛЕНИЯ — (2)	208
7. ИСЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ	250
8. ВЕРОЯТНОСТЬ И СТАТИСТИЧЕСКИЕ ВЫВОДЫ	275
9. СЛУЧАЙНЫЙ ВЫБОР	286
10. СТАНДАРТНЫЕ ОШИБКИ	318
11. ТОЧНЫЕ ВЫБОРОЧНЫЕ РАСПРЕДЕЛЕНИЯ	341
12. АППРОКСИМАЦИЯ ВЫБОРОЧНЫХ РАСПРЕДЕЛЕНИЙ — (1)	380
13. АППРОКСИМАЦИЯ ВЫБОРОЧНЫХ РАСПРЕДЕЛЕНИЙ — (2)	424
14. ПОРЯДКОВЫЕ СТАТИСТИКИ	446
15. МНОГОМЕРНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ И КВАДРАТИЧНЫЕ ФОРМЫ	477
16. РАСПРЕДЕЛЕНИЯ, СВЯЗАННЫЕ С НОРМАЛЬНЫМ	510

Приложение

ТАБЛИЦЫ

1. ФУНКЦИЯ ПЛОТНОСТИ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ	549
2. ФУНКЦИЯ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ	549
3. КВАНТИЛИ χ^2 -РАСПРЕДЕЛЕНИЯ	551
4а. ФУНКЦИЯ χ^2 -РАСПРЕДЕЛЕНИЯ С ОДНОЙ СТЕПЕНЬЮ СВОБОДЫ ДЛЯ $0 \leq \chi^2 \leq 1$	552

46. ФУНКЦИЯ χ^2 -РАСПРЕДЕЛЕНИЯ С ОДНОЙ СТЕПЕНЬЮ СВОБОДЫ, ДЛЯ $1 \leq \chi^2 \leq 10$	553
5. КВАНТИЛИ t -РАСПРЕДЕЛЕНИЯ	554
6. 5%-НЫЕ ТОЧКИ z -РАСПРЕДЕЛЕНИЯ	556
7. 5%-НЫЕ ТОЧКИ F -РАСПРЕДЕЛЕНИЯ	558
8. 1%-НЫЕ ТОЧКИ z -РАСПРЕДЕЛЕНИЯ	559
9. 1%-НЫЕ ТОЧКИ F -РАСПРЕДЕЛЕНИЯ	561
10. СИММЕТРИЧЕСКИЕ ФУНКЦИИ. ФОРМУЛЫ ПРЕДСТАВЛЕНИЯ РАСШИРЕННЫХ СИММЕТРИЧЕСКИХ ФУНКЦИЙ В ТЕРМИНАХ СУММ СТЕПЕНЕЙ И ОБРАТНЫЕ ФОРМУЛЫ.	562
11. МНОГОМЕРНЫЕ k -СТАТИСТИКИ. ТАБЛИЦЫ ПРЕДСТАВЛЕНИЯ t -СТАТИСТИК В ТЕРМИНАХ РАСШИРЕННЫХ СИММЕТРИЧЕСКИХ ФУНКЦИЙ И ОБРАТНЫЕ ТАБЛИЦЫ.	563
ЦИТИРОВАННАЯ ЛИТЕРАТУРА	565
УКАЗАТЕЛЬ	573

«Давайте присядем на это бревно у дороги, — говорю я, — и забудем бессердечность и сквернословие этих поэтов. Настоящую красоту нужно искать в великолепных рядах установленных фактов и общепринятых правил. В этом самом бревне, на котором мы сидим, миссис Сэмпсон, — говорю я, — скрыта статистика более прекрасная, чем любая поэма. Кольца показывают, что ему было шестьдесят лет. На глубине двух тысяч футов оно за три тысячи лет превратилось бы в уголь. Самая глубокая в мире угольная шахта находится в Киллингворте, близ Ньюкастла. В ящик длиной четыре фута, шириной три фута и высотой два фута восемь дюймов войдет тонна угля. Если артерия порезана, сожмите ее выше раны. В ноге человека — тридцать костей. Лондонский Тауэр горел в 1841 году». «Продолжайте, мистер Пратт», — говорит миссис Сэмпсон. — Эти мысли так оригинальны и приятны. Я думаю, ничего нет прекраснее этой статистики».

О. Генри, Справочник Гименея
(O. Henry, The Handbook of Hymen)

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Эта книга в ее первоначальном виде была задумана перед тем, как разразилась вторая мировая война, и большая часть монографии была написана в течение первых лет этого тяжелого времени. Первый том вышел в свет в 1943 году, второй — в 1946 году. С тех пор статистическая теория так далеко шагнула вперед и настолько сильно расширилась область применения статистических методов, что это представляется удивительным даже при сравнении с общим интенсивным развитием научных исследований. Пять переизданий первоначального варианта первого тома и три — второго тома отличались друг от друга, по существу, лишь исправлением ошибок и добавлением ссылок. Но время сделало необходимой полную ревизию. Сейчас издается первый том в новой редакции. Увеличение объема материала привело к тому, что теперь планируются еще два тома. Я огорчен разрастанием объема нового варианта, однако этого увеличения невозможно избежать, если намереваться охватить весь задуманный круг вопросов.

Пересмотренный вариант книги состоит из трех томов. Предметом настоящего нового тома I является теория распределений, включающая выборочные распределения. Во втором томе содержится теория оценивания, общая теория статистических выводов и теория статистической связи. Третий том будет посвящен теории планирования, анализу выборочных данных и экспериментов, многомерному анализу и временным рядам.

В настоящий том вошла значительная часть первоначального тома I. Кроме этого, сделан ряд существенных дополнений, особенно по следующим темам: стандартные распределения, теория аппроксимации выборочных распределений, порядковые статистики, многомерное нормальное распределение, распределения, связанные с нормальным. Три главы этого тома целиком

являются новыми, содержание же остальных глав существенно дополнено новым материалом. Весьма значительные дополнения и изменения будут необходимы в новых томах 2 и 3. Однако до их выхода пройдет несколько лет *), и я надеюсь, что в течение этого времени будет полезен первоначальный том 2.

В настоящем издании добавлено много новых примеров и упражнений. Теперь том 1 содержит свыше 300 упражнений. Если читатель, проработав их, будет чувствовать потребность в дальнейшей тренировке, он может обратиться к решению задач из моего сборника **), после чего он окажется в состоянии сам ставить себе задачи.

Приведенный в конце тома список литературы содержит только те работы, на которые имеются ссылки в тексте книги. Дальнейшая библиография будет приведена в томах 2 и 3.

В какой-то мере я чувствовал себя неспособным в одиночку выполнить нелегкий труд переработки старого издания. К счастью, мне удалось привлечь к этой работе моего коллегу м-ра Алэна Стьюарта, который взялся за нее со свойственными ему энергией и способностями. Новое издание следует рассматривать как результат нашего совместного труда, и мы в равной степени несем за него ответственность.

За разрешение воспроизвести некоторые из таблиц, помещенных в конце книги, мы весьма признательны д-ру Ф. Н. Дэвиду (F. N. David), проф. сэру Рональду Фишеру (sir Ronald Fisher), покойному д-ру Дж. Уишарту (J. Wishart), д-ру Ф. Иэйтсу (F. Yates), м-рам Оливеру и Бойду (Oliver and Boyd) и издателям журнала *Biometrika*. Всем им, так же как и авторам используемых нами работ, мы выражаем свою искреннюю благодарность. Мы благодарны также м-ру Дж. Дурбину (J. Durbin), прочитавшему с большим тщанием гранки книги и сделавшему ряд ценных предложений. Мы пользуемся случаем поблагодарить м-ра Е. В. Бёрка (E. V. Burke) из издательства Charles Griffin and Company Limited, взявшему на себя немалый труд, связанный с выходом книги в свет.

*) Второй том нового издания вышел в 1961 г. (Прим. перев.)

**) Kendall M., *Exercises in Theoretical Statistics. With answers and hints on solutions*, London, Griffin, 1954.

Некоторые читатели предыдущих изданий любезно указали нам на ряд опечаток и неточностей изложения, а также на несколько мест, требующих пояснения. Мы будем благодарны читателям за подобные замечания и к этому новому изданию.

М. К.

Лондон,
май 1958

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Это издание мало отличается от первого. Устранено незначительное число опечаток и ошибок, в ряде мест улучшено изложение, добавлено несколько новых упражнений и приведены ссылки на некоторые последние работы. Мы благодарны читателям, замечания которых способствовали этому улучшению книги.

М. К.,
А. С.

Лондон,
июнь 1962

ВВОДНЫЕ ЗАМЕЧАНИЯ

0.1 В этой книге разделы глав (пункты) нумеруются серийно. Номерам пунктов в серии предшествует номер главы, в которой они находятся, например 14.13 означает тринадцатый пункт главы 14. Подобный способ используется для нумерации таблиц, уравнений и упражнений, например (7.15) соответствует пятнадцатому уравнению главы 7. При ссылках номера пунктов обозначаются жирным шрифтом, номера таблиц, уравнений и т. д. — обычным шрифтом.

0.2 Ссылка на опубликованные работы дается посредством указания фамилии автора и года публикации. В списке литературы в конце книги фамилии авторов расположены по алфавиту. При ссылках на журнальные статьи номер тома печатается жирным шрифтом, а номер первой страницы статьи — обычным шрифтом, например *Ann. Math. Statist.* **10**, 275 обозначает статью, начинающуюся на 275-й странице 10-го тома журнала *The Annals of Mathematical Statistics*. Если после упражнения указывается фамилия автора и год, то это означает, что результат, содержащийся в упражнении, имеется в соответствующей статье, указанной в списке литературы.

0.3 В книге используются обычные математические обозначения, однако некоторые символы следует объяснить.

(1) Знак восклицания, написанный после целого числа, обозначает факториал этого числа. Некоторые авторы используют этот знак более общим образом, полагая для нецелых чисел

$$x! = \Gamma(x+1) = \int_0^{\infty} e^{-t} t^x dx.$$

Это, конечно, согласуется с обозначением факториала, однако не будет использовано в этой книге.

(2) Вместо более старого комбинаторного знака ${}^nC_r$, обозначающего $\frac{n!}{r!(n-r)!}$, используется знак $\binom{n}{r}$.

(3) Σ обозначает знак суммирования, например, $\sum_{j=1}^{j=n} x_j = x_1 + x_2 + \dots + x_n$. Символ $\sum_{j=1}^{j=n}$, как правило, сокра-

щается до $\sum_{j=1}^n$ и во многих случаях до $\sum_j$ или просто до Σ , если из контекста ясно, по какому множеству ведется суммирование.

(4) Используются обычные обозначения для Γ -функции (см. выше), В-функции и гипергеометрической функции:

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx = \frac{\Gamma(p) \Gamma(q)}{\Gamma(p+q)}$$

и

$$F(\alpha, \beta, \gamma, x) = 1 + \frac{\alpha\beta}{1 \cdot \gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1 \cdot 2 \cdot \gamma(\gamma+1)} x^2 + \\ + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1 \cdot 2 \cdot 3 \cdot \gamma(\gamma+1)(\gamma+2)} x^3 + \dots$$

(5) Если экспонента не громоздка, экспоненциальная функция записывается в виде степени e , например $e^{\frac{1}{2}x^2}$. Если же экспонента громоздка, то мы пишем, например, $\exp\left\{-\frac{1}{2}(x^2 - 2\rho xy + y^2)\right\}$, вместо $e^{-\frac{1}{2}(x^2 - 2\rho xy + y^2)}$.

(6) Запись выражения в виде $O(f(n))$ означает, что его порядок не выше порядка $f(n)$, т. е. что предел отношения этого выражения к $f(n)$ есть ограниченная постоянная C . Если $C=1$, то мы скажем, что выражение асимптотически эквивалентно $f(n)$; символически это указывается знаком $\sim$. Если $C=0$, то вместо $O(f(n))$ мы пишем $o(f(n))$ для обозначения того, что порядок выражения меньше порядка $f(n)$.

(7) Символ $\propto$, стоящий между двумя выражениями, означает, что эти выражения равны с точностью до постоянного множителя.

0.4 В некоторых случаях полезно пользоваться различными обозначениями для количества, относящегося к генеральной совокупности, и для оценки этого количества по выборке. Там, где это возможно, первое из них будет обозначаться греческой буквой, а второе — латинской, например: ρ — коэффициент корреляции в генеральной совокупности и r — выборочный коэффициент корреляции. Не всегда, однако, можно сохранить подобное различие: к примеру, прописная греческая буква P («ро») совпадает по написанию с прописной латинской буквой P («пэ»). Полной согласованности обозначений можно достигнуть лишь при отказе от многих прижившихся в статистике обычаев, и при этом весьма вероятно, что в результате получится очень громоздкая символика.

0.5 Для удобства читателя при разборе рабочих примеров и иллюстративного материала мы помещаем в конце этого тома несколько таблиц часто употребляемых функций. Эти таблицы ни в коей мере не заменяют опубликованных к настоящему времени обширных таблиц, являющихся необходимым инструментом для большинства практических и значительного числа теоретических работ. Частые ссылки будут делаться на следующие таблицы:

Biometrika Tables for Statisticians, vol. 1, Cambridge University Press, изданные Е. С. Пирсоном и Х. О. Хартли.

Statistical Tables for use in Biological, Agricultural and Medical Research, Oliver and Boyd, Edinburgh, созданные сэром Рональдом Фишером и Ф. Иэйтсом *).

В настоящее время имеется много более специальных таблиц. Нуждающемуся в них читателю лучше всего обратиться за советом к специалистам. Имеющиеся в тексте указания границ и шага табулирования приводятся в принятой ныне форме: например, $x=0(0,1)10(1)50(2)100$ означает, что функция табулирована для значений x , лежащих между 0 и 10 — с шагом 0,1, между 10 и 50 — с шагом 1, между 50 и 100 — с шагом 2.

*) Возможности этих сборников в значительной мере покрываются возможностями изданных на русском языке книг: Л. Н. Большев, Н. В. Смирнов, Таблицы математической статистики, М., Изд-во «Наука», 1965; Д. Б. Оуэн, Сборник статистических таблиц (перев. с англ.), М., Изд-во ВЦ АН СССР, 1966; Я. Янко, Математико-статистические таблицы (перев. с чешского), М., Госстатиздат, 1961. (Прим. ред.)

ЧАСТОТНЫЕ РАСПРЕДЕЛЕНИЯ

1.1 Фундаментальным понятием в статистической теории является понятие группы или семейства; для него статистики используют специальный термин «генеральная совокупность». Этот термин будет обычно применяться к любому набору рассматриваемых объектов, как одушевленных, так и неодушевленных. Например, мы будем рассматривать генеральные совокупности мужчин, растений, ошибок при снятии показаний приборов, высот ртутного столба барометра в различные дни и даже генеральные совокупности идей или планов, таких как всевозможные планы розыгрыша имеющихся на руке карт. Общим во всех этих примерах является понятие совокупности.

Статистика изучает свойства генеральных совокупностей. При статистическом рассмотрении генеральной совокупности мужчин мы не интересуемся, имеют ли некоторые из них коричневые глаза или являются ли некоторые из них фальшивомонетчиками, а интересуемся скорее тем, сколько индивидуумов имеют коричневые глаза или являются фальшивомонетчиками и связаны ли между собой эти качества. Мы имеем, так сказать, дело со свойствами самой совокупности. Такой подход встречается в физике и в демографии. Например, при рассмотрении свойств газа мы, как правило, интересуемся не столько поведением отдельных молекул, сколько поведением совокупности молекул, образующих газ. Статистик, подобно природе, имеет главным образом дело с видами, а не с отдельными индивидуумами.

1.2 Мы можем, таким образом, начать подход к определению нашего предмета, сказав, что статистика есть ветвь научного метода, изучающая свойства генеральных совокупностей. Это однако, слишком общо. Статистика рассматривает только численные свойства. Словарь, например, представляет собой генеральную совокупность слов, и среди свойств этой совокупности, представляющих собой разумный объект научного исследования, можно указать словопронхождение. Со статистической точки зрения неважно, какие слова имеют латинское, англосак-

сонское или индостанское происхождение. Статистический аспект появится только тогда, когда мы станем исследовать, *сколько* слов произошло из того или иного источника.

1.3 В качестве второго приближения к нашему определению мы можем теперь сказать, что статистика есть ветвь научного метода, изучающая данные, полученные исчислением или измерением свойств генеральной совокупности.

Это опять несколько слишком общо. Таблица логарифмов представляет собой генеральную совокупность чисел, однако вряд ли может быть объектом статистического исследования, поскольку каждое число в ней вполне определяется математическими законами. Статистик же интересуется скорее теми генеральными совокупностями, которые встречаются в Природе и являются объектом многочисленных и разнообразных влияний. Его генеральная совокупность в точности подчиняется простым математическим законам очень редко; именно при отходе от подобных законов часто обнаруживаются темы наибольшего статистического интереса. Учитывая сказанное, мы можем теперь сформулировать наше определение следующим образом:

Статистика есть ветвь научного метода, изучающая данные, полученные исчислением или измерением свойств генеральных совокупностей природных феноменов. В этом определении «природные феномены» включают все возможные объекты внешнего мира, как связанные, так и не связанные с человеком.

1.4 Во избежание недоразумений при толковании этого определения следует отметить, что «статистика», как название научного метода, является собирательным именем существительным и употребляется в единственном числе. Тем же словом «статистика» называют те числовые данные, с которыми работает этот метод, и в этом смысле слово «статистика» означает множество данных. В дальнейших разделах этой книги мы встретимся еще с одним значением слова «статистика» — так называют функцию от элементов выборки из некоторой генеральной совокупности. В этом смысле «статистика» употребляется как в единственном, так и во множественном числе.

Частотные распределения

1.5 Рассмотрим генеральную совокупность элементов, каждому из которых соответствует определенное числовое значение некоторой переменной, например, совокупность мужчин, распределенных по росту, или совокупность цветков, классифицированных в соответствии с числом лепестков в них. Эту переменную мы будем называть *вариантой*. Если она принимает только некоторое множество изолированных значений, то назо-

вом ее дискретной, если же она принимает непрерывное множество значений, то скажем, что она непрерывна. Генеральной совокупности элементов у нас соответствует генеральная совокупность значений варианты, и именно эту последнюю совокупность мы должны рассматривать.

Если генеральная совокупность состоит только из нескольких элементов, мы можем без труда рассматривать соответствующую генеральную совокупность значений варианты. Если же, как это обычно случается, совокупность велика (или бесконечна — в смысле, определяемом ниже), множество значений варианты должно быть некоторым образом редуцировано, чтобы наша мысль могла охватить его. Это осуществляется классифицированием элементов в соответствии с принадлежностью отвечающих им значений варианты некоторым интервалам, на которые разбивается область значений варианты. Настолько, насколько возможно, интервалы должны быть одинаковыми, с тем чтобы количества попаданий в различные интервалы были сравнимы. Эти интервалы называются класс-интервалами (или просто интервалами), а количества элементов, которым соответствуют значения варианты, попадающие в заданные класс-интервалы,

Таблица 1.1

Распределение областей Англии и Уэльса в соответствии с числом новорожденных на каждую 1000 жителей в 1953 году
(Данные из The Registrar-General's Statistical Review of England and Wales for 1953.)

Число новорожденных на 1000 жителей	Число областей	Число новорожденных на 1000 жителей	Число областей
2,5—	1	15,5—	250
3,5—	0	16,5—	178
4,5—	0	17,5—	129
5,5—	1	18,5—	64
6,5—	0	19,5—	39
7,5—	3	20,5—	20
8,5—	7	21,5—	12
9,5—	19	22,5—	7
10,5—	39	23,5—	1
11,5—	96	24,5—	2
12,5—	151	25,5—	3
13,5—	231	26,5—	0
14,5—	221	27,5—	1
		Общее число областей	1 475

З а м е ч а н и е: 2,5 — означает «от 2,5 до 3,5».

Таблица 1.2

Распределение лиц в Соединенном Королевстве, обложенных дополнительным подоходным налогом в течение года, начиная с 6 апреля 1950 г., в соответствии с величиной их годового дохода

(из Annual Abstract of Statistics, 1950)

Годовой доход (в тыс. фунтов)				Число лиц	Оцененная частота на 500-фунтовый интервал
не менее	2	и менее	2,5	60 336	60 336
»	»	2,5	»	41 033	41 033
»	»	3	»	45 532	22 766
»	»	4	»	23 263	11 632
»	»	5	»	13 475	6 737
»	»	6	»	13 456	3 364
»	»	8	»	6 419	1 605
»	»	10	»	3 551	888
»	»	12	»	2 926	488
»	»	15	»	2 007	201
»	»	20	»	820	82
»	»	25	»	399	40
»	»	30	»	376	19
»	»	40	»	134	6
»	»	50	»	128	2
»	»	75	»	45	1
»	»	100	и больше	38	?
Общее число лиц				213 938	—

называются класс-частотами (или просто частотами). Распределение класс-частот по класс-интервалам называется частотным распределением (или просто распределением).

1.6 Таблицы 1.1 и 1.2 содержат частотные распределения наблюдаемых генеральных совокупностей, классифицированных по одной из вариантов. Таблица 1.1 показывает распределение 1475 областей Англии и Уэльса на 1953 год, классифицированных по числу новорожденных на 1000 жителей. Распределение в этой таблице представляется простым и компактным способом, который был бы совершенно невозможен, если бы каждая из 1475 областей указывалась в отдельности. Наибольшее число областей попадает в интервал 15,5—16,5 рождаемостей на тысячу; в обе стороны от этого интервала частоты убывают. Таблица 1.2 показывает число лиц, обложенных добавочным подоходным налогом в Соединенном Королевстве*) в 1950 году, класси-

*) Соединенное Королевство — обычное сокращение для Соединенного Королевства Великобритании и Северной Ирландии. (Прим. перев.)

фицированных в соответствии с величиной дохода. Класс-интервалы здесь неравные (типичный недостаток официальных данных). В последнем столбце для сравнения указана редукция класс-частот; а именно, в каждом из рассматриваемых класс-интервалов вычислены частоты, соответствующие 500 фунтам. Просматривая этот столбец, видим, что максимальная частота на 500 фунтов находится в начале частотного распределения.

1.7 Частотное распределение можно изображать графически. Располагая по оси *икс* значения варианты и по оси *игрек* — значения частоты в класс-интервалах, мы относим абсциссам, соответствующим центрам класс-интервалов, ординаты, равные частоте, приходящейся на единичный интервал в соответствующем класс-интервале. Концы этих ординат последовательно соединяются прямыми линиями. Так полученная ломаная называется полигоном частот. На рис. 1.1 изображен полигон частот для данных таблицы 1.1.

Возможно иное графическое изображение частотного распределения. Мы можем на каждом класс-интервале, расположенном на оси *икс*, построить прямоугольник с площадью, пропорциональной частоте в этом интервале. Так построенная фигура называется гистограммой.

На рис. 1.2 изображена гистограмма данных таблицы 1.2. Очевидно, гистограмма является более подходящей формой изображения при неравных класс-интервалах.

1.8 Сделаем несколько практически важных замечаний относительно табулирования наблюдаемых частотных распределений.

(1) Уже отмечалось, что всякий раз, когда это возможно, класс-интервалы должны быть одинаковыми. Важность этого будет объяснена в последующих главах. Однако и сейчас очевидно, что визуально производить сравнения трудно, если класс-интервалы неодинаковы. Просматривая, например, сверху вниз второй столбец таблицы 1.2, замечаем, что частота в интервале 3—4 больше частоты в непосредственно предшествующем ему интервале. Это происходит, однако, из-за изменения в этом месте ширины интервала, и из третьего столбца видно, что частота на единичный интервал (равный 500 фунтов) постоянно убывает.

(2) Важно точно определять класс-интервалы. Нередко встречаются такие классификации, как «0—10, 10—20, 20—30» и т. д. К какому интервалу отнести элемент со значением варианты, равным 10? Очевидно, классификация неопределенная, если такое значение в действительности может появиться. Мы или должны взять интервалы «больше или равно 0 и меньше 10,

больше или равно 10 и меньше 20», или пользоваться каким-либо другим точным соглашением относительно того, куда причислять значения варианты, попадающие на границу между

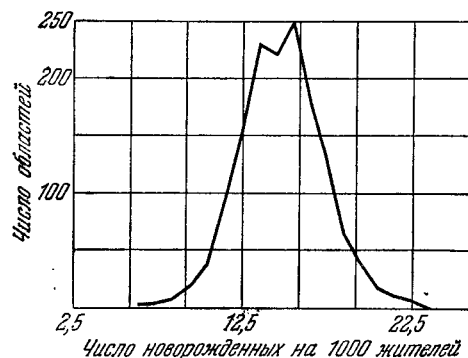


Рис. 1.1. Полигон частот для данных таблицы 1.1.

двумя соседними интервалами. Например, можно условиться относить по половине их числа к каждому из граничащих интервалов. Существуют различные способы обозначения класс-интервалов в практически используемых таблицах, например «10—, 20—, 30—» означает «больше или равно 10 и меньше 20» и т. д. Иногда при рассмотрении случаев, когда варианта непрерывна, имеется некоторая неопределенность в спецификации точности измерений. Например, при измерении длины в целых сантиметрах с округлением в сторону ближайшего целого интервал, указанный как «больше 15 и меньше 18», означает в действительности интервал «больше 15,5 и меньше 17,5». Когда точность измерения известна, можно специфицировать интервал указанием его средней точки, в рассмотренном случае, например, такой точкой будет 16,5.

(3) Несмотря на важность равенства класс-интервалов, отмеченного в замечании (1), бывает полезно рассмотреть более мелкие класс-интервалы на тех участках, где частота меняется очень быстро. Таблица 1.3, например, показывает число смертей от менингита в Англии и Уэльсе в 1953 году в соответствии с вариантом «возраст при смерти». Если бы частота в интервале «0 и меньше 5» не была подразделена, а была показана в общем как 208 на интервале, нам было бы неясно, происходит ли наибольшая часть смертей

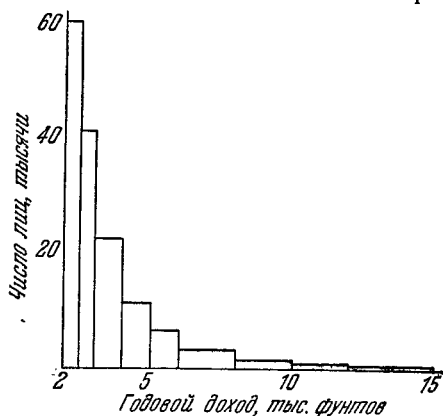


Рис. 1.2. Гистограмма для данных таблицы 1.2.

Таблица 1.3

Смертность от менингита (исключая цереброспинальный и туберкулезный менингит) в Англии и Уэльсе в 1953 году
(из The Registrar-General's Statistical Review for England and Wales for 1953)

Возраст в годах	Число смертей	Возраст в годах	Число смертей
0—	162	35—	5
1—	29	40—	11
2—	11	45—	17
3—	6	50—	21
4—	0	55—	32
5—	10	60—	23
10—	8	65—	22
15—	6	70—	15
20—	4	75—	11
25—	3	80—	10
30—	5	85 и выше	3
		Общее число смертей	414

в первый год жизни. Это в действительности так. Частоты для каждого года из первых пяти лет показывают относительно высокую смертность в первом году.

(4) Пожалуй, нет особой необходимости добавлять, что гистограмма не является хорошим изображением данных, классифицированных по дискретной вариате. Она показывает класс-частоту, равномерно распределенную по всему интервалу, в то время как в случае дискретной варианты частоты концентрируются в некоторых определенных точках.

Частотные распределения: дискретные варианты

1.9 Теперь полезно будет привести несколько примеров частотных распределений, встречающихся в практике.

Таблица 1.4 показывает распределение цифр в числах, взятых из телефонной книги (содержащей четырехзначные номера). Числа выбирались следующим образом: книга открывалась наугад и на открытой странице брались последние две цифры всех числе, за исключением напечатанных жирным шрифтом. Распределение нерегулярно, однако после беглого просмотра таблицы можно предположить, что все цифры встречаются одинаково часто в большой генеральной совокупности, из которой эти 10 000 элементов были выбраны. Мы увидим дальше, что имеющиеся здесь отклонения от средней частоты на цифру 1000

Таблица 1.4

Распределение цифр в числах из лондонской телефонной книги
(M. G. Kendall and Babington Smith, Journ. Roy. Statist. Soc. 101, 147)

Цифра	0	1	2	3	4	5	6	7	8	9	Общее количество цифр
Частота	1 026	1 107	997	966	1 075	933	1 107	972	964	853	10 000

нельзя полностью отнести за счет выборочных флуктуаций. Теперь же достаточно отметить, что эти данные наводят на мысль рассмотреть генеральную совокупность с одинаково частыми элементами.

Таблица 1.5 показывает распределение проб листы апельсиновых деревьев, пораженных черными щитовками (black — scaleinsect *). Нулю соответствует наибольшая частота, частоты для больших значений варианты убывают более или менее регулярно.

Таблица 1.5

Распределение 821 пробы (десять листьев и их стебель) с апельсиновых деревьев по числу черных щитовок в пробе
(W. M. Upholt and R. Craig, 1940, Jour. Econ. Entom. 33, 113)

Число щитовок	0	1	2	3	4	5	6	7	8	> 8	Общее число проб
Частота	199	124	106	65	42	46	36	30	14	159	821

В таблице 1.6, показывающей распределение женщин-самоубийц в некоторых немецких государствах в течение ряда лет

Таблица 1.6

Женские самоубийства в девяти немецких государствах в течение четырнадцати лет

(Von Bortkiewicz, Das Gesetz der kleinen Zahlen, 1898)

Число самоубийств в год	0	1	2	3	4	5	6	7	8	9	10 и больше	Всего
Частота	9	19	17	20	15	11	8	2	3	5	3	112

*) По-видимому, здесь имеется в виду *Parlatoria ziziphi* Lucas. (Прим. перев.)

в соответствии с вариантом «число самоубийц в год», частота достигает максимума на 1—3 самоубийцах в год, а затем медленно убывает.

Частотные распределения: непрерывные варианты

1.10 Таблица 1.7 показывает распределение взрослых мужчин в Соединенном Королевстве (включавшем в момент сбора

Таблица 1.7

Распределение взрослых мужчин, родившихся в Соединенном Королевстве (включая всю Ирландию) по росту

(Из Final Report of the Anthropometric Committee to the British association, 1883, стр. 256.) Поскольку измерения производились,

как указывается, в дюймах, кратных $\frac{1}{8}$ дюйма, класс-интервалами

являются здесь, по-видимому, интервалы $56\frac{15}{16} - 57\frac{15}{16}$,

$57\frac{15}{16} - 58\frac{15}{16}$ и т. д.

Рост без обуви (дюймы)	Число мужчин с ростом в указанных границах	Рост без обуви (дюймы)	Число мужчин с ростом в указанных границах
57—	2	68—	1 230
58—	4	69—	1 063
59—	14	70—	646
60—	41	71—	392
61—	83	72—	202
62—	169	73—	79
63—	394	74—	32
64—	669	75—	16
65—	990	76—	5
66—	1 223	77—	2
67—	1 329		
		Общее число мужчин	8 585

данных всю Ирландию), в соответствии с их ростом в дюймах*). Соответствующий полигон частот изображен на рис. 1.3. Мы видим, что распределение почти симметрично, максимальная ордината приходится на 67 дюймов, по обе стороны от своего максимума частота постоянно убывает.

*) 1 дюйм=2,5 см. (Прим. перев.)

Такое более или менее равномерное убывание частот довольно характерно для рассмотренных распределений, однако

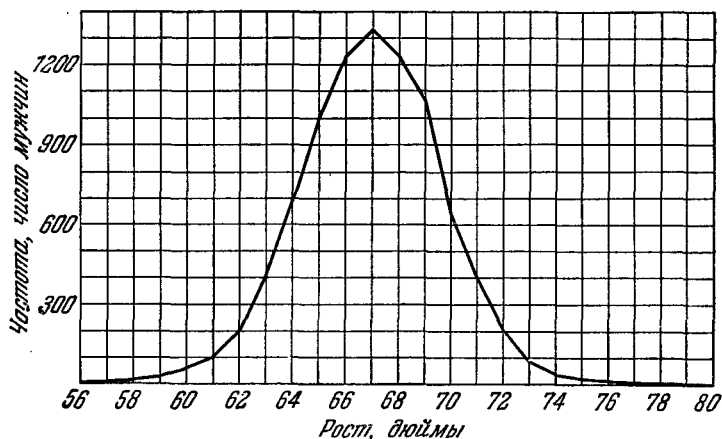


Рис. 1.3. Полигон частот для данных таблицы 1.7. Отмеченные значения абсцисс соответствуют началам класс-интервалов.

Таблица 1.8

Распределение свадеб в Австралии за 1907—1914 годы в соответствии с возрастом жениха (длины класс-интервалов взяты трехлетними)

(данные Преториуса (S. J. Pretorius), (1930), *Biometrika* 22, 210)

Возраст жениха (центр трехлет- него интервала, в годах)	Число свадеб	Возраст жениха (центр трехлет- него интервала, в годах)	Число свадеб
16,5	294	55,5	1 655
19,5	10 995	58,5	1 100
22,5	61 001	61,5	810
25,5	73 054	64,5	649
28,5	56 501	67,5	487
31,5	33 478	70,5	326
34,5	20 569	73,5	211
37,5	14 281	76,5	119
40,5	9 320	79,5	73
43,5	6 236	82,5	27
46,5	4 770	85,5	14
49,5	3 620	88,5	5
52,5	2 190		
		Общее число свадеб	301 785

свойство симметрии наблюдается в них относительно редко. Таблица 1.1 грубо симметрична, а таблицы 1.8 и 1.9, первая из которых показывает распределение браков в Австралии в соответствии с возрастом жениха и вторая — распределение молочных стад в соответствии со стоимостью молочных продуктов, иллюстрируют различные возможные степени асимметрии. Очень асимметричное распределение представлено в таблице 1.3. Таблицу 1.10, содержащую распределение мужчин в соответствии с весом, интересно сравнить с таблицей 1.7, показывающей их распределение в соответствии с ростом. Вторая симметрична, а первая нет.

Таблица 1.9

Распределение молочных стад в выборке из стад Англии и Уэльса в 1948—1949 годы в соответствии со стоимостью молочных продуктов
(Из National Investigation into the Economics of Milk Production — Milk Marketing Board, 1950.)

Цена в пенсах за галлон *)	Число стад	Цена в пенсах за галлон	Число стад
Меньше 12	3	24—	70
12—	19	26—	49
14—	52	28—	31
16—	96	30—	16
18—	121	32—	6
20—	115	34—	8
22—	86	36 и выше	7
		Всего	679

*) 1 галлон = 4,54 л. (Прим. перев.)

1.11 Если асимметрия распределения, как, например, в таблице 1.3, велика, мы можем оказаться не в состоянии определить, происходит ли вблизи его максимума убывание по обе стороны или максимум находится в самом начале распределения. В случае таблицы 1.3 положение было бы именно таким, если бы мы не имели данных относительно более мелкого группирования для первых пяти лет жизни. Такая ситуация наблюдается в таблице 1.2, где максимальная частота приходится, по-видимому, на или очень близко к доходу в 2000 фунтов в год. Асимметричные распределения иногда называют «косыми» (skew). Распределения такие, как представленное в таблице 1.2, называются «J-образными».

Таблица 1.10

Частотное распределение веса взрослых мужчин, родившихся в Соединенном Королевстве (Источник данных тот же, что и в таблице 1.7. Веса измерялись в целых фунтах*), так что действительными класс-интервалами являются 89,5—99,5, 99,5—109,5 и т. д.)

Вес в фунтах	Частота	Вес в фунтах	Частота
90—	2	190—	263
100—	34	200—	107
110—	152	210—	85
120—	390	220—	41
130—	867	230—	16
140—	1 623	240—	11
150—	1 559	250—	8
160—	1 326	260—	1
170—	787	270—	—
180—	476	280—	1
		Полная частота	7 749

*) 1 фунт = 453,6 г. (Прим. перев.)

1.12 В редких случаях распределение может достигать максимума на обоих концах. Таблица 1.11, показывающая распределение дней в соответствии со степенью облачности, доставляет

Таблица 1.11

Частоты оценок интенсивности облачности в Гринвиче с 1890 по 1904 (исключая 1901) годы для июля (Данные Gertrude E. Pearce (1928), Biometrika 20A, 336.)

Степень облачности	Частота	Степень облачности	Частота
10	676	4	45
9	148	3	68
8	90	2	74
7	65	1	129
6	55	0	320
5	45	Полная частота	1 715

соответствующий пример. Распределения такого типа называют «U-образными».

1.13 Встречаются также распределения, общий вид которых походит на части уже рассмотренных типов распределений. J-образное распределение, например, имеет сходство с «хвостом» симметричного распределения таблицы 1.7. Данные о самоубийствах таблицы 1.6 можно рассматривать как некоторое урезанное левее максимума (варианта здесь не может принимать отрицательных значений) симметричное распределение. Подобные соображения иногда полезны при подгонке кривых к наблюдаемым данным. Аналитическая кривая может очень хорошо соответствовать этим данным в некоторой области значений варианты, но и может продолжаться на значения, недостижимые, так сказать, для данных.

1.14 Рассмотренные до сих пор распределения имеют одну общую черту — они имеют только один максимум или, в случае U-образных кривых, только один минимум. Существуют также распределения, имеющие несколько максимумов. Таблицы 1.12 и 1.13 дают соответствующие примеры. Первая из них, показывающая распределение смертности в соответствии с возрастом при смерти, является типичной для подобного рода таблиц. Вблизи начала распределения имеется максимум, после которого частота быстро убывает, затем опять поднимается до ожидаемого максимума на возраст 70—75 лет и после этого второго максимума убывает к нулю. Естественно спросить, не будет ли

Таблица 1.12

Распределение смертности мужчин в Англии и Уэльсе за 1953 год в соответствии с возрастом при смерти

(Источник данных тот же, что и в таблице 1.1)

Возраст при смерти	Число смертей	Возраст при смерти	Число смертей
0—	12 244	45—	9 016
5—	1 043	50—	14 507
10—	665	55—	19 204
15—	1 104	60—	26 802
20—	1 640	65—	34 292
25—	1 932	70—	39 644
30—	2 449	75—	40 162
35—	3 068	85—	29 061
40—	5 104	85 и выше	17 553
		Общее число смертей	259 490

полезным рассмотреть такое распределение как суперпозицию двух или более распределений, в данном случае, например, как суперпозицию J-образного распределения, соответствующего дет-

Таблица 1.13

Распределение трипанозом в *Glossina morsitans* в соответствии с длиной в микронах

(по К. Пирсону (K. Pearson) (1914—1915), *Biometrika* 10, 112. Длина, по-видимому, в целых микронах)

Длина (микроны)	Частота	Длина (микроны)	Частота
15	7	26	110
16	31	27	127
17	148	28	133
18	230	29	113
19	326	30	96
20	252	31	54
21	237	32	44
22	184	33	11
23	143	34	7
24	115	35	2
25	130	Полная частота 2500	

ской смертности, и асимметричного распределения с максимумом на 70—75 годах, обычного для смертности от старости.

Подобное расчленение сложного распределения могло бы быть произведено и для данных таблицы 1.13, показывающей распределение трипанозом в мухе цеце, *Glossina morsitans*, в соответствии с их длиной. Здесь можно подозревать, что распределение образовано сложением нескольких распределений (это, между прочим, привело к предположению, что рассматриваемые трипанозомы представляют смесь различных своих типов).

Функции плотности и функции распределения

1.15 Рассмотренные выше примеры иллюстрируют тот замечательный факт, что большинство встречающихся на практике частотных распределений обладает значительной степенью регулярности. Вид полигонов частот и гистограмм, приведенных выше, естественно подсказывает, что наши данные суть аппроксимации распределений, задаваемых гладкими кривыми и простыми математическими выражениями. Этот подход к понятию *функции* плотности требует, однако, некоторой осторожности, особенно в случае непрерывных распределений.

Рассмотрим сначала дискретное распределение типа, задаваемого таблицей 1.4. Обозначим варианту буквой ξ . Тогда мы можем сказать, что ξ принимает только десять значений 0, 1, ..., 9 и что частота $f(x)$ значения ξ , равного x , определяется таблицей: $f(0) = 1026$, $f(1) = 1107$, $f(2) = 997$ и т. д. Таблица частот, таким образом, определяет функцию плотности. Более того, большая часть частот, содержащихся в таблице, приблизительно равна 1000, и мы можем рассматривать наблюдаемое

распределение как аппроксимацию к распределению

$$f(x) = 1000, x = 0, 1, \dots, 9 \quad (1.1)$$

или, более общо, к

$$f(x) = k, x = 0, 1, \dots, 9. \quad (1.2)$$

Это является, возможно, простейшим примером дискретной функции плотности, поскольку $f(x)$ здесь постоянна при всех возможных значениях x

Варианта в таблице 1.5, тоже дискретна, но может, теоретически говоря, принимать бесконечное множество значений — именно любое положительное число. Практически, конечно, и здесь варианта ограничена, но, так как мы не знаем границ принимаемых значений, мы можем представлять себе ее неограниченной. Функция плотности для этой таблицы по-прежнему определяется просто содержащимися в ней частотами, но если мы хотим перейти к «бесконечному» обобщению подобной таблицы, мы должны допустить, что функция плотности определена для всех целых положительных значений x . Это не представляет трудностей: нужно лишь позаботиться, чтобы можно было придать некоторый смысл полной частоте, т. е. чтобы ряд $\sum_{j=1}^{\infty} f(x_j)$ сходил.

1.16 Рассмотрим теперь случай непрерывной варианты. При обычных опытных данных распределения всегда дискретны, поскольку степень точности измерений ограничена. Например, мы обычно считаем, что рост человека в действительности может быть любым числом в некоторых пределах, скажем от 50 до 80 дюймов, хотя бы 20 π . Практически, однако, мы можем измерить рост только с некоторой степенью точности, например с точностью до одной тысячной дюйма. Измерения образуют при этом последовательность чисел, кратных тысячной дюйма, лежащих между 50 000 и 80 000, а такое значение как 62 831,85 (приблизительно равное 20 000 π) невозможно. Все физические измерения подчинены этому ограничению. Тем не менее мы говорим о «непрерывных» вариантах, предполагая при этом, что измерения суть аппроксимации к числам, которые могут принимать любое действительное значение.

1.17 Имея это в виду, мы можем рассматривать распределение частот при группировании как приближение к понятию функции плотности непрерывной варианты. Если бы в одном из рассмотренных выше распределений, скажем в распределении таблицы 1.7, мы стали разбивать интервалы на более мелкие, то, возможно, обнаружилось бы, что *до некоторого момента*

получающиеся частоты ведут себя все глаже и глаже. Читатель может убедиться в наличии этого эффекта, группируя данные таблицы 1.7 для интервалов длиной 8, 4 и 2 дюйма. Мы не можем, однако, продолжать процесс слишком далеко, поскольку в случае конечной генеральной совокупности дальнейшие подразбиения интервалов приведут к тому, что рано или поздно в каждом интервале будет лишь по нескольку элементов и поведение частот станет нерегулярным. Можно предположить, однако, что для не слишком малых величин Δx распределение может быть выражено функцией $f(x)\Delta x$, равной частоте для интервала длины Δx с центром в точке x , при всех x из области допустимых значений варианты. Можно предположить далее, что при стремлении Δx к нулю генеральная совокупность постоянно растет, благодаря чему предотвращается появление малых и нерегулярных частот. Этим путем мы подходим к понятию функции плотности непрерывной варианты. Мы пишем

$$dF = f(x)dx, \quad (1.3)$$

выражая этим, что элемент частоты dF для интервала между $x - \frac{1}{2}dx$ и $x + \frac{1}{2}dx$ равен $f(x)dx$ для всех x и малых dx .

1.18 Изложенный интуитивный подход к понятию непрерывного частотного распределения представляется наилучшим для статистических целей. Именно он первоначально привел к появлению этого понятия. Однако при формулировке аксиом и постулатов строгой математической теории рассматривается более обща функция F .

Пусть $F(x)$ — функция от x , определенная во всех точках некоторого интервала и непрерывная всюду, за исключением, может быть, счетного множества точек. Потребуем, чтобы F обращалась в нуль в наименьшей точке интервала (которая может равняться $-\infty$), была равна постоянной N в наибольшей точке (равной, может быть, $+\infty$) и не убывала ни в какой точке. Такая функция называется функцией распределения. Она соответствует накопленным частотам частотного распределения, а N соответствует полной частоте. Для таблицы 1.4, например, $F(x) = 0$ при $x < 0$, $F(x) = 1026$ при $0 \leq x < 1$, $F(x) = 2133 (= 1026 + 1107)$ при $1 \leq x < 2$ и т. д. $F(x)$ здесь имеет десять точек разрыва. Точки разрыва называются точками скачков, а функция $F(x)$ подобного вида называется ступенчатой функцией.

Если функция $F(x)$ не имеет скачков, то она всюду непрерывна и монотонно возрастает. Если $F(x)$ обладает производной, то имеет место уравнение

$$dF = F'(x)dx = f(x)dx, \quad (1.4)$$

соответствующее (1.3); $f(x)$ называется функцией плотности. С математической точки зрения изучение распределений сводится к изучению функций $F(x)$ и $f(x)$.

1.19 Так определенные функции в двух отношениях являются более общими, чем те, которые появляются в статистике: (I) $F(x)$ может возрастать монотонно без разрывов в некоторой части интервала и затем иметь скачок, т. е. частота может быть некоторое время непрерывной, а потом вдруг стать разрывной; в статистике варианта или непрерывна, или дискретна всюду на интервале ее изменения; (II) $F(x)$ может не иметь скачков и в то же время может не существовать функции плотности, поскольку непрерывная функция может быть недифференцируемой. Во всех рассматриваемых дальше случаях для непрерывной варианты функция плотности всегда существует.

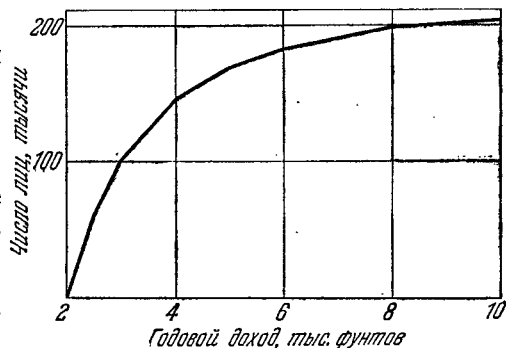


Рис. 1.4. Функция распределения для данных таблицы 1.2.

По причинам, которые станут ясными в главе 7, где излагается теория вероятностей, функцию $F(x)$ иногда называют вероятностной функцией. По существу, однако, она ничего общего с вероятностью не имеет, и мы будем пользоваться только термином «функция распределения».

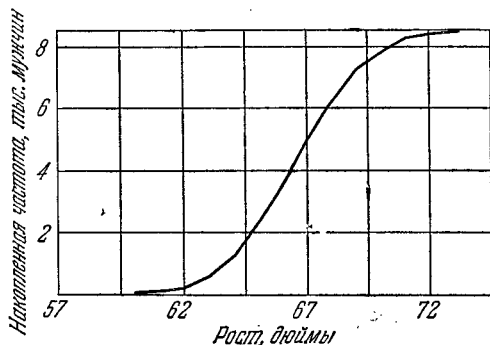


Рис. 1.5. Функция распределения для данных таблицы 1.7.

В обоих случаях мы несколько «сгладили» функции, соединив концы ординат, вместо того чтобы строить ступеньки. Если варианта непрерывна, то такое сглаживание очень часто

На рис. 1.4 показан график функции распределения для данных таблицы 1.2; значения x откладываются по оси абсцисс, значения $F(x)$ — по оси ординат. На рис. 1.5 показана функция распределения для данных табли-

оправдано; оно дает аппроксимацию непрерывного распределения, которому, как мы полагаем, подчиняются рассматриваемые данные.

1.20 Если $f(x)$ — дискретная функция плотности, а $F(x)$ — полная частота значений, не превосходящих x , то

$$F(x_r) = \sum_{j=1}^r f(x_j). \quad (1.5)$$

В непрерывном случае

$$F(x) = \int_a^x dF = \int_a^x f(u) du \quad (1.6)$$

(предполагается, что функция $F(x)$ задана на интервале (a, b)). Примем теперь два следующих соглашения, которые до некоторой степени упростят эти выражения. Будем предполагать, если не оговорено противное, что в этих математических выражениях частоты выражены в отношении к полной частоте, так что полная частота равна единице и сумма или интеграл по всему интервалу, где определена функция плотности, тоже равны единице, т. е. $F(b) = 1$. Далее, чтобы постоянно не указывать интервал (a, b) , мы можем предположить, не ограничивая общности, что $F(x)$ и $f(x)$ равны нулю при x меньших a и $F(x) = 1$, $f(x) = 0$ при x больших, чем b . Приняв эти соглашения, можем написать

$$\left. \begin{aligned} F(x_r) &= \sum_{j=-\infty}^r f(x_j), \\ F(x) &= \int_{-\infty}^x f(u) du \end{aligned} \right\} \quad (1.7)$$

и

$$\left. \begin{aligned} \sum_{j=-\infty}^{\infty} f(x_j) &= F(\infty) - F(-\infty) = 1, \\ \int_{-\infty}^{\infty} f(u) du &= F(\infty) - F(-\infty) = 1. \end{aligned} \right\} \quad (1.8)$$

Если необходимо принимать во внимание полную частоту N , то можно вернуться к безотносительным частотам, умножая на N частоты, задаваемые функцией плотности. Мы будем предполагать всегда, что функция $F(x)$ непрерывна справа.

1.21 Очень часто мы будем сокращать выражение «функция плотности» до ф. п., а «функция распределения» — до ф. р.

(в главе 4 встретится выражение «характеристическая функция», которое сокращается до χ ф.).

Еще одно соглашение поможет нам иногда избежать громоздких обозначений. Строго говоря, мы должны пользоваться различными символами для варианты (скажем ξ) и для тех значений, которые она может принимать (например x). Так $F(x)$, как функция от x , есть частота значений $\xi \leq x$. Часто удобно, однако, использовать один и тот же символ как для варианты ξ , так и для переменной x . Мы говорим при этом о варианте x с ф. п. $f(x)$ и ф. р. $F(x)$. Но всякий раз, когда это может привести к недоразумению, мы будем возвращаться к более строгим обозначениям.

Сведения об интегрировании

1.22 Мы уже отмечали выше, что функции, появляющиеся и рассматриваемые в статистической практике, менее сложны, чем общие функции распределения, рассматриваемые математиками. С этим фактом связаны постоянные трудности, которые встречаются при построении систематического курса теоретической статистики. Современная математика развила строгие теории большой общности, однако ценой этому является известная их трудность и абстрактность. Для большинства статистических целей этот сложный аппарат не требуется и встречающиеся в статистике функции могут быть одинаково исследованы более простыми методами без потери строгости. В настоящей книге мы, как правило, будем предпочитать эти более простые методы. Такой путь не позволит, однако, преодолеть все встречающиеся трудности. Некоторые из результатов, которые нам потребуются, не могут быть получены простыми способами, так называемые элементарные доказательства других слишком длинны и утомительны, чтобы их предпочесть более тонким методам, элементарное изложение третьих может затемнить их статистическую сущность.

1.23 Обратимся к современной теории интегрирования. Для многих целей было бы достаточно рассматривать интеграл Коши, определяемый обычно при начальном изложении. Пусть функция $f(x)$ определена на интервале (a, b) . Разделим его точками $x_1, x_2, \dots, x_n$ и рассмотрим сумму

$$s_n = f(a)(x_1 - a) + f(x_1)(x_2 - x_1) + \dots + f(x_n)(b - x_n). \quad (1.9)$$

Можно показать, что при некоторых условиях (таких как непрерывность $f(x)$ на (a, b)) суммы s_n стремятся к пределу, когда длины интервалов разбиения стремятся к нулю независимо от того, как выбираются точки деления и каким образом

происходит стремление к нулю. Этот предел и называется интегралом Коши: $\int_a^b f(x) dx$.

Определение более общего процесса интегрирования принадлежит Риману и Стильтесу. Пусть $\psi(x)$ — непрерывная функция от x . Выберем ξ_1 в промежутке от a до x_1 , ξ_2 в промежутке от x_1 до x_2 и т. д. Рассмотрим сумму

$$s'_n = \psi(\xi_1)\{F(x_1) - F(a)\} + \psi(\xi_2)\{F(x_2) - F(x_1)\} + \dots + \psi(\xi_n)\{F(b) - F(x_n)\}. \quad (1.10)$$

Можно показать, что, когда длина интервалов разбиения равномерно стремится к нулю, суммы s'_n стремятся к пределу, не зависящему от выбора точек ξ и точек деления. Предел этот записывают в виде

$$\int_a^b \psi(x) dF \quad (1.11)$$

и называют интегралом Стильтеса от функции $\psi(x)$ по функции $F(x)$. Риман рассматривал частный случай, когда $F(x) = x$.

1.24 Преимущество интеграла Стильтеса состоит в том, что он сводится к интегралу Коши, если функция плотности $f(x)$ непрерывна, и к обычному суммированию, если $f(x)$ дискретна. Чтобы не выписывать все формулы дважды, один раз для непрерывной ф. п. и другой раз для дискретной, мы будем обычно пользоваться интегралом этого типа, заменяя его в нужных случаях и в зависимости от обстоятельств обычными интегралом или суммой.

Многие результаты из теории обычного интегрирования верны и для интеграла Стильтеса. Нам часто придется пользоваться следующими его свойствами:

$$\left| \int_a^b \psi dF \right| \leq \int_a^b |\psi| dF \leq \quad (1.12)$$

$$\leq M \int_a^b dF \leq M, \quad (1.13)$$

где M — верхняя грань $|\psi(x)|$ в интервале (a, b) ;

$$\int_a^b \psi dF = \psi(\xi) \int_a^b dF, \quad (1.14)$$

где ξ — некоторая точка, принадлежащая интервалу (a, b) .

Если ряд $\sum f_j(x)$ сходится равномерно на (a, b) , то

$$\int_a^b \sum_{j=1}^{\infty} f_j(x) dF = \sum_{j=1}^{\infty} \int_a^b f_j(x) dF. \quad (1.15)$$

Обычные правила интегрирования по частям также применимы к интегралу Стильтьеса.

1.25 Более сложный интеграл был определен Лебегом. Его значение совпадает с интегралом Римана — Стильтьеса, если последний существует, но область его определения шире, чем у интеграла Римана — Стильтьеса. В обычной теории распределений интеграл Лебега, как правило, не применяется, однако для строгого изложения преобразований Фурье, в котором мы нуждаемся в ряде разделов теории временных рядов, его использование представляется существенным *).

Преобразования вариант

1.26 Предположим, что переменная y функционально связана с переменной x соотношением

$$x = x(y), \quad y = y(x), \quad (1.16)$$

причем функция y непрерывна и дифференцируема по x так же, как и x относительно y .

Тогда

$$dx = \frac{dx}{dy} dy. \quad (1.17)$$

Следовательно, для непрерывного распределения с ф. р. $F(x)$

$$F(x) = \int_{-\infty}^x dF(x) = \int_{-\infty}^x f(x) dx = \int_{-\infty}^{x(y)} f(x) \frac{dx}{dy} dy.$$

Можно написать, таким образом,

$$dF = f\{x(y)\} \frac{dx}{dy} dy. \quad (1.18)$$

Это соотношение выражает тот факт, что элемент частоты в интервале между $y - \frac{1}{2} dy$ и $y + \frac{1}{2} dy$ равен $f\{x(y)\} dx/dy$. Если

*) Читатель, чувствующий, что его математические знания нужно углубить или дополнить, может воспользоваться любым курсом высшей математики для вузов. Этого достаточно для занятий статистикой на вводном уровне. Более специальные сведения содержатся в книге Г. Крамера, Математические методы статистики, ИЛ, 1948.

ξ — варианта, соответствующая x , то это соотношение можно рассматривать как определяющее варианту η , соответствующую y . Уравнение, определяющее функцию плотности, преобразуется, стало быть, как если бы это было уравнение в дифференциалах. Такие преобразования вариант весьма важны в теории непрерывных распределений. С их помощью многие определенные математически распределения могут быть сведены, в точности или приблизительно, к известным формам.

Например, в выборочной теории нам придется изучать распределение

$$dF = \frac{1}{2^{\frac{1}{2}v-1} \Gamma\left(\frac{1}{2}v\right)} e^{-\frac{1}{2}\chi^2} \chi^{v-1} d\chi, \quad 0 \leq \chi \leq \infty. \quad (1.19)$$

Легко проверяется, что $F(\infty) = 1$.

Посредством преобразования $t = \frac{1}{2}\chi^2$ это распределение сводится к

$$dF = \frac{1}{\Gamma\left(\frac{1}{2}v\right)} e^{-t} t^{\frac{1}{2}v-1} dt, \quad 0 \leq t \leq \infty. \quad (1.20)$$

Полученное распределение имеет хорошо известную в анализе форму: его функция распределения есть неполная Γ -функция

$$F(t) = \Gamma_t\left(\frac{1}{2}v\right) / \Gamma\left(\frac{1}{2}v\right). \quad (1.21)$$

Распределение

$$dF = \frac{y_0}{(1+t^2/v)^{\frac{1}{2}(v+1)}} dt, \quad -\infty \leq t \leq \infty \quad (1.22)$$

(y_0 выбирается так, чтобы $F(\infty) = 1$) симметрично, одновершинно, существенно не ограничено. Оно напоминает распределение рис. 1.3. Преобразование $t = \sqrt{v} \operatorname{tg} \theta$ переводит его в

$$dF = \frac{y_0 \sqrt{v} \sec^2 \theta d\theta}{\sec^{v+1} \theta} = y_0 \sqrt{v} \cos^{v-1} \theta d\theta, \quad -\frac{1}{2}\pi \leq \theta \leq \frac{1}{2}\pi. \quad (1.23)$$

Полученное распределение сосредоточено на отрезке $\left[-\frac{1}{2}\pi, \frac{1}{2}\pi\right]$, но по-прежнему симметрично. Положим $\sin \theta = u$. Тогда

$$dF = y_0 \sqrt{v} (1-u^2)^{\frac{1}{2}v-1} du, \quad -1 \leq u \leq 1. \quad (1.24)$$

и, заменив еще u^2 на x ,

$$dF = y_0 \sqrt{v} x^{-\frac{1}{2}} (1-x)^{\frac{1}{2}v-1} dx, \quad 0 \leq x \leq 1. \quad (1.25)$$

По поводу последнего из этих преобразований нужно отметить следующее. Когда u меняется от -1 до $+1$, x меняется от $+1$ до 0 и затем опять до $+1$. При этом отрезку $(0, x)$ для x -распределения соответствует отрезок $(-u^2, u^2)$ для u -распределения. В тех случаях, когда применяемые преобразования не взаимно однозначны, следует обращать внимание на явления, подобные только что указанному.

1.27 Остановимся специально на следующем преобразовании варианты. Пусть

$$dF = f(x) dx,$$

Положим

$$y = \int_{-\infty}^x f(u) du.$$

Тогда

$$dF = f(x) \frac{dx}{dy} dy = \frac{f(x)}{f(x)} dy = dy, \quad 0 \leq y \leq 1. \quad (1.26)$$

Мы видим, что при этом преобразовании распределение преобразуется в распределение, имеющее очень простую «прямоугольную форму»: оно сосредоточено на отрезке $(0, 1)$ и все значения этого отрезка одинаково часты. Всякое непрерывное распределение может быть преобразовано к прямоугольной форме. Любое непрерывное распределение может быть преобразовано в любое другое непрерывное распределение: можно взять суперпозицию двух преобразований, первое из которых переводит одно распределение в прямоугольное, а второе представляет собой обратное преобразование, переводящее прямоугольное во второе распределение.

Генезис частотных распределений

1.28 До сих пор мы не интересовались происхождением различных наблюдаемых частотных распределений, рассмотренных выше в целях иллюстрации. Некоторые из них могут рассматриваться вне всяких вопросов об их происхождении из большей совокупности. Одним из таких примеров является распределение смертности, приведенное в таблице 1.12. Если интересоваться только распределением смертности мужчин в Англии и Уэльсе в 1953 году, то мы имеем перед собой всю рассматриваемую генеральную совокупность.

Однако в большей части случаев совокупность, которую мы в состоянии исследовать, представляет собой лишь часть той большей генеральной совокупности, на которой сосредоточено наше основное внимание. Распределение ростов в таблице 1.7 относится только к части генеральной совокупности мужчин Соединенного Королевства, живших во время обследования, и основной его интерес заключен в той информации, которую оно дает об этой генеральной совокупности. Точно так же распределение молочных стад в таблице 1.9 интересно главным образом в силу той информации, которая в нем содержится относительно стоимостей молочных продуктов во всей стране.

1.29 В только что упомянутых двух случаях (рост мужчин и стоимость молочных продуктов) мы обладаем информацией о некоторой выборке элементов, извлеченных из реально существующей генеральной совокупности. Только недостаток времени и соответствующих условий мешает нам исследовать всю генеральную совокупность. Но иногда случается, что имеющиеся

Таблица 1.14

Распределение результатов бросаний 12 костей по числу успехов (по числу выпаданий 4, 5 и 6)

(Данные Уэлдона (Weldon), цитированные Эджвортом (F. Y. Edgeworth, *Encyclopedia Britannica*, 11th ed., 22, 39).)

Число успехов	Частота	Число успехов	Частота
0	0	7	847
1	7	8	536
2	60	9	257
3	198	10	71
4	430	11	11
5	731	12	0
6	948		
		Полная частота	4 096

данные не относятся к существующей конечной генеральной совокупности. Соответствующий пример дает таблица 1.14, показывающая распределение результатов бросаний игральных костей.

Ясно, что в некотором смысле осуществленные бросания не составляют всей генеральной совокупности, поскольку процесс бросаний мы можем продолжить. В то же время эти дальнейшие бросания не существуют в том смысле, в каком существуют необследованные мужчины Соединенного Королевства или необследован-

ные молочные стада Англии и Уэльса. Их существование в некотором роде гипотетично и «даровано» самим понятием бросания костей.

Даже распределения, представляющиеся на первый взгляд отвечающими реальным совокупностям, могут рассматриваться в этом свете. Например, распределение трипанозом в таблице 1.13 получено обследованием определенных мух цеце. Можно считать, что выборка произведена из совокупности всех существ-

вующих мух цеце, содержащих трипанозомы или нет, реально существующей совокупности, но можно также считать ее выборкой из гипотетической совокупности, если предположить, что все мухи цеце заражены трипанозомами.

Исходная для наблюденного распределения генеральная совокупность имеет фундаментальное значение для статистических выводов. Мы еще вернемся к этому вопросу в последующих главах, где рассматривается проблема выбора. Отметим же здесь мы это потому, что нам придется с этим иногда сталкиваться до упомянутых глав. Следует заметить, что различие реальных и воображаемых генеральных совокупностей не является просто предметом онтологической спекуляции (если бы это было так, мы могли бы без ущерба игнорировать его), оно практически важно, когда делаются выводы о генеральной совокупности по извлеченной из нее выборке.

Многомерные распределения

1.30 В предыдущих пунктах мы рассматривали элементы генеральной совокупности в соответствии с одной вариантой. Построенные там частотные распределения могут быть названы, следовательно, одномерными. Содержание этих параграфов легко может быть обобщено так, чтобы включить в теорию генеральные совокупности, рассматриваемые в соответствии с двумя или более вариантами. При этом будут получаться двумерные, трехмерные и т. д. многомерные частотные распределения. Таблица 1.15, например, показывает распределение партии бобов в соответствии с их длиной и шириной. Нижняя строка и крайний правый столбец содержат частоты одномерных распределений в соответствии с длиной и шириной бобов в отдельности, а средняя часть таблицы показывает совместное изменение этих качеств.

Что касается одномерных вариантов, если их рассматривать в отдельности, то они могут быть как дискретными, так и непрерывными, и иногда встречаются случаи, когда одна варианта — одного типа, а другая — другого.

1.31 При обобщении полигона частот и гистограммы для представления двумерного распределения строятся трехмерные фигуры. Представим себе горизонтальную плоскость с парой перпендикулярных осей на ней, разделенную на клетки подобно шахматной доске, причем прямые, образующие клетки, проведены перпендикулярно осям в концах класс-интервалов. В центре каждого двумерного интервала (в клетке) перпендикулярно восстанавливаем отрезок, пропорциональный по длине частоте, отвечающей интервалу. Вершины построенных вертикальных отрезков

Распределение бобов по их длине и ширине

(Данные Йохансена (Johannsen), цитированные Преториусом (S. J. Pretorius), (1930), Biometrika 22, 110.)

Длина в мм (центральные значения)

	17	16,5	16	15,5	15	14,5	14	13,5	13	12,5	12	11,5	11	10,5	10	9,5	Всего
9,125	—	2	—	—	3	—	—	—	—	—	—	—	—	—	—	—	5
8,875	4	8	17	19	—	—	—	—	—	—	—	—	—	—	—	—	48
8,625	2	23	101	156	93	23	2	—	—	—	—	—	—	—	—	—	400
8,375	—	18	105	494	574	227	56	9	—	—	—	—	—	—	—	—	1 483
8,125	—	4	41	375	956	913	362	73	12	3	—	—	—	—	—	—	2 742
7,875	—	—	7	81	385	871	794	330	89	19	3	—	—	—	—	—	2 579
7,625	—	—	1	4	65	236	469	361	175	55	27	4	—	—	—	—	1 397
7,375	—	—	—	—	6	23	91	137	124	78	37	22	11	—	1	—	530
7,125	—	—	—	—	—	1	13	18	28	35	25	32	11	6	1	—	170
6,875	—	—	—	—	—	—	—	1	9	8	21	12	13	7	1	—	72
6,625	—	—	—	—	—	—	—	—	—	—	2	—	1	4	3	—	10
6,375	—	—	—	—	—	—	—	—	—	1	—	—	—	1	1	1	4
Всего	6	55	275	1 129	2 082	2 294	1 787	929	437	199	115	70	36	18	7	1	9 440

Ширина в мм (центральные значения)

соединяются: вершина отрезка, проведенного из центра некоторой клетки, с вершинами четырех других отрезков, проведенных из центров тех соседних клеток, которые имеют общую сторону с исходной клеткой. Получившаяся фигура называется двумерным полигоном частот, или стереограммой.

Подобным образом на каждой клетке можно построить прямоугольный параллелепипед, имеющий основанием эту клетку и по объему пропорциональный частоте, отвечающей этой клетке. В результате получается двумерная гистограмма. На рис. 1.6 показана гистограмма для данных таблицы 1.15.

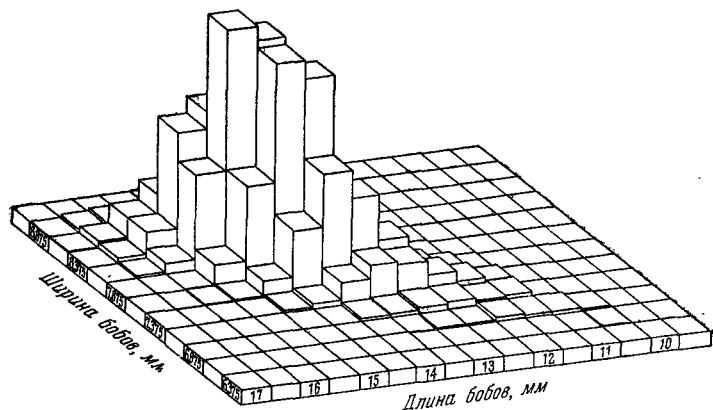


Рис. 1.6. Двумерная гистограмма для данных таблицы 1.15.

1.32 Непрерывное двумерное распределение вариант x_1 и x_2 можно записать в виде

$$dF = f(x_1, x_2) dx_1 dx_2. \quad (1.27)$$

При выполнении обычных соглашений двумерная функция распределения равна тогда

$$F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(u_1, u_2) du_1 du_2. \quad (1.28)$$

В случае дискретного распределения интеграл в (1.28) заменяется суммой.

Если зафиксировать какое-либо значение одной варианты, скажем $x_2 = k$, то другая варианта будет иметь отвечающее ей одномерное распределение. Так как в соответствии с принятыми соглашениями полная частота должна равняться единице, то

это распределение имеет вид

$$dF_0(x_1) = \frac{\int_{-\infty}^{\infty} f(x_1, k) dx_1}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, k) dx_1 dk}. \quad (1.29)$$

Распределение (1.29) называется условным распределением x_1 при условии $x_2 = k$.

Маргинальные (частные) распределения

1.33 Суммы на краях таких таблиц, как таблица 1.15, задают частотные распределения каждой из вариантов в отдельности. Частота любого значения варианты x_2 получается суммированием по всем значениям x_1 так, что если $F(x_1, x_2)$ — ф. р. вариант x_1 и x_2 , то одномерная ф. р. x_1 равна $F(x_1, \infty)$, а одномерная ф. р. x_2 есть $F(\infty, x_2)$. В терминах функции плотности

$$dF_1(x_1) = \left\{ \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 \right\} dx_1, \quad (1.30)$$

$$dF_2(x_2) = \left\{ \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 \right\} dx_2. \quad (1.31)$$

Эти одномерные распределения называют маргинальными (или частными) распределениями. В экономическом контексте термина «маргинальный» в таком понимании лучше избегать, поскольку в экономике ему обычно придают другой смысл; вообще же он удобен и в практике его применение едва ли вызывает какие-либо трудности.

Независимость

1.34 Если

$$F(x_1, x_2) = F(x_1, \infty) F(\infty, x_2) = F_1(x_1) F_2(x_2), \quad (1.32)$$

то x_1 и x_2 называются независимыми. В том случае, когда существуют ф. п., имеем

$$\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{x_1} f_1(x_1) dx_1 \int_{-\infty}^{x_2} f_2(x_2) dx_2,$$

и, следовательно,

$$f(x_1, x_2) = f_1(x_1) f_2(x_2). \quad (1.33)$$

Таким образом, ф. р. и ф. п. распадаются на два множителя, которые суть ф. р. и ф. п. для каждой варианты в отдельности. Легко видеть, что это определение статистической независимости согласуется как с обычным, так и с математическим пониманием независимости. Условное распределение x_2 при любом фиксированном x_1 (например, распределение в строке или в столбце двумерной таблицы частот) одно и то же, т. е. распределение x_2 независимо от x_1 .

Две варианты, не являющиеся независимыми, называются зависимыми. Очевидно, варианты в таблице 1.15 зависимы, поскольку распределения как в строках, так и в столбцах далеко не одинаковы.

Пример 1.1

Рассмотрим распределение

$$dF = z_0 \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2} \right) \right\} dx_1 dx_2, \\ -\infty \leq x_1, x_2 \leq \infty. \quad (1.34)$$

Число z_0 здесь выбрано так, чтобы полная частота равнялась единице. Положим

$$y_1 = \frac{1}{(1-\rho^2)^{\frac{1}{2}}} \left(\frac{x_1}{\sigma_1} - \frac{\rho x_2}{\sigma_2} \right), \quad y_2 = \frac{x_2}{\sigma_2}.$$

Имеем

$$\frac{\partial (y_1, y_2)}{\partial (x_1, x_2)} = \begin{vmatrix} \frac{1}{\sigma_1 (1-\rho^2)^{\frac{1}{2}}} & -\frac{\rho}{\sigma_2 (1-\rho^2)^{\frac{1}{2}}} \\ 0 & \frac{1}{\sigma_2} \end{vmatrix} = \frac{1}{\sigma_1 \sigma_2 (1-\rho^2)^{\frac{1}{2}}}$$

и

$$y_1^2 + y_2^2 = \frac{1}{1-\rho^2} \left\{ \frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2} \right\}.$$

Распределение, следовательно, принимает вид

$$dF = z_0 \sigma_1 \sigma_2 (1-\rho^2)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y_1^2 + y_2^2) \right\} dy_1 dy_2 = \\ = z_0 \sigma_1 \sigma_2 (1-\rho^2)^{\frac{1}{2}} e^{-\frac{1}{2} y_1^2} dy_1 e^{-\frac{1}{2} y_2^2} dy_2. \quad (1.35)$$

Исходные варианты, очевидно, зависимы, так как (1.34) не может быть разложено в произведение функции, зависящей только

от x_1 , на функцию, зависящую только от x_2 . Из формулы (1.35) вытекает, что преобразованные варианты независимы.

Заметим, что (1.35) позволяет легко вычислить z_0 . Действительно, так как

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du = \sqrt{2\pi},$$

то, интегрируя правую часть (1.35) по y_1 и y_2 и приравнявая полученное выражение к единице, получаем

$$z_0 = 1 / \left\{ 2\pi\sigma_1\sigma_2(1 - \rho^2)^{\frac{1}{2}} \right\}.$$

1.35 В рассмотренном примере мы использовали якобиан для перехода от одной пары вариантов к другой. Это преобразование имеет, конечно, общий характер. Если

$$dF = f(x_1, x_2) dx_1 dx_2$$

и

$$x_1 = x_1(y_1, y_2), \quad x_2 = x_2(y_1, y_2),$$

то функция плотности вариант y_1 и y_2 дается формулой

$$f\{x_1(y_1, y_2), x_2(y_1, y_2)\} J, \quad (1.36)$$

где

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix},$$

причем якобиан J в (1.36) должен браться по абсолютной величине. Если J меняет знак в области интегрирования, то могут возникнуть некоторые неясности, требующие специального внимания. Мы будем останавливаться на этом каждый раз, когда это необходимо.

1.36 Переход от двумерного случая к n -мерному не представляет затруднений. Например, если ф. р. $F(x_1, \dots, x_n)$ может быть записана в виде

$$F(x_1, x_2, \dots, x_n) = F(x_1, \infty, \dots, \infty) F(\infty, x_2, \dots, \infty) \dots F(\infty, \infty, \dots, x_n), \quad (1.37)$$

то соответствующие варианты независимы. Здесь следует сделать замечание: чтобы n вариант были независимы, недостаточно, чтобы каждые две из них были независимы, т. е. уравнение (1.37) может не выполняться, в то время как

$$\begin{aligned} F(x_1, x_2, \infty, \dots, \infty) &= \\ &= F(x_1, \infty, \infty, \dots, \infty) F(\infty, x_2, \infty, \dots, \infty) \end{aligned} \quad (1.38)$$

и аналогично для остальных x . Когда речь идет о независимости совокупности вариантов, число которых больше двух, важно, следовательно, точно указывать, что имеется в виду. Варианты, удовлетворяющие условию (1.37), назовем *взаимно* независимыми. Если каждые две варианты из совокупности удовлетворяют соотношению типа (1.38), то скажем, что эти варианты попарно независимы.

Производящие функции

1.37 При изучении распределений, заданных математически, особенно дискретных, часто удобно пользоваться представляющими их производящими функциями. Предположим, что частота варианты x , принимающей значения $0, 1, 2, \dots$, равна f_r для $x=r$. Производящей функцией варианты x называется функция

$$P(t) \equiv \sum_{r=0}^{\infty} f_r t^r. \quad (1.39)$$

Найдем, для примера, производящую функцию так называемого биномиального распределения, которое мы изучим довольно подробно позднее. Для этого распределения частота значения $x=r$ ($r=0, 1, \dots, n$) равна

$$f_r = \binom{n}{r} p^r (1-p)^{n-r}, \quad 0 \leq p \leq 1, \quad (1.40)$$

и производящая функция, следовательно, имеет вид

$$P(t) = \sum_{j=0}^n \binom{n}{j} p^j (1-p)^{n-j} t^j = \{pt + (1-p)\}^n. \quad (1.41)$$

УПРАЖНЕНИЯ

1.1 Построить полигоны частот и гистограммы следующих распределений:

Т а б л и ц а 1.16

Распределение результатов 4026 бросаний 12 костей в соответствии с числом успехов; успехом считается выпадение 6 очков

(Данные Уэлдона; источник тот же, что и в таблице 1.14)

Число успехов	0	1	2	3	4	5	6	7 и больше	Общее число бросаний
Число бросаний	447	1 145	1 181	796	380	115	24	8	4096

Таблица 1.17

Распределение предложений по их длине в отрывках из очерков Макося о Бэкоме и Чатаме

(По Юлу (G. Uduy Yulu) (1939), Biometrika, 33, 363.)

Длина предложения в словах	Число предложений	Длина предложения в словах	Число предложений
1—5	46	66—	2
6—	204	71—	4
11—	252	76—	8
16—	200	81—	2
21—	186	86—	2
26—	108	91—	1
31—	61	96—	2
36—	68	101—	1
41—	38	106—	—
46—	24	111—	1
51—	20	116—	—
56—	12	121—	1
61—	8		
		Общее число предложений	1 251

Таблица 1.18

Распределение черепов древних египтян в соответствии с длиной левой затылочной кости (в мм)

(По Ву (T. L. Woo) (1930), Biometrika 22, 324.)

Длина (центральные значения)	Частота	Длина (центральные значения)	Частота
84,5	12	102,5	74
86,5	12	104,5	68
88,5	32	106,5	36
90,5	48	108,5	18
92,5	79	110,5	7
94,5	116	112,5	4
96,5	104	114,5	4
98,5	126	116,5	—
100,5	123	118,5	1
		Полная частота	864

Таблица 1.19

Распределение племенных конематок (скаковых лошадей) по плодовитости; отношение числа произведенных годов алых жеребят к числу покрытий (учитывались только кобылы, покрытые не меньше восьми раз)

(Пирсон, Ли и Мур (Pearson, Lee and Moore) (1899), Phil. Trans. A192, 303. Из случаев, отвечающих границе между двумя интервалами, половина относилась к каждому из граничащих интервалов.)

Плодовитость	Число кобыл, плодовитость которых находится в указанных границах	Плодовитость	Число кобыл, плодовитость которых находится в указанных границах
1/30—3/30	2,0	17/30—19/30	315,0
3/30—5/30	7,5	19/30—21/30	337,0
5/30—7/30	11,5	21/30—23/30	293,5
7/30—9/30	21,5	23/30—25/30	204,0
9/30—11/30	55,0	25/30—27/30	127,0
11/30—13/30	104,5	27/30—29/30	49,0
13/30—15/30	182,0	29/30—1	19,0
15/30—17/30	271,5		
		Общее число кобыл	2 000

Т а б л и ц а 1.20

Частотное распределение участков площадью 1/500 акра, взятых на пшеничном поле, в соответствии с урожаем зерна в фунтах

(Мерсер и Холл (Mercer and Hall) (1911), Journ. Agr. Science, 4, 107.)

Урожай зерна в фунтах на 1/500 акра (центральное значение интервала)														Общее число участков
2,8	3,0	3,2	3,4	3,6	3,8	4,0	4,2	4,4	4,6	4,8	5,0	5,2		
Число участков	4	15	20	47	63	78	88	69	59	35	10	8	4	500

Т а б л и ц а 1.21

Частотное распределение 166 водителей лондонских автобусов в соответствии с числом случившихся с ними несчастных происшествий в течение:
а) одного года, б) пяти лет

(Фармер и Чеимберс (E. Farmer and E. G. Chambers) (1939)) (Industrial Health Board Report 84, London, H. M. Stationery Office.)

Число несчастных происшествий	Число водителей, с которыми случилось указанное число несчастных происшествий		Число несчастных происшествий	Число водителей, с которыми случилось указанное число несчастных происшествий	
	в течение одного года	в течение пяти лет		в течение одного года	в течение пяти лет
0	45	1	8	1	14
1	36	2	9	—	12
2	40	3	10	—	13
3	19	14	11	—	9
4	12	17	12	—	6
5	8	21	13	—	2
6	3	17	14	—	6
7	2	14	15 или больше	—	15
			Общее число водителей	166	166

Таблица 1.22

Распределение группы сделавших аборт женщины по времени, прошедшем между зачатием и абортом (в неделях) (по Пирсу (T. V. Pearce) (1930), Biometrika 22, 250)

Время (недели)	Частота	Время (недели)	Частота
4	3	17	13
5	7	18	14
6	10	19	8
7	13	20	4
8	14	21	2
9	29	22	10
10	22	23	4
11	21	24	4
12	18	25	3
13	28	26	4
14	16	27	6
15	19	28	1
16	10		
		Общее число женщин	283

1.2 Построить графики следующих функций и сравнить их вид с полигами частот и гистограммами распределений предыдущего упражнения:

$$\begin{aligned}
 y &= y_0 e^{-\frac{1}{2}x^2}, & -\infty \leq x \leq \infty, \\
 y &= y_0 e^{-x}, & -\infty \leq x \leq \infty, \\
 y &= y_0 x^{-\gamma}, & \gamma > 1, \quad 0 \leq x < \infty, \\
 y &= \frac{y_0}{(1+x^2)^n}, & n > 0, \quad -\infty \leq x \leq \infty, \\
 y &= y_0 (1-x)^a x^b, & a, b > 1, \quad 0 \leq x \leq 1, \\
 y &= y_0 e^{-x} x^\gamma, & \gamma > 1, \quad 0 \leq x < \infty, \\
 y &= y_0 (1-x^2)^a, & a < 0, \quad -1 \leq x \leq 1.
 \end{aligned}$$

1.3 Показать, что следующие распределения:

$$\begin{aligned}
 dF &= r_0 (1-r^2)^{\frac{1}{2}(n-4)} dr, & -1 \leq r \leq 1, \\
 dF &= \frac{t_0}{(1+t^2/v)^{\frac{1}{2}(v+1)}} dt, & -\infty \leq t \leq \infty, \\
 dF &= \frac{z_0 e^{v_1 z}}{(v_1 e^{2z} + v_2)^{\frac{1}{2}(v_1+v_2)}} dz, & -\infty \leq z \leq \infty,
 \end{aligned}$$

могут быть преобразованы к виду

$$dF = y_0(1-x)^{p-1} x^{q-1} dx, \quad 0 \leq x \leq 1,$$

и найти соответствующие преобразования. (Все эти распределения имеют важное значение в статистике. Распределение, к которому они сводятся, называется распределением I типа или В-распределением.)

1.4. Построить стереограммы или двумерные гистограммы следующих распределений:

Таблица 1.23

Распределение обследованных студентов Лондонского университета (1955 г.) в соответствии с (1) числом прочтенных и (2) числом просмотренных газет в течение дня

(Букер (H. S. Booker), неопубликованные данные.)

(2) Число только просмотренных газет

	0	1	2	3	4	5 или больше	Всего
0	77	75	19	10	1	2	184
1	179	136	65	20	15	2	417
2	86	70	45	18	1	3	223
3	17	21	13	3	4	—	58
4	4	2	2	2	—	—	10
5 или больше	2	2	—	2	—	—	6
Всего	365	306	144	55	21	7	898

1.5 Доказать, что для того, чтобы функция

$$f(x_1, x_2) = z_0 \exp \{Ax_1^2 + 2Hx_1x_2 + Bx_2^2\}, \quad -\infty \leq x_1, x_2 \leq \infty,$$

могла быть функцией плотности, необходимы следующие условия:

(а) $A \leq 0$,

(б) $B \leq 0$,

(в) $AB - H^2 \geq 0$.

Показать также, что если эти условия выполнены и интеграл от $f(x_1, x_2)$ по обоим переменным в пределах от $-\infty$ до ∞ равен единице, то

$$z_0 = \frac{1}{\pi} \left| \begin{array}{cc} -A & H \\ H & -B \end{array} \right|^{\frac{1}{2}}.$$

1.6 Показать, что двумерное распределение

$$dF = \frac{ky}{(1+x)^4} \exp\left(-\frac{y}{1+x}\right) dx dy, \quad 0 \leq x, y \leq \infty,$$

посредством преобразования $u=y/(1+x)$, $v=1/(1+x)$ сводится к распределению

$$dF = kue^{-u} du dv, \quad 0 \leq u \leq \infty, \quad 0 \leq v \leq 1,$$

и что, стало быть, $k=1$.

Таблица 1.24

Распределение 4912 эрширских коров в соответствии (1) с возрастом в годах,
(2) с удойностью за неделю

(по Точеру (J. F. Tocher), (1928), Biometrika 20B, 106)

(1) Возраст в годах

(2) Удойность за неделю (галлоны) (центральное значение интервала)																			Всего
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18			
8	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	1		
9	—	2	2	—	1	—	—	—	—	—	—	—	—	—	—	—	5		
10	3	5	1	1	3	—	—	—	—	—	—	—	—	—	—	—	13		
11	2	10	8	7	1	—	1	—	2	1	1	—	—	—	—	—	33		
12	2	25	17	9	5	4	4	2	1	1	—	—	—	1	—	—	71		
13	9	76	29	18	9	2	4	1	1	1	—	1	—	—	—	—	151		
14	11	76	57	38	23	9	7	6	4	2	3	—	—	—	—	—	236		
15	11	115	79	43	34	24	11	8	4	5	1	2	1	—	1	—	339		
16	15	149	119	74	59	23	23	16	9	7	4	—	—	—	1	—	499		
17	16	148	131	94	58	34	32	15	12	6	5	—	1	—	—	—	552		
18	11	146	132	83	73	49	39	22	17	6	5	1	1	—	—	—	585		
19	10	117	112	113	87	51	35	33	11	10	2	3	1	—	—	1	586		
20	8	97	107	79	69	51	25	30	13	10	3	3	—	—	1	—	496		
21	3	63	93	88	70	49	31	29	9	7	4	—	1	—	1	—	448		
22	5	42	63	49	45	32	14	18	10	3	1	2	—	—	—	—	284		
23	1	19	33	38	38	27	17	17	12	7	1	2	2	—	—	—	214		
24	2	20	23	34	27	19	13	9	3	2	1	—	—	—	—	—	153		
25	3	10	15	22	17	20	8	10	3	4	—	—	—	—	—	—	112		
26	—	7	13	7	4	15	2	4	2	3	1	—	—	—	—	—	58		
27	—	2	7	9	5	5	4	2	—	—	—	—	—	1	—	—	35		
28	—	—	2	1	4	2	1	1	2	—	—	—	—	—	—	—	13		
29	—	—	2	2	4	1	3	—	3	—	—	—	—	—	—	—	15		
30	—	—	—	—	—	2	—	—	2	—	—	—	—	—	—	—	4		
31	—	—	2	1	—	—	—	—	—	—	—	—	—	—	—	—	5		
32	—	—	—	2	—	—	—	—	—	—	—	—	—	—	—	—	2		
33	—	—	—	—	—	—	—	—	—	—	1	—	—	—	—	—	1		
34	—	—	—	—	—	—	—	—	1	—	—	—	—	—	—	—	1		
Всего	112	1129	1047	812	636	419	276	223	122	75	32	15	7	2	4	1	4912		

1.7 Показать, что в двумерном распределении

$$dF = \frac{k(1+x+y)}{(1+x)^4(1+y)^4} dx dy, \quad 0 \leq x, y \leq \infty,$$

k равно $9/2$ и что маргинальное распределение x есть

$$dF = \frac{3}{4} \cdot \frac{2x+3}{(1+x)^4} dx, \quad 0 \leq x \leq \infty.$$

1.8 Построить график функции плотности распределения

$$dF = k \operatorname{sech}^n x dx, \quad -\infty \leq x \leq \infty,$$

и показать, что

$$k = \frac{\Gamma\left\{\frac{1}{2}(n+1)\right\}}{\Gamma\left(\frac{1}{2}n\right)\sqrt{\pi}}.$$

1.9. Три варианты независимы и имеют одинаковое распределение

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx, \quad -\infty \leq x \leq \infty.$$

Переходя к новым вариантам посредством преобразования

$$\begin{aligned} y_1 &= (x_1 - x_2)/\sqrt{2}, \\ y_2 &= (x_1 + x_2 - 2x_3)/\sqrt{6}, \\ y_3 &= (x_1 + x_2 + x_3)/\sqrt{3}, \end{aligned}$$

показать, что игреки тоже взаимно независимы и имеют распределение того же вида.

1.10 Применить к вариантам предыдущего упражнения сферическое преобразование

$$\begin{aligned} x_1 &= r \cos \theta_1 \cos \theta_2, \\ x_2 &= r \cos \theta_1 \sin \theta_2, \\ x_3 &= r \sin \theta_1. \end{aligned}$$

Показать, что варианты r , θ_1 и θ_2 взаимно независимы и что распределение первой из них есть

$$dF = \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}r^2} r^2 dr, \quad 0 \leq r \leq \infty.$$

1.11 Показать, что маргинальное распределение x для двумерного распределения

$$dF = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right\} dx dy,$$

$-\infty \leq x, y \leq \infty$, имеет вид

$$dF = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_1^2}\right) dx, \quad -\infty \leq x \leq \infty.$$

Сделать набросок совместного распределения x и y и рассмотреть предельный случай при $\rho \rightarrow 1$.

1.12 Показать, что для двумерного распределения

$$\begin{aligned} dF &= \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \times \\ &\times \frac{n-1}{n-2} \frac{dx dy}{\left[1 + \frac{1}{2(n-2)(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right]^n}, \quad -\infty \leq x, y \leq \infty, \end{aligned}$$

маргинальное распределение x имеет вид

$$dF = \frac{1}{\sigma_1 \sqrt{(2\pi)(n-2)}} \frac{\Gamma\left(n - \frac{1}{2}\right)}{\Gamma(n-1)} \frac{1}{\left(1 + \frac{x^2}{2(n-2)\sigma_1^2}\right)^{n-\frac{1}{2}}} dx, \quad -\infty \leq x \leq \infty.$$

(К. Пирсон, (1923), *Biometrika* 15, 231.)

1.13 Пуассоновским называется распределение, для которого частота значения $x=r$ ($r=0, 1, 2, \dots$) равна $e^{-\lambda} \lambda^r / r!$. Показать, что производящая функция этого распределения есть

$$P(t) = \exp\{\lambda(t-1)\}.$$

1.14. Дискретное распределение, для которого частота значения $x=r$ ($r=0, 1, 2, \dots$) равна

$$f_r = \left(1 + \frac{m}{k}\right)^{-k} \frac{\Gamma(k+r)}{r! \Gamma(k)} \left(\frac{m}{m+k}\right)^r$$

(k — целое), называется отрицательным биномиальным распределением. Показать, что

$$P(t) = \left\{1 + \frac{m}{k}(1-t)\right\}^{-k}$$

есть производящая функция этого распределения.

1.15 Распределение варианты сосредоточено на интервале $0 \leq x \leq \infty$ и имеет вид $dF = f(x)dx$. Написать распределение этой варианты в интервале $x_0 \leq x \leq \infty$ ($x_0 > 0$). Показать, что если оно имеет тот же вид, что и исходное распределение, то $f(x) = ke^{-kx}$. Можно предположить, что функция $f(x)$ дифференцируема.

1.16 Задано n -мерное распределение

$$dF = k \exp\left\{-\sum_{i,j=1}^n a_{ij}x_i x_j\right\} dx_1 \dots dx_n, \quad -\infty \leq x_1, \dots, x_n \leq \infty,$$

где $a_{ij} = a_{ji}$. Показать, что показатель экспоненты здесь может быть записан в форме

$$-a_{nn}\left(x_n + \frac{1}{a_{nn}} \sum_{i=1}^{n-1} a_{ni}x_i\right)^2 - \sum_{i,j=1}^{n-1} b_{ij}x_i x_j,$$

где $b_{ij} = a_{ij} - \frac{a_{in}a_{jn}}{a_{nn}}$, $i, j = 1, 2, \dots, n-1$. Пользуясь этим фактом, найти линейное преобразование исходных переменных во взаимно независимые переменные, имеющие каждое распределение

$$dF = \frac{1}{\sigma \sqrt{2\pi}} e^{-u^2/2\sigma^2} du, \quad -\infty \leq u \leq \infty.$$

1.17 Пусть $f(x, y)$ — дифференцируемая функция плотности двумерного распределения. Доказать, что линейное преобразование, переводящее варианты в независимые, существует тогда и только тогда, когда

$$\left(A \frac{\partial^2}{\partial x^2} + 2H \frac{\partial^2}{\partial x \partial y} + B \frac{\partial^2}{\partial y^2}\right) \log f = 0,$$

где A , H и B — постоянные. Показать далее, что для распределения из упражнения 1.11 существует бесконечно много линейных преобразований, переводящих варианты в независимые.

1.18 Три варианты x , y и z неотрицательны, удовлетворяют неравенству $x+y+z \leq 1$ и имеют распределение

$$dF = kx^{l-1}y^{m-1}z^{n-1} dx dy dz.$$

Показать, что

$$k = \frac{\Gamma(l+m+n+1)}{\Gamma(l)\Gamma(m)\Gamma(n)}.$$

1.19. n вариант имеют распределение

$$dF = kx_1^{l_1-1}x_2^{l_2-1} \dots x_n^{l_n-1} g(x_1+x_2+\dots+x_n) dx_1 \dots dx_n.$$

Все они неотрицательны и удовлетворяют неравенству $\sum_{i=1}^n x_i \leq 1$. Показать, что

$$\frac{1}{k} = \frac{\Gamma(l_1)\Gamma(l_2)\dots\Gamma(l_n)}{\Gamma(l_1+l_2+\dots+l_n)} \int_0^1 v^{L-1} g(v) dv,$$

где $L = \sum_{i=1}^n l_i$.

1.20. Показать, что если функцию $g(x)$ можно разложить в равномерно сходящийся ряд по степеням x , то

$$\int_{-\infty}^{\infty} \frac{\sin^2 x}{x^2} g(\sin^2 x) dx = \int_0^{\pi} g(\sin^2 x) dx.$$

Сделать набросок распределения

$$dF = \frac{k \sin^2 x \cos^2 x}{x^2 (1 + \sin^2 x)} dx, \quad -\infty \leq x \leq \infty,$$

и показать, что $1/k = \pi(\sqrt{2}-1)$.

1.21. Каждая из независимых вариант x_1 , x_2 имеет распределение

$$dF = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx, \quad -\infty \leq x \leq \infty.$$

Показать, что варианты

$$y_1 = \exp\left\{-\frac{1}{2}(x_1^2 + x_2^2)\right\},$$

$$y_2 = \frac{1}{2\pi} \arctg\left(\frac{x_1}{x_2}\right)$$

независимы и имеют прямоугольное распределение на $(0, 1)$.

(Бокс и Мюллер (Box and Muller), 1958, Ann. Math. Statist. 29, 610.)

1.22 $F_1(x_1)$, $F_2(x_2)$ — одномерные ф. р. Показать, что функция

$$F(x_1, x_2) = F_1(x_1)F_2(x_2)[1 + \theta\{1 - F_1(x_1)\}\{1 - F_2(x_2)\}],$$

где θ — постоянная ($-1 \leq \theta \leq 1$), есть двумерная ф. р.

(Ср. Гамбел (E. J. Gumbel), 1960, J. Amer. Statist. Ass. 55, 698.)

МЕРЫ РАСПОЛОЖЕНИЯ И РАССЕЯНИЯ

2.1 Из материала главы 1 видно, что частотные распределения, возникающие в статистической практике, чрезвычайно разнообразны. Некоторые из них сосредоточены на конечных интервалах, другие — на бесконечных. Одни симметричны, другие же, наоборот, имеют заметно асимметричную форму. Некоторые распределения обладают только одним максимумом, другие же — несколькими. Однако среди этого разнообразия форм распределений можно выделить четыре общих типа: (а) симметричные распределения с одним максимумом (например, распределение в таблице 1.7); (б) асимметричные или косые распределения с одним максимумом (например, распределения в таблицах 1.8 и 1.9); (в) крайне асимметричные, или J-образные, распределения (см., к примеру, таблицу 1.2); (г) U-образные распределения, примером которых служит распределение в таблице 1.11. Если желать сделать эту классификацию исчерпывающей, надо было бы в пятый класс отнести все те распределения, которые не вошли в четыре первых класса.

Распределения с единственным максимумом в дальнейшем будут называться «унимодальными» (одновершинными).

2.2 В статистической практике часто возникает необходимость сравнения двух распределений. Если одно из них унимодально, а другое, скажем, J-образно или мультимодально (многовершинно), то ясно, что провести хотя бы в общих чертах их сравнение, ограничиваясь рассмотрением небольшого числа их характерных признаков, будет затруднительно. В этом случае, возможно, придется полностью специфицировать каждое из распределений. Но если оба распределения принадлежат к одному и тому же типу (а именно такие случаи чаще всего возникают), то мы можем сделать вполне удовлетворительное сравнение, ограничиваясь сопоставлением их некоторых главных характеристик. Например, если оба распределения унимодальны, то весьма правдоподобно, что вполне достаточно выявить: (а) расположение их некоторых главных значений, таких как, например, максимумы, дающие представление о расположении

распределения; (б) степень разброса распределений около этих значений и (в) степень отклонения распределений от симметричной формы, так называемую асимметрию.

Та же проблема возникает, конечно, и тогда, когда распределения задаются математическими функциями. Если, например, мы имеем два распределения типа

$$dF = k \exp \left\{ -\frac{(x-m)^2}{2\sigma^2} \right\} dx, \quad -\infty \leq x \leq \infty,$$

каждое из которых симметрично около точки $x=m$, то полное их сопоставление может быть сделано путем сравнения констант m и σ , определяющих распределения указанного типа. Такие константы называются параметрами распределения *).

Меры расположения: арифметическое среднее

2.3 Существуют три группы мер, применяемых для описания характера расположения распределений: среднее (арифметическое, геометрическое и гармоническое), медиана и мода. Мы рассмотрим их по очереди.

Арифметическое среднее — одна из наиболее употребительных статистических мер. Оно употреблялось еще задолго до появления статистики как науки. Если функция плотности равна $f(x)$, то арифметическое среднее $\mu'_1 = \mu'_1(a)$ относительно точки $x=a$ определяется формулой **)

$$\mu'_1(a) = \int_{-\infty}^{\infty} (x-a)f(x) dx = \int_{-\infty}^{\infty} (x-a)dF. \quad (2.1)$$

Последний интеграл надо понимать как интеграл Стильтьеса, и, следовательно, в дискретном случае он превращается в сумму. Частным случаем этой формулы является арифметическое среднее из конечного числа значений.

В соответствии с нашими обычными соглашениями в формуле (2.1) частота $f(x)$ предполагается нормированной, т. е.

*) В прошлом существовала некоторая разница в понимании параметров распределения. Мы будем употреблять этот термин только для обозначения величин таких, например, как m и σ , которые явно входят в формулу функции распределения. Поэтому среднее и дисперсия в нашем смысле, вообще говоря, не являются параметрами, хотя в рассмотренном примере m равно среднему, а σ^2 — дисперсии распределения.

**) Часто в дальнейшем под средним μ'_1 (без указания точки, относительно которой оно берется) понимается среднее относительно нуля. (Прим. перев.)

$\int_{-\infty}^{\infty} f(x) dx = 1$. Если фактические частоты обозначить $g(x)$,

а полная частота равна N , т. е. $\int_{-\infty}^{\infty} g(x) dx = N$, то в непрерывном случае

$$\mu'_1 = \frac{1}{N} \int_{-\infty}^{\infty} (x - a) g(x) dx$$

и в дискретном случае

$$\mu'_1 = \frac{1}{N} \sum_{j=-\infty}^{\infty} (x_j - a) g(x_j).$$

Значение арифметического среднего, таким образом, зависит от значения a , т. е. от точки, относительно которой производятся измерения. Вообще говоря, для распределений, задаваемых формулами, интеграл (2.1) может и не сходиться. В таком случае говорят, что арифметическое среднее не существует.

2.4 Вычисление арифметического среднего осуществляется довольно просто для распределений, задаваемых численно (т. е. таких, для которых частотные распределения задаются в форме таблиц, подобных рассмотренным в главе 1). Если генеральная совокупность содержит относительно малое число значений, то мы просто их суммируем и делим на их общее число N . Если же эти значения представлены в форме таблиц, содержащих и частоты значений, то предпочтительнее воспользоваться более формальной процедурой, хотя, в принципе, той же самой. В чем состоит эта процедура, наглядно иллюстрируется в следующем примере.

Пример 2.1

Вычислить арифметическое среднее значение роста мужчин, используя данные таблицы 1.7.

Прежде чем переходить к подсчетам, заметим, что если b есть некоторое другое значение, отличное от a , то

$$\begin{aligned} \mu'_1(a) &= \int_{-\infty}^{\infty} (x - a) dF = \int_{-\infty}^{\infty} (x - b) dF + \int_{-\infty}^{\infty} (b - a) dF = \\ &= \mu'_1(b) + (b - a). \end{aligned} \quad (2.2)$$

Другими словами, мы можем найти среднее значение относительно произвольной точки, если только известно среднее отно-

сительно какой-нибудь одной точки. Поэтому при вычислении среднего относительно нужной точки можно его вычислять сначала около произвольной точки, которую следует выбрать так, чтобы упростить выкладки, а затем уже простым пересчетом найти нужное среднее. Для облегчения вычислений эту вспомогательную точку рекомендуется выбирать где-нибудь вблизи максимума частотного распределения.

Некоторое затруднение возникает в случае группированных данных таких, как в этом примере. Мы не знаем точно индивидуальных значений варианты внутри любого данного класс-интервала. Поэтому мы условимся считать эти значения совпадающими с центром соответствующего интервала. Поправки на искажения, возникающие в результате такой процедуры, будут рассмотрены в главе 3. При этом в действительности для арифметического среднего никаких поправок не потребуется, если только функция плотности на обоих концах распределения имеет касание оси абсцисс достаточно высокого порядка.

Таблица 2.1

Вычисление арифметического среднего для распределения, заданного таблицей 1.7

(1) Рост в дюймах	(2) Частота f	(3) Отклонения $\xi = x - a$	(4) Произведение ξf
57—	2	—10	20
58—	4	—9	36
59—	14	—8	112
60—	41	—7	287
61—	83	—6	498
62—	169	—5	845
63—	394	—4	1 586
64—	669	—3	2 007
65—	990	—2	1 980
66—	1 223	—1	1 223
67—	1 329	0	—8 584
68—	1 230	+1	1 230
69—	1 063	+2	2 126
70—	646	+3	1 938
71—	392	+4	1 568
72—	202	+5	1 010
73—	79	+6	474
74—	32	+7	224
75—	16	+8	128
76—	5	+9	45
77—	2	+10	20
Суммы	8 585	—	+8 763

В рассматриваемом нами случае мы можем выбрать в качестве произвольного начала отсчета a середину интервала 67 дюймов, т. е. точку $67 \frac{7}{16}$ дюймов, и подсчитаем отклонения $\xi (=x - a)$. Второй столбец в таблице 2.1 дает значения частот, третий столбец — значения ξ и четвертый — значения ξf . С учетом знака отклонений мы находим

$$\sum (\xi f) = 8763 - 8584 = 179.$$

Следовательно, среднее значение (в дюймах) относительно точки $x=0$ равно $67 \frac{7}{16} + \frac{179}{8585} = 67,46$.

Пример 2.2

Для распределений, задаваемых математическими формулами, нахождение среднего сводится к вычислению интеграла (2.1), когда он существует. Например, для распределения

$$dF = \frac{1}{B(p, q)} (1-x)^{p-1} x^{q-1} dx, \quad 0 \leq x \leq 1, \quad p, q > 0,$$

мы имеем

$$\begin{aligned} \mu'_1 &= \frac{1}{B(p, q)} \int_0^1 (1-x)^{p-1} x^q dx = \\ &= \frac{B(p, q+1)}{B(p, q)} = \frac{\Gamma(p) \Gamma(q+1)}{\Gamma(p+q+1)} \frac{\Gamma(p+q)}{\Gamma(p) \Gamma(q)} = \frac{q}{p+q}. \end{aligned}$$

Пример 2.3

Если распределение задается производящей функцией, то среднее может быть вычислено следующим образом.

Пусть

$$P(t) = \sum_{j=0}^{\infty} f_j t^j. \quad (2.3)$$

Тогда

$$\left[\frac{dP}{dt} \right]_{t=1} = \left[\sum f_j t^{j-1} j \right]_{t=1} = \sum f_j j = \mu'_1(0). \quad (2.4)$$

Например, для биномиального распределения

$$P(t) = \{pt + (1-p)\}^n$$

и

$$\left[\frac{dP}{dt} \right]_{t=1} = np \{p + (1-p)\}^n = np,$$

что и дает среднее относительно нуля.

Геометрическое и гармоническое средние

2.5 Эти виды средних употребляются в элементарной статистике, хотя и не играют важной роли в высшей теории.

Геометрическое среднее из N значений варианты равно корню N -й степени из их произведения. В том случае, когда некоторые из этих значений равны нулю или отрицательны, геометрическое среднее не определяется. Если частоты, соответствующие значениям x , равны $f(x)$, то

$$\left. \begin{aligned} G &= \prod_{j=-\infty}^{\infty} (x_j^{f_j}) \\ \log G &= \sum_{j=-\infty}^{\infty} f_j \log x_j. \end{aligned} \right\} \quad (2.5)$$

Для ненормированных частот $g(x)$, в сумме равных N , имеем

$$\left. \begin{aligned} G &= \Pi (x_j^{g_j})^{\frac{1}{N}}, \\ \log G &= \frac{1}{N} \sum g_j \log x_j. \end{aligned} \right\} \quad (2.6)$$

Гармоническое среднее N значений варианты есть обратная величина арифметического среднего для обратных величин этих значений. В обычных обозначениях

$$\frac{1}{H} = \int_{-\infty}^{\infty} \frac{dF}{x} = \int_{-\infty}^{\infty} \frac{f(x) dx}{x}, \quad (2.7)$$

или для ненормированных частот

$$\frac{1}{H} = \frac{1}{N} \int_{-\infty}^{\infty} \frac{g(x) dx}{x} \quad (2.8)$$

в предположении, конечно, что интеграл существует.

Пример 2.4

Найдем геометрическое и гармоническое средние для распределения

$$dF = \frac{1}{B(p, q)} (1-x)^{p-1} x^{q-1} dx, \quad 0 \leq x \leq 1, \quad p, q > 0.$$

Имеем

$$\log G = \frac{1}{B(p, q)} \int_0^1 (1-x)^{p-1} x^{q-1} \log x dx.$$

Согласно определению В-функции

$$\int_0^1 (1-x)^{p-1} x^{q-1} dx = B(p, q),$$

откуда, дифференцируя обе части равенства по q (что законно в силу равномерной сходимости интеграла и существования выражения, получаемого после дифференцирования), мы получим

$$\int_0^1 (1-x)^{p-1} x^{q-1} \log x dx = \frac{\partial}{\partial q} B(p, q).$$

Поэтому

$$\begin{aligned} \log G &= \frac{1}{B(p, q)} \frac{\partial}{\partial q} B(p, q) = \frac{\partial}{\partial q} \log \frac{\Gamma(p) \Gamma(q)}{\Gamma(p+q)} = \\ &= \frac{\partial}{\partial q} \{\log \Gamma(q) - \log \Gamma(p+q)\}. \end{aligned}$$

Гармоническое среднее дается выражением

$$\frac{1}{H} = \frac{1}{B(p, q)} \int_0^1 (1-x)^{p-1} x^{q-2} dx.$$

Однако для того чтобы это выражение было конечным, надо потребовать $q > 1$. В этом случае

$$\frac{1}{H} = \frac{B(p, q-1)}{B(p, q)} = \frac{\Gamma(q-1)}{\Gamma(p+q-1)} \frac{\Gamma(p+q)}{\Gamma(q)} = \frac{p+q-1}{q-1},$$

так что

$$H = \frac{q-1}{p+q-1}.$$

Заметим, что арифметическое среднее $\frac{q}{p+q}$ больше, чем гармоническое. В самом деле,

$$\frac{q}{p+q} = 1 - \frac{p}{p+q}, \quad \frac{q-1}{p+q-1} = 1 - \frac{p}{p+q-1},$$

откуда требуемое соотношение $\mu'_1 > H$ непосредственно следует из того, что

$$\frac{p}{p+q-1} > \frac{p}{p+q}.$$

2.6 Вообще можно показать, что для распределений неотрицательных вариантов

$$H \leq G \leq \mu'_1. \quad (2.9)$$

В самом деле, рассмотрим величину

$$A(t) = \left| \frac{1}{N} (x_1^t + x_2^t + \dots + x_N^t) \right|^{1/t},$$

где все x — действительные числа. Покажем, что $A(t)$ — возрастающая функция от t , т. е. $A(t_1) > A(t_2)$ при $t_1 > t_2$. В тривиальном случае, когда все значения x равны, эти неравенства превращаются в равенства.

Заметим, что при $t=1$ функция $A(t)$ есть не что иное, как арифметическое среднее; когда $t=-1$, мы имеем гармоническое среднее, а при $t \rightarrow 0$ получаем геометрическое среднее, поскольку

$$\lim_{t \rightarrow 0} \log A = \lim_{t \rightarrow 0} (\log \sum x^t / N) / t = \lim_{t \rightarrow 0} (\sum x^t \log x) / N = \sum \log x / N.$$

Положим

$$I = \frac{1}{N} \sum x^t, \quad F = t^2 \frac{d}{dt} \log A.$$

Тогда

$$F = t^2 \frac{d}{dt} \left(\frac{\log I}{t} \right) = \frac{t}{I} \frac{dI}{dt} - \log I$$

и, следовательно,

$$\frac{dF}{dt} = \frac{t}{I^2} \left\{ I \frac{d^2 I}{dt^2} - \left(\frac{dI}{dt} \right)^2 \right\}. \quad (2.10)$$

Выражение в скобках в формуле (2.10), совпадающее с величиной

$$\left(\frac{1}{N} \sum x^t \right) \left(\frac{1}{N} \sum x^t \log^2 x \right) - \left(\frac{1}{N} \sum x^t \log x \right)^2,$$

в силу неравенства Коши — Буняковского неотрицательно. Следовательно, знак производной dF/dt совпадает со знаком t , и поэтому функция $F(t)$ имеет минимум в точке $t=0$. Но при $t=0$ $F=0$, значит, функция F неотрицательна. Поэтому производная $d \log A / dt$ также неотрицательна, и в силу того, что $A \geq 0$, и $dA/dt \geq 0$.

Отсюда следует, что функция $A(t)$ — неубывающая. На самом же деле она является возрастающей, за исключением того случая, когда все значения x равны между собой.

Неравенства (2.9) вытекают теперь очевидным образом из сказанного выше.

Отметим, что при их доказательстве мы ограничились для простоты случаем дискретных вариантов. Но нетрудно видеть, что проведенный анализ остается справедливым, если сумму заменить интегралом Стильтьеса, так что соотношения (2.9) остаются в силе и для общего случая.

В дальнейшем всюду арифметическое среднее будет называться «среднее».

2.7 Неравенство Коши — Буняковского *) в математической статистике используется довольно часто. Поэтому представляется полезным в связи с этим неравенством сделать некоторые замечания. Неравенство, найденное Коши в 1821 году, устанавливает, что если a_i, b_i ($i=1, \dots, n$) — действительные числа, то

$$\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) > \left(\sum_{i=1}^n a_i b_i\right)^2,$$

за исключением того случая, когда все a и b пропорциональны. В этом последнем случае строгое неравенство превращается в равенство. Соответствующий результат для интегралов был получен Буняковским в 1859 и Шварцем в 1885 году. Из неравенства Коши — Буняковского можно получить ряд других полезных неравенств, например неравенство Гельдера (1889 г.); если $a_i, b_i, \dots, l_i$ ($i=1, \dots, n$) неотрицательны, $\alpha, \beta, \dots, \lambda$ положительны и удовлетворяют условию $\alpha + \beta + \dots + \lambda = 1$, то

$$\sum a^\alpha b^\beta \dots l^\lambda < \sum a^\alpha \sum b^\beta \dots \sum l^\lambda,$$

за исключением того случая, когда по крайней мере одно из множеств (a), (b) и т. д. состоит из нулей или когда элементы одного из множеств пропорциональны элементам всех других. В этих двух случаях неравенство превращается в равенство. Ясное и исчерпывающее изложение затронутых вопросов читатель найдет в книге Г. Г. Харди, Дж. Е. Литтлвуд и Г. Полиа «Неравенства», М., ИЛ, 1948.

Медиана

2.8 Медиана есть такое значение варианты, которое разделяет все частоты на две равные части, иначе говоря, медиана μ_e — значение, удовлетворяющее соотношению

$$\int_{-\infty}^{\mu_e} f(x) dx = \int_{\mu_e}^{\infty} f(x) dx = \frac{1}{2}. \quad (2.11)$$

В этом определении существует некоторая неопределенность для случая дискретных распределений, но ее можно избежать, если принять следующее соглашение.

Если генеральная совокупность состоит из $(2N+1)$ членов, то под медианой понимается $(N+1)$ -е по величине значение варианты. Если же совокупность состоит из $2N$ членов, то условимся считать медиану равной среднему арифметическому между N -м и $(N+1)$ -м по величине значениями варианты.

*) В оригинале это неравенство называется неравенством Коши — Шварца. (Прим. перев.)

Для распределений, задаваемых численно по класс-интервалам, характерна обычная неопределенность вследствие группировки. Эту неопределенность можно преодолеть подобно тому, как это делается в следующем примере.

Пример 2.5

Найти значение медианы для распределения ростов, рассмотренного в примере 2.1.

Половина из 8585 наблюдений составляет	4292,5
Число наблюдений, меньших или равных $66\frac{15}{16}$ дюймов	3589
Разность этих чисел	703,5
Частота в следующем интервале	1329
Следовательно, медиана равна (в дюймах)	

$$66\frac{15}{16} + \frac{703,5}{1329} = 67,47.$$

Среднее значение (см. пример 2.1) равно 67,46 дюйма, что практически совпадает с медианой.

Графический способ определения медианы будет рассмотрен в этой главе несколько позже (2.15).

2.9 С математической точки зрения неопределенности, возникающей в определении медианы, можно избежать следующим образом. Нетрудно установить, что для множества значений $x_1, \dots, x_N$ сумма $\sum |\xi - x_i|^p$, рассматриваемая как функция от ξ , имеет единственный минимум для некоторого ξ_p , если только $p > 1$. Если p стремится к единице, то ξ_p стремится к пределу, который, по определению, можно принять за медиану (см. Джексон (1921)).

Распространение понятия медианы на многомерный случай несколько сложно. Для двумерного распределения под медианой понимается пара чисел, каждое из которых есть медиана маргинального распределения [см. Холдейн (1948)].

Мода

2.10 Мода, или модальное значение, есть то значение варианты, при котором функция плотности достигает своего максимального значения. Если функция плотности непрерывна и дифференцируема, то мода удовлетворяет соотношениям

$$f'(x) = \frac{d}{dx} f(x) = 0, \quad f''(x) = \frac{d^2}{dx^2} f(x) < 0. \quad (2.12)$$

Если $f'(x) = 0$ и $f''(x) > 0$, то в точке x достигается минимум и такая точка иногда называется «антимодой».

Для распределений, задаваемых численно, а также для дискретных распределений точное определение моды часто оказывается затруднительным. По существу, понятие моды связано

только с распределениями, обладающими непрерывной плотностью. Так, если частотное распределение сосредоточено в конечном числе точек и значения частот в этих точках равны, то мода в смысле определения, данного выше, не существует. Но, однако, в тех случаях, когда число точек, в которых сосредоточено частотное распределение, достаточно велико, то при группировании обычно будет существовать интервал, в котором частота достигает максимума, и тогда можно считать, что значение моды принадлежит этому интервалу. Вообще же у частотного распределения возможно наличие многих максимумов. В этом случае распределение называется мультимодальным. Например, для распределения роста мужчин, заданного таблицей 1.7, мода может рассматриваться как значение, лежащее где-то в интервале 67 дюймов. Чтобы оценить моду более точно, надо данное распределение аппроксимировать непрерывной кривой и взять моду этой кривой. Этот процесс «подгонки» будет подробно рассмотрен в главе 6.

2.11 Для симметричных распределений среднее, медиана и мода (а для случая U-образных распределений антимода) совпадают. Для асимметричных распределений они вообще различны. Для унимодальных распределений, не слишком отличающихся от симметричных, между этими тремя величинами существует интересная связь, найденная эмпирически. Она состоит в том, что

$$\text{среднее} - \text{мода} = 3 (\text{среднее} - \text{медиана}). \quad (2.13)$$

Математическое объяснение этого соотношения дается в упражнении 6.20.

Если термин «среднее значение» заменить термином «математическое ожидание», то можно подметить полезное мнемоническое правило: величины математического ожидания, медианы и моды располагаются в том же (или в обратном) порядке, в каком соответствующие термины расположены в словаре. И точно так же как в словаре, медиана располагается ближе к математическому ожиданию, чем к моде.

В элементарной статистической теории медиана и мода призваны служить мерами расположения. Они легко интерпретируются в привычных терминах (медиана — точка, являющаяся серединой частотного распределения, мода — наиболее часто встречающееся значение), и к тому же в случае распределений, задаваемых численно, медиана более легко поддается определению, чем среднее значение. То, что арифметическое среднее в высшей теории статистики играет более значительную роль, объясняется его превосходными математическими достоинствами и некоторыми выборочными свойствами. Но зато медиана

имеет другие преимущества; так, например, она менее зависит от формы распределения, чем среднее.

2.12 Упомянем также о другой, малоупотребительной мере расположения, о *центре*. Эта величина определяется как значение среднего арифметического между крайними точками распределения, если таковые существуют. Поскольку это определение зависит от наличия у распределения крайних точек, то для распределений, простирающихся до бесконечности, оно непригодно. Однако центр, так же как и размах (см. ниже, 2.17), имеет некоторое распространение в выборочной теории и легко поддается вычислению. Центр, определяемый по выборочным значениям, иногда называют «*срединой размаха*»^{*)}.

Квантили

2.13 С целью более точного указания расположения кривой частотного распределения можно обобщить понятие медианы, введя несколько подобных ей характеристик. Можно, например, для этого найти три значения варианты, которые разбивают все распределение на четыре равные части. Средняя из этих величин будет, очевидно, медианой, другие два значения соответственно называются нижними и верхними квантилями. Аналогично, можно отыскать 9 значений варианты, подразделяющих все распределение на десять равных частей — децили. Вообще можно найти $(n - 1)$ точек, разбивающих распределение на n равных частей. Эти точки называются квантилями. Очевидно, что указание квантилей для достаточно больших n , таких, скажем, как 10, дает очень хорошее представление о форме кривой частотного распределения. Даже только указания значения квантилей и медианы дают вполне хорошие сведения о характере распределения. Статистики указанных типов будут систематически изучаться в главе 14.

2.14 Нахождение квантилей для распределений, заданных численно, осуществляется точно так же, как и определение медианы. Неопределенность, возникающая при отыскании квантилей, может быть разрешена соглашениями, аналогичными указанным выше. Для распределений, задаваемых математическими формулами, j -й квантиль есть решение x уравнения

$$\frac{j}{n} = \int_{-\infty}^x dF. \quad (2.14)$$

Это решение без труда вычисляется с помощью интерполяции в тех случаях, когда интеграл от dF табулирован.

*) В оригинале «mid-range». (Прим. перев.)

Пример 2.6

Найти квартили для распределения по росту, рассмотренного в примере 2.1.

Четверть полной частоты равна $8585/4 =$	2146,25	
До интервала 65 — имеется	990	»
Разность	1376	членов
В интервале 65 — имеется	770,45	»
Поэтому нижний квартиль равен		
$64 \frac{15}{16} + \frac{770,25}{990} =$	65,71	дюйма
Верхний квартиль равен	69,21	»
Ранее мы нашли (пример 2.5), что медиана μ_e равна	67,47	»

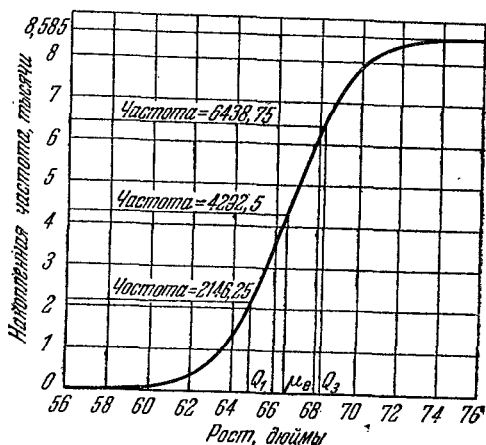


Рис. 2.1. Функция распределения для данных таблицы 1.7. (На этом рисунке данные соответствуют таблице; например, накопленные частоты, отвечающие точке 64 дюйма, представляют собой сумму частот всех интервалов вплоть до 64 — включительно, т. е. до значения $64 \frac{15}{16}$ дюйма.)

Обозначая нижний и верхний квартили Q_1 и Q_3 , мы видим, что

$$Q_1 - \mu_e = -1,76 \text{ дюйма,}$$

$$Q_3 - \mu_e = 1,74 \text{ дюйма,}$$

так что медиана лежит почти посредине между квартилями — указание на симметрию распределения.

2.15 Квартили можно легко определять по графику функции распределения. На рис. 2.1 представлена кривая накопленных частот для данных таблицы 1.7. Чтобы найти, например, медиану, мы определяем ординату, соответствующую абсциссе $\frac{1}{2}N$. Положения квартилей и медианы показаны на рис. 2.1, и читателю рекомендуем сравнить значения, найденные по графику на глаз, с теми, которые получены в примере 2.6.

Меры рассеяния

2.16 Мы сейчас переходим к рассмотрению величин, характеризующих степень рассеяния распределения. Они состоят из трех групп:

(а) Мер, являющихся расстояниями между точками, в которых частотное распределение имеет некоторые характерные особенности. К ним относятся такие меры, как размах, интердецильная ширина или интерквартильная ширина.

(б) Мер, образованных из отклонений каждого члена генеральной совокупности от некоторого центрального значения. К их числу относятся среднее отклонение от среднего значения, среднее отклонение от медианы и среднее квадратичное отклонение.

(в) Мер, составленных из отклонений между всеми членами генеральной совокупности; к числу таких мер относится, например, средняя разность.

В высшей теории наибольшую роль играет среднее квадратичное отклонение, но и все указанные выше меры заслуживают некоторого внимания.

Размах и интерквартильные разности

2.17 Размах распределения для ограниченной варианты определяется как разность между наибольшим и наименьшим ее значениями. Как описательная мера рассеяния распределения размах употребляется очень редко. Приблизительное знание крайних значений интервала изменения варианты, очевидно, мало, что говорит о поведении распределения внутри самого интервала. Для неограниченных же распределений это понятие вообще неприменимо.

Более полезные меры, дающие хотя бы приблизительное представление о рассеянии распределения, можно получить, используя квантили. Среди таких мер наиболее употребительны интерквартильная и интердецильная широты. Интерквартильная ширина определяется как расстояние между верхним и нижним квартилями. Тем самым определяемый этими квартилями интер-

вал содержит половину полной частоты. Интердецильная ширина есть расстояние между девятым и первым децилями. Обе эти меры очевидно дают некоторое представление о «размазанности» распределения и легко поддаются вычислению. В силу этих обстоятельств они довольно часто используются в элементарной описательной статистике. Что же касается высшей теории, то они мало употребляются в силу того, что в выборочной теории они трудно поддаются математическому анализу.

Средние отклонения (первые абсолютные моменты)

2.18 Степень разброса в генеральной совокупности можно измерять величинами, образованными отклонениями от среднего значения $\mu'_1 = \mu'_1(0)$. Легко видеть, что сумма этих отклонений (с учетом знака) равна нулю. Поэтому естественно ввести величину

$$\delta_1 = \int_{-\infty}^{\infty} |x - \mu'_1| dF, \quad (2.15)$$

где отклонения от среднего берутся по модулю, и считать δ_1 мерой рассеяния. Эту величину мы будем называть средним отклонением от среднего. Аналогично для медианы можем ввести

$$\delta_2 = \int_{-\infty}^{\infty} |x - \mu_e| dF \quad (2.16)$$

и называть δ_2 средним отклонением от медианы.

В дальнейшем под «средним отклонением» всегда будет пониматься среднее отклонение от среднего значения. В элементарной теории обе введенные меры обладают тем достоинством, что они довольно просто вычисляются, но, однако, практически они исключаются из рассмотрения в высшей математической статистике в силу трудности их математической обработки в выборочной теории.

Стандартное отклонение

2.19 Мы видели, что среднее относительно произвольной точки a дается формулой

$$\mu'_1 = \mu'_1(a) = \int_{-\infty}^{\infty} (x - a) dF.$$

По аналогии с терминологией статистики эту величину можно назвать первым моментом. Введем теперь второй момент

относительно точки a

$$\mu'_2 = \int_{-\infty}^{\infty} (x - a)^2 dF. \quad (2.17)$$

Второй момент относительно среднего будем обозначать без штриха

$$\mu_2 = \int_{-\infty}^{\infty} (x - \mu'_1)^2 dF \quad (2.18)$$

и называть дисперсией. Положительное значение квадратного корня из дисперсии называется стандартным (или средним квадратичным) отклонением и обычно обозначается буквой σ . Так что имеем

$$\sigma = |\sqrt{\mu_2}|. \quad (2.19)$$

Таким образом, дисперсия есть среднее из квадратов отклонений от среднего значения. Может показаться несколько искусственным способ определения стандартного отклонения посредством операций возведения в квадрат и извлечения квадратного корня из суммарного результата. Но этот способ определения оправдывается теми преимуществами, которыми стандартное отклонение обладает в выборочной теории по сравнению, например, со средним отклонением.

Для вычисления дисперсии и стандартного отклонения можно использовать методы, аналогичные употребляемым для подсчета среднего значения. В частности, если b есть некоторое произвольное значение, то

$$\begin{aligned} \mu'_2(a) &= \int_{-\infty}^{\infty} (x - a)^2 dF = \int_{-\infty}^{\infty} [(x - b)^2 + 2(b - a)(x - b) + \\ &+ (b - a)^2] dF = \mu'_2(b) + 2(b - a)\mu'_1(b) + (b - a)^2. \end{aligned} \quad (2.20)$$

Поэтому, если b есть среднее значение, то

$$\mu'_2 = \mu_2 + (\mu'_1 - a)^2$$

или

$$\mu_2 = \mu'_2 - (\mu'_1 - a)^2. \quad (2.21)$$

Тем самым дисперсию легко найти, зная второй момент относительно произвольной точки, которую следует выбирать из соображений упрощения вычислений.

Пример 2.7

Найти среднее и стандартное отклонения для распределения роста мужчин, рассмотренного в примере 2.1 (таблица 1.7).

Для среднего отклонения в случае группированного распределения сумму отклонений удобно подсчитывать сначала относительно центра того интервала группирования, в котором лежит среднее, а затем уже редуцировать к собственно среднему значению. Так вышло, что в таблице 2.1 среднее попадает в тот же самый интервал (67—), относительно центра которого производился подсчет среднего значения. Тем самым некоторые предварительные арифметические действия, необходимые нам для подсчета среднего отклонения, уже содержатся в указанной таблице. Сумма положительных отклонений равна 8763, отрицательных —8584. Без учета знака сумма отклонений относительно центра интервала 67 — равна 17 347, где единицей служит длина интервала группирования.

Рассматривая отклонения относительно среднего (а не относительно середины интервала, которому оно принадлежит), мы замечаем, что если число наблюдений, меньших среднего значения, равно N_1 , а больших среднего равно N_2 и $d = \mu'_1 - a$, то к сумме отклонений относительно центра интервала нужно прибавить $N_1 d$ и вычесть $N_2 d$. В рассматриваемом случае $d = 0,02$ (пример 2.1), $N_1 = 4918$, $N_2 = 3667$. Отсюда следует, что при подсчете среднего отклонения мы должны учесть значение, равное $(4918 - 3667)0,02 = 25$. Следовательно,

$$\delta_1 = \frac{17\,347 + 25}{8585} = 2,02 \text{ дюйма}^*).$$

При нахождении стандартного отклонения рекомендуется проводить вычисления по схеме, указанной в таблице 2.2.

Столбец (4) дает сумму $\sum \xi^2 f$, где f — истинная частота. Поэтому для второго момента, подсчитываемого относительно точки a , мы получаем

$$\mu'_2 = \frac{56\,809}{8585} = 6,6172.$$

В примере 2.1 было найдено, что

$$\mu'_1 - a = \frac{179}{8585} = 0,0209.$$

Тем самым в силу (2.21)

$$\mu_2 = 6,6172 - (0,0209)^2 = 6,6168,$$

$$\sigma = \sqrt{\mu_2} = 2,57 \text{ дюйма}.$$

*) Эти вычисления могут быть несколько усовершенствованы, см. упражнение 2.21.

Таблица 2.2

Вычисление стандартного отклонения для распределения, заданного таблицей 1.7 (некоторые предварительные вычисления содержатся в таблице 2.1)

(1) Рост в дюймах	(2) Частота f	(3) Отклонения ξ	(4) $\xi^2 f$
57—	2	—10	200
58—	4	—9	324
59—	14	—8	896
60—	41	—7	2 009
61—	83	—6	2 988
62—	169	—5	4 225
63—	394	—4	6 304
64—	669	—3	6 021
65—	990	—2	3 960
66—	1 223	—1	1 223
67—	1 329	0	0
68—	1 230	1	1 230
69—	1 063	2	4 252
70—	646	3	5 814
71—	392	4	6 272
72—	202	5	5 050
73—	79	6	2 844
74—	32	7	1 568
75—	16	8	1 024
76—	5	9	405
77—	2	10	200
Всего	8585	—	56 809

Можно заметить, что в рассмотренном примере среднее отклонение составляет приблизительно 80% от стандартного отклонения. Это соотношение довольно часто выполнено приближенно для одновершинных распределений, близких к симметричным. Происхождение этого факта станет более ясным, когда мы будем изучать так называемое «нормальное» распределение*) (глава 5).

*) Чакраварти (M. C. Chakravarty), (1948), Calcutta Stat. Ass. Bull. 1, 187, показал, что вообще для генеральной совокупности, состоящей из $n > 2$ элементов,

$$\left(\frac{2}{n}\right)^{1/2} \leq \frac{\delta_1}{\sigma} \leq \begin{cases} (1 - n^{-2})^{1/2}, & n \text{ нечетно,} \\ 1, & n \text{ четно,} \end{cases}$$

причем равенства достигаются.

Пример 2.8

Найти дисперсию распределения

$$dF = \frac{1}{B(p, q)} (1-x)^{p-1} x^{q-1} dx, \quad 0 \leq x \leq 1, \quad p, q > 0.$$

Относительно нуля имеем

$$\mu'_2 = \frac{1}{B(p, q)} \int_0^1 (1-x)^{p-1} x^{q+1} dx = \frac{B(p, q+2)}{B(p, q)} = \frac{(q+1)q}{(p+q+1)(p+q)}.$$

К тому же мы уже нашли (пример 2.2), что

$$\mu'_1 = \frac{q}{p+q}.$$

Поэтому

$$\mu_2 = \mu'_2 - (\mu'_1)^2 = \frac{(q+1)q}{(p+q+1)(p+q)} - \frac{q^2}{(p+q)^2} = \frac{pq}{(p+q+1)(p+q)^2}.$$

Пример 2.9

Так же как и среднее значение, дисперсию можно найти дифференцированием производящей функции. Мы имеем (ср. пример 2.3)

$$\frac{d^2 P}{dt^2} = \sum \{j(j-1)t^{j-2} f_j\},$$

и, полагая $t=1$, получим для моментов, подсчитываемых относительно нуля,

$$\left[\frac{d^2 P}{dt^2} \right]_{t=1} = \sum \{j(j-1) f_j\} = \mu'_2 - \mu'_1. \quad (2.22)$$

В случае биномиального распределения

$$\left[\frac{d^2}{dt^2} \{pt + (1-p)\}^n \right]_{t=1} = n(n-1)p^2,$$

откуда

$$\mu'_2 = n(n-1)p^2 + np,$$

$$\mu'_2 = n(n-1)p^2 + np - n^2 p^2 = np(1-p). \quad (2.23)$$

Поправки Шеппарда

2.20 Естественно, что группирование значений варианты путем приписывания соответствующих им частот середине интервала группирования дает представление о характере частотного распределения лишь с некоторым приближением. В ряде случаев, однако, оказывается возможным ввести поправки, назы-

ваемые «поправками Шеппарда», на те искажения, которые вносятся в результате группирования. Подробно они будут изучаться в следующей главе. Сейчас же мы только укажем (без доказательства) необходимые поправки для второго момента.

Если распределение непрерывно и имеет достаточно высокий порядок соприкосновения с осью значений варианты на его концах, т. е. если «хвосты» частотного распределения достаточно быстро убывают, то для корректирования второго момента, грубо подсчитанного на основе сгруппированных данных, надо вычесть из него $h^2/12$, где h — ширина интервала группирования. Так, в примере 2.7 $h=1$ и исправленный второй момент равен

$$6,6168 - 0,0833 = 6,5335.$$

Исправленное значение σ равно $\sqrt{6,5335} = 2,56$ в отличие от неисправленного, равного 2,57.

Средняя разность

2.21 Коэффициент средней разности (не путать со средним отклонением!) определяется по формуле

$$\begin{aligned} \Delta_1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - y| dF(x) dF(y) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - y| f(x) f(y) dx dy. \end{aligned} \quad (2.24)$$

Для дискретного случая Δ_1 можно определить двумя разными формулами:

$$\Delta_1 = \frac{1}{N(N-1)} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |x_j - x_k| f(x_j) f(x_k), \quad j \neq k \quad (2.25)$$

(средняя разность без учета совпадений x_j и x_k),

$$\Delta_1 = \frac{1}{N^2} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |x_j - x_k| f(x_j) f(x_k) \quad (2.26)$$

(с учетом совпадений). Разница между ними состоит только в делителе. Очевидно, что при больших N она становится незначительной.

Средняя разность представляет собой среднее значение разностей всех возможных пар значений варианты без учета их знака. В коэффициенте, определенном с учетом совпадений, рассматриваются и отклонения каждого значения от него же, что

конечно ничего не вносит в сумму отклонений, но зато общее число возможных пар будет равно N^2 . В коэффициенте же без совпадений рассматриваются разности только различных значений, число которых, очевидно, равно $N(N-1)$. Именно из этих соображений и выбраны делители в формулах (2.25) и (2.26).

2.22 Средняя разность, введенная Джини (1912), представляет некоторый теоретический интерес как величина, характеризующая разброс значений варианты друг относительно друга и не зависящая от какого-либо центрального значения. Однако эта величина трудно поддается вычислению по сравнению со стандартным отклонением. К тому же наличие в формулах, определяющих среднюю разность, знака модуля вносит известные трудности при использовании ее в выборочной теории. Можно было бы думать, что эти трудности исчезнут, если вместо средней разности ввести коэффициент

$$E^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-y)^2 dF(x) dF(y).$$

Однако это есть не что иное, как удвоенное значение дисперсии. В самом деле,

$$\begin{aligned} E^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{x^2 - 2xy + y^2\} dF(x) dF(y) = \int_{-\infty}^{\infty} x^2 dF(x) \int_{-\infty}^{\infty} dF(y) - \\ &\quad - 2 \int_{-\infty}^{\infty} x dF(x) \int_{-\infty}^{\infty} y dF(y) + \int_{-\infty}^{\infty} dF(x) \int_{-\infty}^{\infty} y^2 dF(y) = \\ &= 2\mu'_2 - 2(\mu'_1)^2 = 2\mu_2. \end{aligned} \quad (2.27)$$

Это интересное соотношение показывает, что дисперсию можно определить как половину среднего значения квадратов всевозможных разностей значений варианты, т. е., иначе говоря, определить, не обращаясь к рассмотрению отклонений от некоторого центрального значения (среднего).

Коэффициенты вариации; нормированное распределение

2.23 Введенные выше меры рассеяния распределений были размерными величинами. Поэтому могут возникнуть трудности при их сравнении для разных генеральных совокупностей. Это привело к поиску мер, которые не зависели бы от шкалы значений варианты, т. е. которые были бы просто числами.

Подобные меры можно сконструировать различными способами. Например, можно взять среднее отклонение и поделить его на среднее или медиану. Однако наиболее часто в практике употребляются две характеристики. Одна из них — это коэффициент вариации Карла Пирсона

$$V = 100 \frac{\sigma}{\mu'_1}, \quad (2.28)$$

а другая — коэффициент рассеяния Джини

$$G = \frac{\Delta_1}{2\mu'_1}. \quad (2.29)$$

Оба эти коэффициента страдают тем недостатком, что они в значительной мере зависят от среднего μ'_1 , подсчитываемого, вообще говоря, относительно произвольной точки. Поэтому их применение ограничивается теми случаями, когда существует некоторое естественное начало отсчета или когда распределения имеют сходные начала отсчета.

2.24 Для наших же целей сравнение различных распределений удобнее проводить другим путем. Возьмем величину σ за новую единицу масштаба и перейдем к варианту

$$y = \frac{x - \mu'_1}{\sigma}. \quad (2.30)$$

Ее среднее значение равно нулю, дисперсия равна единице. Распредельные варианты y можно назвать нормированным распределением. Нормированные распределения обладают тем удобством, что их легко можно сравнивать с точки зрения их формы, асимметрии и других показателей, за исключением, конечно, средних и дисперсий.

Рассеяние

2.25 К определению коэффициента рассеяния Джини можно прийти следующим естественным путем. Пусть, как обычно,

$$\left. \begin{aligned} F(x) &= \int_{-\infty}^x f(x) dx, \\ \Phi(x) &= \frac{1}{\mu'_1} \int_{-\infty}^x x f(x) dx. \end{aligned} \right\} \quad (2.31)$$

и пусть

Конечно, $\Phi(x)$ существует, если только существует μ'_1 . Подобно тому как $F(x)$ изменяется от 0 до 1, функция $\Phi(x)$ также меняется в пределах от 0 до 1, если только начало отсчета взято левее начала частотного распределения, которое в дальнейшем предполагается равным нулю. Функция $\Phi(x)$ может быть названа неполным первым моментом.

Уравнение (2.31) можно рассматривать как соотношение между переменными F и Φ , если x считать параметром. Кривая $\Phi = \Phi(F)$ называется кривой рассеяния. График такой кривой показан на рис. 2.2.

Кривая рассеяния выпукла. В самом деле,

$$\mu'_1 \frac{d\Phi}{dF} = \frac{x f'(x)}{f(x)} = x$$

и

$$\mu'_1 \frac{d^2\Phi}{dF^2} = \frac{dx}{dF} = \frac{1}{f(x)} > 0.$$

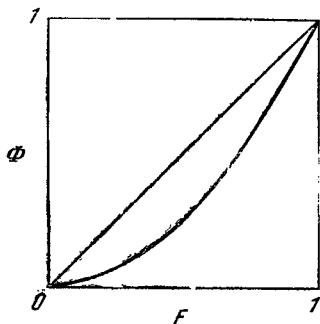


Рис. 2.2. Кривая рассеяния.

Поэтому в силу того, что μ'_1 взято относительно нуля, касательная к кривой рассеяния имеет положительный угол с осью F и этот угол возрастает при увеличении F . Другими словами, кривая выпукла.

Площадь S , заключенная между кривой рассеяния и линией $\Phi = F$, называется площадью рассеяния. Мы сейчас покажем, что эта площадь численно равна половине коэффициента рассеяния. Действительно, из рис. 2.2 следует, что

$$2S = \int_0^1 F d\Phi - \int_0^1 \Phi dF,$$

и поэтому

$$\begin{aligned} 2\mu'_1 S &= \int_{-\infty}^{\infty} F(x) x dF(x) - \mu'_1 \int_{-\infty}^{\infty} \Phi(x) dF(x) = \\ &= \int_{-\infty}^{\infty} x dF(x) \int_{-\infty}^x dF(y) - \int_{-\infty}^{\infty} dF(x) \int_{-\infty}^x y dF(y) = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^x (x - y) dF(x) dF(y). \end{aligned}$$

Но $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-y) dF(x) dF(y) = 0$ и, следовательно,

$$2\mu'_1 S = \frac{1}{2} \left[\int_{-\infty}^{\infty} \int_{-\infty}^x (x-y) dF(x) dF(y) + \right. \\ \left. + \int_{-\infty}^{\infty} \int_x^{\infty} (y-x) dF(x) dF(y) \right] = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x-y| dF(x) dF(y) = \frac{1}{2} \Delta_1.$$

Тем самым площадь рассеяния равна $\Delta_1/4\mu'_1$, что составляет половину коэффициента рассеяния G , определяемого формулой (2.29).

2.26 Были предложены разные способы вычисления средней плотности. Приводимый ниже способ, вероятно, самый простой, особенно для распределений с равными интервалами группирования.

Без ограничения общности будем считать, что начало отсчета находится в точке начала распределения. Тогда можно написать

$$\sum_{j=1}^N \sum_{k=1}^N |x_j - x_k| = 2 \sum' (x_j - x_k),$$

где суммирование $\sum'$ производится по всем тем индексам, для которых $j \geq k$. Далее

$$(x_j - x_k) = (x_j - x_{j-1}) + (x_{j-1} - x_{j-2}) + \dots + (x_{k+1} - x_k),$$

и поэтому

$$\sum' (x_j - x_k) = \sum_{h=1}^{N-1} C_h (x_{h+1} - x_h),$$

где C_h — число членов вида $(x_j - x_k)$ в $\sum'$, содержащих $x_{h+1} - x_h$. Поскольку h — число значений j , меньших или равных h (начало отсчета взято в точке начала распределения), и $(N-h)$ — число значений, больших или равных $h+1$, то $C_h = h(N-h)$ и поэтому

$$\Delta_1 = \frac{2}{N^2} \sum' (x_j - x_k) = \frac{2}{N^2} \sum_{h=1}^{N-1} h(N-h) (x_{h+1} - x_h). \quad (2.32)$$

Эта формула особенно удобна, когда все интервалы равны. Тогда если F_h — функция распределения для x_h , то

$$\Delta_1 = \frac{2}{N^2} \sum_{h=1}^{N-1} (NF_h)(N - NF_h) = 2 \sum_{h=1}^{N-1} F_h(1 - F_h). \quad (2.33)$$

В том случае, когда истинные накопленные частоты для x_h равны G_h ,

$$\Delta_1 = \frac{2}{N^2} \sum_{h=1}^{N-1} G_h (N - G_h). \quad (2.34)$$

Эта формула очень удобна для практических целей.

Пример 2.10

Вернемся еще раз к распределению роста мужчин, рассмотренному в предшествующих примерах. Сумму $\sum G_h (N - G_h)$ можно вычислить по схеме, предложенной в таблице 2.3.

Таблица 2.3

Вычисление средней разности для распределения роста мужчин, заданного в таблице 1.7

(1) Рост в дюймах	(2) Частота	(3) G_h	(4) $N - G_h$	(5) $G_h (N - G_h)$
57—	2	2	8 583	17 166
58—	4	6	8 579	51 474
59—	14	20	8 565	171 300
60—	41	61	8 524	519 964
61—	83	144	8 441	1 215 504
62—	169	313	8 272	2 589 136
63—	394	707	7 878	5 569 746
64—	669	1 376	7 209	9 919 584
65—	990	2 366	6 219	14 714 154
66—	1 223	3 589	4 996	17 930 644
67—	1 329	4 918	3 667	18 034 306
68—	1 230	6 148	2 437	14 982 676
69—	1 063	7 211	1 374	9 907 914
70—	646	7 857	728	5 719 896
71—	392	8 249	336	2 771 664
72—	202	8 451	134	1 132 434
73—	79	8 530	55	469 150
74—	32	8 562	23	196 926
75—	16	8 578	7	60 046
76—	5	8 583	2	17 166
77—	2	8 585	—	—
Всего	8 585	—	—	105 990 850

Используя формулу (2.34) для средней разности с учетом повторений, мы получим

$$\Delta_1 = \frac{2 \cdot 105\,990\,850}{8585^2} = 2,88 \text{ (дюймов)},$$

в то время как среднее отклонение (в дюймах) равно 2,02, а стандартное отклонение равно 2,57 (пример 2.7). Отсюда видно, что существует некоторое различие между этими величинами, но в этом, конечно, нет ничего удивительного, поскольку эти величины различны по самой своей природе. Поэтому нет оснований считать, что их численные значения в частных случаях окажутся примерно одинаковыми.

Средние значения; общий случай

2.27 Понятие арифметического среднего варианты x можно легко распространить на функции $h(x)$ от нее. Определим

$$M(h) = \int_{-\infty}^{\infty} h(x) dF(x) \quad (2.35)$$

при условии, конечно, что выражение справа существует. Так определенное среднее значение для $h(x)$ играет фундаментальную роль в теории вероятностей и статистике. Происхождение символа $M(\cdot)$ объясняется тем, что в теории вероятностей такие значения называют «математическими ожиданиями». Весьма замечательным и плодотворным является тот факт, что в выборочной теории часто оказывается значительно легче определить некоторые средние характеристики, чем соответствующее частотное распределение.

Если мы имеем функцию $h(x_1, \dots, x_n)$ от многих вариантов $x_1, \dots, x_n$, то положим

$$M(h) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h(x_1, \dots, x_n) dF(x_1, \dots, x_n). \quad (2.36)$$

В том частном случае, когда h является функцией только одной переменной (скажем, x_1), мы имеем

$$M(h) = \int_{-\infty}^{\infty} h(x_1) dF(x_1), \quad (2.37)$$

где $F(x_1)$ — маргинальная функция распределения для x_1 . Это значение совпадает со значением (2.35). Отсюда видно, что среднее значение варианты (или некоторой функции от нее) будет тем же самым, независимо от того, является ли распределение этой варианты частью многомерного распределения или же нет.

2.28 Отметим два простых, но важных свойства среднего значения:

(а) Среднее значение суммы равно сумме средних значений. Если h_1 и h_2 — две функции, то

$$M(h_1 + h_2) = M(h_1) + M(h_2). \quad (2.38)$$

Это свойство сразу следует из определений.

(б) Среднее значение произведения двух функций есть произведение их средних значений, если множества значений вариантов, от которых зависят эти функции, независимы между собой. Для примера, если x_1 и x_2 независимые варианты, то их совместная функция распределения равна произведению их функций распределений и

$$\begin{aligned} M(h_1 h_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(x_1) h_2(x_2) dF(x_1, x_2) = \\ &= \int_{-\infty}^{\infty} h_1(x_1) dF(x_1, \infty) \int_{-\infty}^{\infty} h_2(x_2) dF(\infty, x_2) = M(h_1) M(h_2). \end{aligned} \quad (2.39)$$

Равенство (2.38) остается справедливым, даже если варианты зависимы; равенство же (2.39), вообще говоря, для зависимых вариантов не верно.

УПРАЖНЕНИЯ

2.1 Показать, что среднее отклонение относительно произвольной точки будет наименьшим, когда эта точка совпадает с медианой.

2.2 Показать, что среднее (относительно нуля) дискретного распределения с частотами

$$e^{-\lambda} \left(1, \lambda, \frac{\lambda^2}{2!}, \dots, \frac{\lambda^r}{r!}, \dots \right)$$

для значений $0, 1, 2, \dots, r, \dots$ равно λ ; дисперсия также равна λ .

2.3 Показать, что если отклонения малы по сравнению со средним, то приближенно для геометрических и гармонических средних справедливы формулы

$$G = \mu'_1 \left(1 - \frac{1}{2} \frac{\sigma^2}{(\mu'_1)^2} \right),$$

$$H = \mu'_1 \left(1 - \frac{\sigma^2}{(\mu'_1)^2} \right)$$

и, следовательно,

$$\mu'_1 - 2G + H = 0.$$

2.4 Показать, что среднее отклонение относительно среднего не больше, чем стандартное отклонение σ ; установить также что средняя разность не может превысить значения $\sqrt{2} \sigma$.

2.5 Показать, что для равномерного распределения $dF = dx$, $0 \leq x \leq 1$, $\mu'_1(0) = \frac{1}{2}$, $\mu_2 = \frac{1}{12}$, среднее отклонение $= \frac{1}{4}$, $\Delta_1 = \frac{1}{3}$.

2.6 Установить, что для распределения $dF = ke^{-x/\sigma} dx$, $0 \leq x \leq \infty$ (отрицательное показательное распределение), среднее, стандартное отклонение и средняя разность равны σ ; и что интерквартильная широта равна $\sigma \ln 3$.

2.7 Показать, что для распределения

$$dF = ke^{-\frac{1}{2}\chi^2} \chi^{v-1} d\chi, \quad 0 \leq \chi \leq \infty,$$

$$\mu'_1(0) = \sqrt{2} \Gamma\left\{\frac{1}{2}(v+1)\right\} / \Gamma\left(\frac{1}{2}v\right),$$

$$\mu'_2 = v.$$

2.8 Показать, что если интервал длиной в 6σ содержит по меньшей мере 18 интервалов группирования, то поправки Шеппарда для стандартного отклонения меньше 0,5% от неисправленного значения стандартного отклонения.

2.9 Показать, что для непрерывного распределения, среднее которого существует,

$$\Delta_1 = 2 \int_{-\infty}^{\infty} F(x) \{1 - F(x)\} dx = 4 \int_{-\infty}^{\infty} x \left\{ F(x) - \frac{1}{2} \right\} dF.$$

2.10 Если значения некоторой варианты в порядке их возрастания суть $x_1, \dots, x_N$ и

$$s_r = \sum_{j=1}^r x_j, \quad t_r = \sum_{j=1}^r x_{N-j+1}, \quad U = \sum_{r=1}^N s_r, \quad V = \sum_{r=1}^N t_r,$$

тогда

$$\Delta_1 = \frac{2}{N^2} (V - U) = \frac{2}{N^2} \{N(N+1)\mu'_1 - 2U\}.$$

2.11 Показать, что для отрицательного биномиального распределения (см. упражнение 1.14) среднее и дисперсия соответственно равны m и $m\{m/k+1\}$.

2.12 Показать, что дисперсия распределения

$$dF = \frac{k}{1+x^4} dx, \quad -\infty \leq x \leq \infty,$$

равна единице.

2.13 Найти среднее значение варианты $r^2 = x_1^2 + x_2^2 + x_3^2$, где x_1, x_2, x_3 — варианты из упражнения 1.10, а) непосредственно и б) из распределения для r^2 .

2.14 Для распределения в упражнении 1.11 показать, что среднее значение x^2 равно σ_1^2 и для y^2 равно σ_2^2 . Дифференцируя $\int dF = 1$ по ρ , показать, что $M(xy) = \rho\sigma_1\sigma_2$.

2.15 Пусть варианта принимает неотрицательные значения и имеет среднее значение μ'_1 . Показать, что для каждого положительного t

$$F(t) > 1 - \mu'_1/t.$$

2.16 Установить, что если распределение

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx, \quad -\infty \leq x \leq \infty,$$

преобразовать по формуле $x = a \log(y-b) + c$, то распределение y будет таково, что для него

$$\frac{\text{среднее} - \text{мода}}{\text{среднее} - \text{медиана}} = \frac{\exp\left(\frac{1}{2a^2}\right) - \exp\left(-\frac{1}{a^2}\right)}{\exp\left(\frac{1}{2a^2}\right) - 1},$$

и что это отношение стремится к 3 при $a \rightarrow \infty$.

2.17 Некоторая варианта с невозрастающей функцией плотности $f(x)$, определенной на интервале от 0 до ∞ , имеет среднее μ . Показать, что

$$\mu \geq \frac{1}{2f(0)}. \quad \text{Проиллюстрировать это на примере распределения}$$

$$dF = \frac{4}{\pi a^2} (a^2 - x^2)^{1/2} dx, \quad 0 \leq x \leq a.$$

2.18 По аналогии с предшествующим упражнением найти неравенство для второго момента. Используя полученное неравенство, показать, что для непрерывной дифференцируемой плотности $f(x)$ с единственной модой ордината моды не меньше, чем $1/(\sqrt{3}\sigma)$, где σ есть корень квадратный из второго момента относительно моды.

2.19 Показать, что если для распределения существует среднее, то существует и средняя разность Δ_1 . Проиллюстрировать это на примере распределения Парето

$$dF = \frac{k}{x^\alpha} dx, \quad 0 < a \leq x \leq \infty, \quad \alpha > 1,$$

для которого среднее может существовать, тогда как дисперсия при $\alpha \leq 3$ не существует.

2.20 Для распределения в упражнении 1.11 показать, что если $M(x^2)M(y^2) = M(x^2y^2)$, тогда варианты независимы.

2.21 При вычислении среднего отклонения для группированных данных в примере 2.7 интервал, в котором лежит среднее, вносит в сумму отклонений величину $f \cdot |d|$. Показать, что если частота f , отвечающая интервалу группирования $(x, x+h)$, в котором лежит μ'_1 , разбивается на две части f_1 и f_2 пропорционально длинам $(\mu'_1 - x)$ и $(x+h - \mu'_1)$ и частоты f_1 и f_2 соответствуют срединам интервалов (x, μ'_1) и $(\mu'_1, x+h)$, то сумма отклонений возрастает на $f|h\left(\frac{|d|}{h} - \frac{1}{2}\right)^2$. Убедиться в том, что в примере 2.7 эта поправка увеличивает δ_1 на 0,04.

МОМЕНТЫ И СЕМИИНВАРИАНТЫ

Определение моментов

3.1 В предыдущей главе мы определили первый момент (арифметическое среднее) относительно произвольной точки a интегралом Стильтьеса

$$\mu'_1 = \int_{-\infty}^{\infty} (x - a) dF, \quad (3.1)$$

а второй момент относительно этой точки как

$$\mu'_2 = \int_{-\infty}^{\infty} (x - a)^2 dF. \quad (3.2)$$

Обобщая эти понятия, мы можем определить последовательность чисел μ'_r , $r = 1, 2, \dots$, положив

$$\mu'_r = \int_{-\infty}^{\infty} (x - a)^r dF. \quad (3.3)$$

Величина μ'_r называется моментом порядка r относительно точки a . В том случае, когда a совпадет со средним μ'_1 , вычисленным относительно нуля, штрих будет опускаться:

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu'_1)^r dF. \quad (3.4)$$

В частности,

$$\mu_1 = 0.$$

Определим еще момент нулевого порядка, положив

$$\mu'_0 = \mu_0 = \int_{-\infty}^{\infty} dF = 1.$$

Мы будем предполагать, что когда говорится об r -м моменте какого-нибудь определенного распределения, соответствующий интеграл (3.3) для этого распределения сходится. Как будет видно из дальнейшего, некоторые из встречающихся в статистике распределений не обладают моментами всех порядков, часть из них имеет только конечное число моментов, есть и такие, которые вообще не имеют моментов, за исключением, конечно, момента нулевого порядка. Очевидно, что если момент r -го порядка существует, то существуют и все моменты более низкого порядка.

3.2 Пусть a, b — два значения варианты, $b - a = c$ и $\mu'(a)$, $\mu'(b)$ — моменты относительно a и b соответственно. Имеем по формуле бинома

$$(x - a)^r = (x - b + b - a)^r = (x - b + c)^r = \sum_{j=0}^r \binom{r}{j} (x - b)^{r-j} c^j.$$

Следовательно,

$$\begin{aligned} \mu'_r(a) &= \int_{-\infty}^{\infty} (x - a)^r dF = \int_{-\infty}^{\infty} \sum_{j=0}^r \binom{r}{j} (x - b)^{r-j} c^j dF = \\ &= \sum_{j=0}^r \binom{r}{j} c^j \int_{-\infty}^{\infty} (x - b)^{r-j} dF = \sum_{j=0}^r \binom{r}{j} \mu'_{r-j}(b) c^j. \end{aligned} \quad (3.5)$$

Это соотношение дает представление r -го момента относительно a через r -й и более низкие моменты относительно b . Его можно записать в следующей удобной для запоминания символической форме:

$$\mu'_r(a) = \{\mu'(b) + c\}^r,$$

где предполагается, что правая часть должна быть записана по биномиальной формуле и величины $\{\mu'(b)\}^j$ заменены на $\mu'_j(b)$.

Равенство (3.5) особенно важно, когда a или b совпадает со средним значением распределения. В этом случае имеем

$$\mu'_r = \sum_{j=0}^r \binom{r}{j} \mu_{r-j} (\mu'_1)^j, \quad (3.6)$$

$$\mu_r = \sum_{j=0}^r \binom{r}{j} \mu'_{r-j} (-\mu'_1)^j. \quad (3.7)$$

В частности,

$$\left. \begin{aligned} \mu'_2 &= \mu_2 + (\mu'_1)^2, \\ \mu'_3 &= \mu_3 + 3\mu'_1\mu_2 + (\mu'_1)^3, \\ \mu'_4 &= \mu_4 + 4\mu'_1\mu_3 + 6(\mu'_1)^2\mu_2 + (\mu'_1)^4 \end{aligned} \right\} \quad (3.8)$$

и

$$\left. \begin{aligned} \mu_2 &= \mu'_2 - (\mu'_1)^2, \\ \mu_3 &= \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3, \\ \mu_4 &= \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4. \end{aligned} \right\} \quad (3.9)$$

Вычисление моментов

3.3 Для распределений, заданных численно таблицей частот, вычисление моментов третьего и более высоких порядков ведется примерно так же, как вычисление первых двух моментов. Для группированных данных (моменты высоких порядков вряд ли когда-либо требуются для негруппированных данных) наблюдения полагаются сосредоточенными в средних точках интервалов; выбирается (удобным образом) какое-либо a и вычисляются моменты относительно этого a ; затем, если это нужно, находят по формулам (3.6) и (3.7) моменты относительно среднего значения. В некоторых случаях эффект группирования может быть учтен соответствующей поправкой.

На практике для численных распределений моменты порядка, большего четырех, используются редко. Они очень чувствительны к выборочным флуктуациям и, вычисленные по небольшому количеству наблюдений, могут нести в себе значительные ошибки.

Существуют два обычно используемых метода для нахождения моментов относительно произвольного начала. Первый из них есть непосредственное обобщение метода, использованного в главе 2 для первых двух моментов. Второй будет рассмотрен в 3.10 в связи с факториальными моментами.

Пример 3.1

Найдем первые четыре момента относительно среднего значения распределения браков в Австралии, приведенного в таблице 1.8.

В процессе вычисления в качестве единичного будем считать трехлетний интервал. Сначала найдем моменты относительно 28,5. Для проверки счета мы пользуемся тождествами вида

$$(x+1)^3 = x^3 + 3x^2 + 3x + 1,$$

$$(x+1)^4 = x^4 + 4x^3 + 6x^2 + 4x + 1.$$

Так, например, наряду со значениями $g(x)x^r$ вычисляются значения $g(x)(x+1)^r$, а затем производится проверка с помощью равенств типа

$$\sum g(x)(x+1)^3 = \sum g(x)x^3 + 3 \sum g(x)x^2 + 3 \sum g(x)x + \sum g(x).$$

Здесь $g(x)$ — действительная (ненормированная) частота. Вычислительная работа показана в таблице 3.1.

Из таблицы находим

$$\begin{aligned}\sum (xg) &= 88\,832, \\ \sum (x^2g) &= 2\,155\,838, \\ \sum (x^3g) &= 12\,798\,362, \\ \sum (x^4g) &= 137\,306\,162.\end{aligned}$$

Деля эти числа на полную частоту, равную 301 785, получаем моменты относительно 28,5

$$\begin{aligned}\mu'_1 &= 0,294\,355\,253, \\ \mu'_2 &= 7,143\,622\,115, \\ \mu'_3 &= 42,408\,873\,867, \\ \mu'_4 &= 454,980\,075\,219.\end{aligned}$$

Отсюда, воспользовавшись формулами (3.9), находим моменты относительно среднего значения

$$\begin{aligned}\mu_2 &= 7,056\,977, \\ \mu_3 &= 36,151\,595, \\ \mu_4 &= 408,738\,210.\end{aligned}$$

Эти значения соответствуют единице, равной по величине класс-интервалу, т. е. трем годам. Чтобы вычислить моменты для случая единицы, равной одному году, нужно найденный r -й момент умножить на 3^r . Например,

$$\mu_2 = 7,056977 \times 9 = 63,51279.$$

3.4 Если распределение задано математической формулой, то нахождение моментов сводится к вычислению некоторых сумм или интегралов. Обычно приходится рассматривать вопрос о существовании моментов. Следующие примеры иллюстрируют относящиеся сюда общие принципы.

Пример 3.2

Рассмотрим биномиальное распределение (ср. примеры 2.3 и 2.9). Обозначим $q=1-p$. Тогда производящая функция

Таблица 3.1

Вычисление первых четырех моментов распределения браков, приведенного в таблице 1.8

Средние значения интервалов (в годах)	g	x	xg	$(x+1)g$	x^2g	$(x+1)^2g$	x^3g	$(x-1)^3g$	x^4g	$(x+1)^4g$
16,5	294	-4	-1 176	-882	4 704	2 646	-18 816	-7 938	75 264	23 814
19,5	10 995	-3	-32 985	-21 990	98 935	43 980	-293 865	-87 960	890 595	175 920
22,5	61 001	-2	-122 002	-61 001	244 004	61 001	-488 008	-61 001	976 016	61 001
25,5	73 054	-1	-73 054	-83 873	73 054	—	-73 054	-156 899	73 054	—
28,5	56 501	0	-229 217	56 501	—	56 501	-876 743	56 501	—	56 501
31,5	33 478	1	33 478	66 956	33 478	133 912	33 478	267 824	33 478	535 648
34,5	20 569	2	41 138	61 707	82 276	185 121	164 552	555 363	329 104	1 666 089
37,5	14 281	3	42 843	57 124	128 529	228 496	385 587	913 984	1 156 761	3 655 936
40,5	9 320	4	37 280	46 600	149 120	233 000	595 480	1 165 000	2 385 920	5 825 000
43,5	6 236	5	31 180	37 416	155 900	224 496	779 500	1 346 976	3 897 500	8 081 856
46,5	4 770	6	28 620	33 390	171 720	233 730	1 030 320	1 636 110	6 181 920	11 452 770
49,5	3 620	7	25 340	28 960	177 380	231 680	1 241 660	1 853 440	8 691 620	14 827 520
52,5	2 190	8	17 520	19 710	140 160	177 390	1 121 280	1 596 510	8 970 240	14 368 590
55,5	1 655	9	14 895	16 550	134 055	165 560	1 206 495	1 655 600	10 858 455	16 550 000
58,5	1 100	10	11 000	12 100	110 000	133 100	1 078 110	1 464 100	11 000 000	16 105 100
61,5	810	11	8 910	9 720	98 010	116 640	1 078 110	1 399 680	11 859 210	16 796 160
64,5	649	12	7 788	8 437	93 456	109 681	1 121 472	1 425 853	13 457 664	18 536 089
67,5	487	13	6 331	6 818	82 303	95 452	1 069 939	1 336 328	13 909 207	18 708 592
70,5	326	14	4 564	4 890	63 896	73 350	894 544	1 100 250	12 523 616	16 503 750
73,5	211	15	3 165	3 376	47 475	54 016	712 125	864 256	10 681 875	13 828 096
76,5	119	16	1 904	2 023	30 464	34 391	487 424	584 647	7 798 784	9 938 999
79,5	73	17	1 241	1 314	21 097	23 652	358 649	425 736	6 097 033	7 663 248
82,5	27	18	485	513	8 748	9 747	157 464	185 193	2 834 352	3 518 667
85,5	14	19	266	280	5 054	5 600	96 026	112 000	1 824 494	2 240 000
88,5	5	20	100	105	2 000	2 205	40 000	46 305	800 000	972 405
Суммы по- ложитель- ных членов	301 785	—	318 049	474 490	2 155 838	2 635 287	13 675 105	19 991 056	137 306 162	202 091 751

запишется в виде $(pt+q)^n$. Подобно тому как было в примерах 2.3 и 2.9,

$$\begin{aligned} \left[\frac{d^3 P}{dt^3} \right]_{t=1} &= \sum_{j=0}^n j(j-1)(j-2) f_j = \\ &= \sum (j^3 f_j) - 3 \sum (j^2 f_j) + 2 \sum (j f_j) = \mu'_3 - 3\mu'_2 + 2\mu'_1. \end{aligned}$$

Моменты здесь вычисляются относительно нуля. Отсюда

$$\mu'_3 = 3\mu'_2 - 2\mu'_1 + n(n-1)(n-2)p^3.$$

Мы уже видели, что

$$\mu'_2 = n(n-1)p^2 + np, \quad \mu'_1 = np.$$

Подставляя эти выражения в предыдущую формулу, находим

$$\mu'_3 = n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np.$$

Обычно более интересны моменты относительно среднего. Согласно (3.9)

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3.$$

Подставляя в правую часть этого равенства найденные выражения и преобразуя, получаем

$$\mu_3 = np(p-1)(2p-1) = \quad (3.10)$$

$$= npq(q-p). \quad (3.11)$$

Подобным же образом находим

$$\mu_4 = 3n^2 p^2 (1-p)^2 + np(1-p)(1-6p+6p^2) = \quad (3.12)$$

$$= 3n^2 p^2 q^2 + npq(1-6pq). \quad (3.13)$$

Очевидно, аналогичным методом можно вычислить моменты сколь угодно высокого порядка. Позднее мы увидим, однако, что существуют более легкие способы нахождения моментов биномиального распределения.

Пример 3.3

Рассмотрим распределение

$$dF = \frac{k}{(1+x^2)^m} dx, \quad -\infty \leq x \leq \infty, \quad m \geq 1.$$

Это распределение унимодально и симметрично относительно $x=0$. Поэтому все существующие моменты нечетного порядка

относительно нуля равны нулю. Постоянная k определяется из уравнения

$$1 = k \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^m} = k \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)}{\Gamma(m)}.$$

Момент относительно среднего порядка $2r$, если он существует, дается интегралом

$$\mu_{2r} = k \int_{-\infty}^{\infty} \frac{x^{2r}}{(1+x^2)^m} dx,$$

сходящимся тогда и только тогда, когда $2m > 2r + 1$. Таким образом, моменты относительно нуля порядка $< 2m - 1$ существуют, а более высокого порядка — нет.

Отметим, что при $m=1$ интеграл $\int_{-\infty}^{\infty} \frac{kx}{1+x^2} dx$ в обычном смысле не сходится, т. е. предел

$$\lim_{n \rightarrow \infty, n' \rightarrow \infty} \int_{-n}^{n'} \frac{kx}{1+x^2} dx$$

не существует, в то время как главное значение

$$\lim_{n \rightarrow \infty} \int_{-n}^n \frac{kx}{1+x^2} dx$$

существует и равно нулю.

Обладает распределение средним значением или нет, является в этом случае вопросом соглашения *). При $m > 1$ среднее существует и равно нулю.

Пользуясь преобразованием $z = \frac{1}{1+x^2}$ в формуле для μ_{2r} , находим

$$\mu_{2r} = k \int_0^1 (1-z)^{r-\frac{1}{2}} z^{m-r-\frac{3}{2}} dz = k \frac{\Gamma\left(r + \frac{1}{2}\right) \Gamma\left(m - r - \frac{1}{2}\right)}{\Gamma(m)},$$

*) Если условиться о существовании среднего значения, то недостатком этого соглашения будет то обстоятельство, что для вариант такого типа среднее значение суммы не обязательно равно сумме средних значений. Ср. Фреше (1937) стр. 45 и упражнение 7.19.

и, подставляя значение k , получим

$$\mu_{2r} = \frac{\Gamma\left(r + \frac{1}{2}\right) \Gamma\left(m - r - \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(m - \frac{1}{2}\right)} \quad \text{при } 2m > 2r + 1.$$

Пример 3.4

Рассмотрим «нормальное» распределение

$$dF = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx, \quad -\infty \leq x \leq \infty. \quad (3.14)$$

Оно симметрично относительно нуля, все его моменты существуют, моменты нечетного порядка равны нулю и

$$\mu_{2r} = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2r} e^{-\frac{x^2}{2\sigma^2}} dx.$$

Интеграл в правой части можно вычислить интегрированием по частям, но мы, однако, предпочтем следующий более прямой метод.

Рассмотрим интеграл

$$M(t) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{x^2}{2\sigma^2}} dx = \exp\left(\frac{1}{2} \sigma^2 t^2\right). \quad (3.15)$$

Для всех действительных значений t

$$e^{tx} e^{-\frac{x^2}{2\sigma^2}} = \sum_{r=0}^{\infty} \left(\frac{t^r}{r!} x^r e^{-\frac{x^2}{2\sigma^2}} \right).$$

Так как ряд в правой части можно интегрировать почленно, то

$$M(t) = \sum_{r=0}^{\infty} \left(\frac{t^r}{r!} \mu_r \right).$$

Иными словами, моменты μ_r суть коэффициенты при $t^r/r!$ в разложении $\exp\left(\frac{1}{2} \sigma^2 t^2\right)$. Следовательно,

$$\mu_{2r} = \frac{\sigma^{2r}}{2^r} \frac{(2r)!}{r!}. \quad (3.16)$$

Производящие функции моментов и характеристические функции

3.5 Предыдущий пример показывает, что в некоторых случаях по функции распределения или функции плотности можно построить функцию $M(t)$, для которой моменты распределения служат коэффициентами при разложении ее по степеням t . В соответствии с этим свойством такие функции называются производящими функциями моментов. Если $P(t)$ — производящая функция, то производящей функцией моментов будет просто $P(e^t)$. Термин «производящая функция моментов» условимся сокращать до п. ф. м.

Для многих распределений интеграл $\int_{-\infty}^{\infty} e^{tx} dF$ или сумма $\sum e^{tx} f(x_j)$ при действительных значениях t не существует. Так обстоит дело, например, с распределением $dF = k(1+x^2)^{-m}dx$ при конечных положительных m . Более полезной вспомогательной функцией является

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF, \quad (3.17)$$

где i — чисто мнимая величина, иногда записываемая в виде $\sqrt{-1}$. Эта функция называется характеристической функцией (сокращенно х. ф.) и имеет большое теоретическое значение. В главе 4 будет показано, что характеристическая функция определяет функцию распределения и в свою очередь, конечно, сама полностью ею определяется. Х. ф. являются также важным инструментом выборочной теории.

Так как в силу самого определения функции распределения интеграл $\int_{-\infty}^{\infty} dF$ сходится, то

$$|\varphi(t)| \leq \int_{-\infty}^{\infty} |e^{itx}| dF = 1,$$

и, следовательно, интеграл Стильтьеса (3.17) сходится абсолютно и равномерно по t . Поэтому в (3.17) можно интегрировать и дифференцировать под знаком интеграла по t всякий раз, когда получающиеся при этом интегралы существуют и равномерно сходятся. Например, обозначая $\frac{d}{dt}$ символом D_t , имеем

$$D_t \varphi(t) = i \int_{-\infty}^{\infty} e^{itx} x dF,$$

и, следовательно, полагая $t=0$, получаем

$$\mu'_r = (-i)^r [D^r \varphi(t)]_{t=0}, \quad (3.18)$$

если только μ_r существует. Отсюда вытекает, что при разложении $\varphi(t)$ по степеням t величины μ'_r должны совпадать с коэффициентами при $(it)^r/r!$. Таким образом, характеристическая функция является также производящей функцией моментов. Во многих случаях при формальных выкладках можно писать, скажем, θ вместо it и игнорировать тот факт, что θ мнимо.

Пример 3.5

Рассмотрим опять биномиальное распределение, частоты которого даются элементами разложения $(p+q)^n$. Производящая функция здесь равна $(pt+q)^n$, и, следовательно, производящей функцией моментов будет

$$M(t) = (pe^t + q)^n.$$

Отсюда

$$\begin{aligned} \mu'_1 &= \left[\frac{d}{dt} M(t) \right]_{t=0} = np, \\ \mu'_2 &= \left[\frac{d^2}{dt^2} M(t) \right]_{t=0} = np + n(n-1)p^2 \text{ и т. д.} \end{aligned}$$

Пример 3.6

Рассмотрим распределение

$$dF = \frac{a^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-ax} dx, \quad 0 \leq x < \infty, \quad a > 0, \quad \gamma > 0.$$

Распределения этого вида образуют III тип в системе Пирсона (см. главу 6). В зависимости от значения γ они могут иметь различную форму, но все моменты у них существуют, поскольку

интеграл $\int_0^\infty x^r a^{-ax} dx$, определяющий Γ -функцию, сходится при $r > -1$. Для характеристической функции имеем выражение

$$\varphi(t) = \frac{a^\gamma}{\Gamma(\gamma)} \int_0^\infty e^{x(-a+it)} x^{\gamma-1} dx.$$

После подстановки $z = x(a - it)$ получаем

$$\varphi(t) = \frac{a^\gamma}{\Gamma(\gamma) (a - it)^\gamma} \int_0^\infty e^{-z} z^{\gamma-1} dz = \frac{1}{\left(1 - \frac{it}{a}\right)^\gamma}, \quad (3.19)$$

поскольку $\int_0^{\infty} e^{-z} z^{\gamma-1} dz = \Gamma(\gamma)$. Следовательно,

$$\varphi(t) = 1 + \gamma \frac{it}{a} + \frac{\gamma(\gamma+1)}{2!} \left(\frac{it}{a}\right)^2 + \dots,$$

откуда

$$\mu'_1 = \frac{\gamma}{a}, \quad \mu'_2 = \frac{\gamma(\gamma+1)}{a^2}, \quad \mu'_3 = \frac{\gamma(\gamma+1)(\gamma+2)}{a^3}$$

и т. д. В частности,

$$\mu_2 = \frac{\gamma}{a^2}, \quad \mu_3 = \frac{2\gamma}{a^3}. \quad (3.20)$$

Абсолютные моменты

3.6 Величина

$$\nu'_r = \int_{-\infty}^{\infty} |x - a|^r dF \quad (3.21)$$

называется абсолютным моментом порядка r относительно a . Абсолютные моменты относительно среднего обозначаются теми же буквами, но без штриха. Ясно, что если существует момент порядка r , то существуют абсолютные моменты всех порядков, не превосходящих r .

Если r четно, то абсолютный момент совпадает с обычным моментом. Если распределение сосредоточено на полуоси, лежащей вправо от какой-либо точки, то абсолютные моменты относительно любой точки, находящейся левее начала распределения, совпадают с соответствующими обычными моментами того же порядка.

Абсолютные моменты удовлетворяют ряду интересных неравенств. Пользуясь неубыванием функции $A(t)$, введенной в 2.6, и полагая $t=1, 2, \dots$, находим, что

$$(\nu'_1)^{1/1} \leq (\nu'_2)^{1/2} \leq (\nu'_3)^{1/3} \leq \dots \leq (\nu'_r)^{1/r}.$$

Имеет место также более общее неравенство, принадлежащее Ляпунову (1901):

$$(\nu'_b)^{a-c} \leq (\nu'_c)^{a-b} (\nu'_a)^{b-c}, \quad a \geq b \geq c \geq 0. \quad (3.22)$$

Доказательство этого результата содержится в упражнении 3.15.

Факториальные моменты

3.7 Произведение

$$x(x-h)(x-2h) \dots \{x-(r-1)h\}$$

удобно записывать в виде $x^{[r]}$. Это обозначение указывает на аналогию со степенью x^r . Беря первые разности (по x) с h в качестве приращения, имеем

$$\Delta x^{[r]} = (x+h)^{[r]} - x^{[r]} = (x+h)x(x-h) \dots \\ \dots \{x - (r-2)h\} - x(x-h) \dots \{x - (r-1)h\} = rx^{[r-1]}h.$$

Это соотношение похоже на уравнение в дифференциалах

$$dx^r = rx^{r-1} dx.$$

Обратно, равенство

$$\sum_{x=0}^x x^{[r]} = \frac{1}{(r+1)h} (x+h)^{[r+1]}$$

соответствует равенству

$$\int_0^x x^r dx = \frac{1}{r+1} x^{r+1}.$$

Далее, r -й факториальный момент относительно произвольной точки a естественно определить как

$$\mu'_{[r]} = \sum_{j=-\infty}^{\infty} (x_j - a)^{[r]} f(x_j). \quad (3.23)$$

Здесь написан знак суммирования Σ вместо интеграла Стильтеса, поскольку факториальные моменты почти исключительно используются лишь для дискретных или группированных в интервалы длины h непрерывных распределений. Факториальные моменты не играют особой роли в статистике, но они позволяют записать в очень компактном виде формулы для моментов некоторых дискретных распределений, аналогичных биномиальному.

Как обычно, когда нужно сделать различие между факториальными моментами относительно произвольной точки и факториальными моментами относительно среднего, в обозначениях этих последних штрих опускается.

3.8 Факториальные моменты преобразуются по формуле, подобной формуле (3.5) для преобразования обычных моментов. В самом деле, имеем *)

$$(a+b)^{[r]} = \sum_{j=0}^r \binom{r}{j} a^{[r-j]} b^{[j]},$$

*) Ясно, что выражение $(a+b)^{[r]}$ является полиномом r -й степени по a .

Следовательно, его можно записать в виде $\sum_{j=0}^r k_j a^{[r-j]}$, где k_j — полиномы

следовательно,

$$(x-a)^{[r]} = (x-b+c)^{[r]} = \sum_{j=0}^r \binom{r}{j} (x-b)^{[r-j]} c^{[j]}, \quad \text{где } c = b-a,$$

и

$$\mu'_{[r]}(a) = \sum_{j=0}^r \binom{r}{j} \mu'_{[r-j]}(b) c^{[j]}. \quad (3.24)$$

Символически это можно записать в виде

$$\mu'_{[r]}(a) = \{\mu'(b) + c\}^{[r]}.$$

3.9 Развертывая факториальные выражения в (3.23), получаем

$$\left. \begin{aligned} \mu'_{[1]} &= \mu'_1, \\ \mu'_{[2]} &= \mu'_2 - h\mu'_1, \\ \mu'_{[3]} &= \mu'_3 - 3h\mu'_2 + 2h^2\mu'_1, \\ \mu'_{[4]} &= \mu'_4 + 6h\mu'_3 + 11h^2\mu'_2 - 6h^3\mu'_1. \end{aligned} \right\} \quad (3.25)$$

Обратно,

$$\left. \begin{aligned} \mu'_2 &= \mu'_{[2]} + h\mu'_{[1]}, \\ \mu'_3 &= \mu'_{[3]} + 3h\mu'_{[2]} + h^2\mu'_{[1]}, \\ \mu'_4 &= \mu'_{[4]} + 6h\mu'_{[3]} + 7h^2\mu'_{[2]} + h^3\mu'_{[1]}. \end{aligned} \right\} \quad (3.26)$$

Так как первые моменты совпадают, то эти соотношения останутся верными, если опустить штрихи и отбросить члены, содержащие первые моменты.

Фришмен (1926) были получены следующие общие формулы, выражающие факториальные моменты относительно одной точки через обычные моменты относительно другой точки и обратно:

$$\mu'_r(a) = \sum_{j=0}^r \left\{ \binom{r}{j} B_{r-j}^{(-j)}(c) h^{r-j} \mu'_{[j]}(b) \right\}, \quad (3.27)$$

$$\mu'_{[r]}(a) = \sum_{j=0}^r \left\{ \binom{r}{j} B_{r-j}^{(r+1)}(c+h) h^{r-j} \mu'_j(b) \right\}, \quad (3.28)$$

где $B_r^{(n)}(x)$ — полиномы Бернулли n -го порядка степени r по x (см. 3.25).

от b и h , не зависящие от a . Полагая $a=0$, получаем $b^{[r]} = k_r$. Беря первые разности относительно a и полагая $a=0$, получаем $rb^{[r-1]} = k_{r-1}$. Дальнейшие последовательные разности дают значения остальных k_j . В результате получается нужная формула.

Вычисление факториальных моментов

3.10. Факториальные моменты для группированных данных могут быть вычислены посредством последовательного суммирования, проиллюстрированного в таблице 3.2.

Таблица 3.2

(1) Частоты	(2) Первое суммирование	(3) Второе суммирование	(4) Третье суммирование
f_1	$f_1 + \dots + f_n$	$f_2 + 2f_3 + \dots + (n-1)f_n$	$f_3 + 3f_4 + \dots$
f_2	$f_2 + \dots + f_n$	$f_3 + 2f_4 + \dots + (n-2)f_n$	$\dots + \frac{(n-1)(n-2)}{2} f_n$
f_3	$f_3 + \dots + f_n$	$f_4 + 2f_5 + \dots + (n-3)f_n$	$f_4 + 3f_5 + \dots$
f_4	$f_4 + \dots + f_n$	$f_5 + 2f_6 + \dots + (n-4)f_n$	$\dots + \frac{(n-2)(n-3)}{2} f_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_{n-2}	$f_{n-2} + f_{n-1} + f_n$	$f_{n-2} + 2f_{n-1} + 3f_n$	$f_{n-2} +$ $+ 3f_{n-1} + 6f_n$
f_{n-1}	$f_{n-1} + f_n$	$f_{n-1} + 2f_n$	$f_{n-1} + 3f_n$
f_n	f_n	f_n	f_n
Суммы по столб- цам	$\sum (if_j) = \mu'_{[1]}$	$\sum \left\{ \frac{j(j-1)}{2} f_j \right\} = \frac{1}{2!} \mu'_{[2]}$	$\sum \left\{ \frac{j(j-1)(j-2)}{3!} f_j \right\} =$ $= \frac{1}{3!} \mu'_{[3]}$

В первом столбце выписаны относительные частоты $f_1, \dots, f_n$ в n последовательных интервалах. Элементы второго столбца строятся сложением частот первого столбца, начиная снизу; в n -ю строку пишем f_n , в $(n-1)$ -ю пишем сумму $f_n + f_{n-1}$, в $(n-2)$ -ю пишем сумму $f_n + f_{n-1} + f_{n-2}$ и т. д. до первой строки, содержащей $f_n + f_{n-1} + \dots + f_1$.

В третьем столбце процесс повторяется по строкам второго столбца, только обрывается он на второй строке. Именно, в n -й строке пишем f_n , в $(n-1)$ -й пишем сумму $(f_n + f_{n-1}) + f_n = 2f_n + f_{n-1}$ и т. д. до второй строки, содержащей $(n-1)f_n + (n-2)f_{n-1} + \dots + 2f_3 + f_2$.

В четвертом столбце продолжается тот же процесс над элементами третьего столбца, только обрывается он на третьей строке, и т. д.

Рассмотрим теперь сумму элементов второго столбца. Частота f_1 в нем встречается один раз, f_2 — два раза, ..., f_n встречается n раз. Поэтому эта сумма равна $\sum_{j=1}^n (j f_j) = \mu'_{[1]}$.

В третьем столбце f_2 встречается один раз, f_3 — два раза, ..., f_n встречается $\frac{1}{2} n(n-1)$ раз. Следовательно, сумма элементов этого столбца равна

$$\sum_{j=1}^n \left(\frac{j(j-1)}{2} f_j \right) = \frac{1}{2} \mu'_{[2]}.$$

Вообще сумма элементов $(r+1)$ -го столбца равна $\frac{1}{r!} \mu'_{[r]}$.

Если вместо относительных частот используются действительные частоты, то суммы следует разделить на полную частоту N .

Таким образом, указанный процесс суммирования непосредственно дает значения факториальных моментов.

Пример 3.7

Рассмотрим опять данные таблицы 1.7, показывающие распределения 8585 мужчин в соответствии с их ростом в дюймах. В правых столбцах таблицы 3.3 выписаны последовательные суммы. Наверху каждого из трех последних столбцов помещены в скобках числа, которые получились бы, если суммирование было бы продолжено на один шаг дальше, чем это нужно для вычисления сумм внизу. Эти числа в скобках полезно иметь для проверки, поскольку каждое из них равно сумме, выписанной ниже предыдущего столбца.

Из таблицы находим, что

$$\begin{aligned} \mu'_{[1]} &= 11,020\,850\,320\,33, \\ \mu'_{[2]} &= 117,055\,096\,097\,84, \\ \mu'_{[3]} &= 1194,957\,483\,983\,69, \\ \mu'_{[4]} &= 11\,702,727\,082\,119\,98. \end{aligned}$$

Отсюда, пользуясь соотношениями (3.26), получаем значения обычных моментов

$$\begin{aligned} \mu'_1 &= 11,020\,850\,320\,33, \\ \mu'_2 &= 128,075\,946\,418\,2, \\ \mu'_3 &= 1557,143\,622\,597\,5, \\ \mu'_4 &= 19\,702,878\,509\,027\,3 \end{aligned}$$

Таблица 3.3

Вычисление факториальных моментов распределения мужчин по их росту в дюймах (таблица 1.7)

Рост	Частота	Первое суммирование	Второе суммирование	Третье суммирование	Четвертое суммирование
57—	2	8 585	(94 614)	—	—
58—	4	8 583	86 029	(502 459)	—
59—	14	8 579	77 446	416 430	(1 709 785)
60—	41	8 565	68 867	338 984	1 293 355
61—	83	8 524	60 302	270 117	954 371
62—	169	8 441	51 778	209 815	684 254
63—	394	8 272	43 337	158 037	474 439
64—	669	7 878	35 065	114 700	316 402
65—	990	7 209	27 187	79 635	201 702
66—	1 223	6 219	19 978	52 448	122 067
67—	1 329	4 996	13 759	32 470	69 619
68—	1 230	3 667	8 763	18 711	37 419
69—	1 063	2 437	5 096	9 948	18 438
70—	646	1 374	2 659	4 852	8 490
71—	392	728	1 285	2 193	3 638
72—	202	336	557	908	1 445
73—	79	134	221	351	537
74—	32	55	87	130	186
75—	16	23	32	43	56
76—	5	7	9	11	13
77—	2	2	2	2	2
Суммы по столбцам	8 585	94 614	502 459	1 709 785	4 186 163

и с их помощью находим моменты относительно среднего

$$\mu_2 = 6,616\,805,$$

$$\mu_3 = -0,207\,840,$$

$$\mu_4 = 137,689\,185.$$

В качестве единицы здесь берется один дюйм.

Производящие функции факториальных моментов

3.11 Если производящая функция равна

$$P(t) = \sum_{j=0}^{\infty} f_j t^j,$$

то, заменяя t на $1+t$, имеем

$$\begin{aligned} P(1+t) &= \sum f_j (1+t)^j = \sum_{j=0}^{\infty} f_j \sum_{i=1}^j \binom{j}{i} t^i = \\ &= \sum_{i=0}^{\infty} \frac{t^i}{i!} \sum_{j=0}^{\infty} [f_j j(j-1) \dots \{j-i+1\}] = \sum_{i=0}^{\infty} (t^i \mu'_{[i]}/i!). \quad (3.29) \end{aligned}$$

Таким образом, $P(1+t)$ является производящей функцией факториальных моментов (п. ф. ф. м.), поскольку в ее разложении факториальный момент порядка i совпадает с коэффициентом при $t^i/i!$. Мы предполагали здесь, что возможные значения отстоят друг от друга на единицу.

Пример 3.8

Рассмотрим опять биномиальное распределение. Производящая функция его равна $\{pt + (1-p)\}^n$. Имеем

$$P(1+t) = (1+pt)^n = \sum_{j=0}^n p^j t^j \binom{n}{j}.$$

Следовательно, факториальные моменты относительно нуля суть

$$\mu'_{[r]} = p^r n^{[r]}, \quad r \leq n,$$

$$\mu'_{[r]} = 0, \quad r > n.$$

Отсюда мы можем, если угодно, получить обычные моменты посредством (3.26). Например,

$$\mu'_1 = \mu_{[1]} = pn,$$

$$\mu'_2 = p^2 n(n-1) + pn \quad \text{и т. д.}$$

Семиинварианты

3.12 Моменты представляют собой совокупность характеризующих распределение постоянных, полезную для измерения его свойств и, при некоторых условиях, для его определения. С этой точки зрения моменты еще будут рассмотрены в последующих главах. Однако они не являются единственной и тем более наилучшей для этих целей совокупностью постоянных. Иное множество постоянных — так называемых семиинвариантов — обладает свойствами, более полезными с теоретической точки зрения.

Формально семиинварианты $\kappa_1, \kappa_2, \dots, \kappa_r, \dots$ определяются следующим тождеством по t :

$$\begin{aligned} \exp \left\{ \kappa_1 t + \frac{\kappa_2 t^2}{2!} + \dots + \frac{\kappa_r t^r}{r!} + \dots \right\} = \\ = 1 + \mu'_1 t + \frac{\mu'_2 t^2}{2!} + \dots + \frac{\mu'_r t^r}{r!} + \dots \quad (3.30) \end{aligned}$$

Иногда более удобно писать это тождество, заменив t на it :

$$\begin{aligned} \exp \left\{ \kappa_1 (it) + \kappa_2 \frac{(it)^2}{2!} + \dots + \kappa_r \frac{(it)^r}{r!} + \dots \right\} = \\ = 1 + \mu'_1 \frac{it}{1!} + \dots + \mu'_r \frac{(it)^r}{r!} + \dots = \int_{-\infty}^{\infty} e^{itx} dF = \varphi(t). \quad (3.31) \end{aligned}$$

Мы видим, что, тогда как μ'_r является коэффициентом при $(it)^r/r!$ в разложении характеристической функции $\varphi(t)$, κ_r представляет собой коэффициент при $(it)^r/r!$ в разложении $\log \varphi(t)$, если такое разложение существует. Функция $\log \varphi(t)$ может поэтому быть названа производящей функцией семиинвариантов (п. ф. с.).

3.13 Если в соотношении (3.31) в качестве начала отсчета взять вместо 0 точку c , то $\varphi(t)$ умножится на e^{-itc} , поскольку $\int e^{itx} dF$ переходит в $\int e^{it(x-c)} dF$. К $\log \varphi(t)$ при этом прибавится лишь член $-itc$. Поэтому коэффициенты разложения $\log \varphi(t)$ останутся неизменными, за исключением первого, из которого вычтется c .

Таким образом, все семиинварианты, исключая первый, инвариантны относительно изменения начала отсчета. В этом они существенно отличаются от моментов относительно произвольной точки.

Как семиинварианты, так и моменты относительно нуля обладают следующим свойством инвариантности: если значения случайной переменной умножаются на постоянную a , то μ'_r и κ_r умножаются на a^r . Это непосредственно следует из их определений. Таким образом, при любом линейном преобразовании вида

$$\xi = lx + m \quad (3.32)$$

семиинварианты κ_r лишь умножаются на l^r . Исключение составляет первый семиинвариант, совпадающий со средним значением. В частности, если перейти к нормированному распределению, то семиинварианты κ_r лишь умножатся на σ^{-r} , где $\sigma = \frac{1}{\kappa_2^{1/2}}$ — стандартное отклонение, равное, как мы вскоре увидим,

Величины κ_r названы семиинвариантами из-за тех инвариантных свойств, которыми они обладают. Впервые семиинварианты рассматривались Тиле (1903).

Соотношения между моментами и семиинвариантами

3.14 Если выполнены условия существования, то из (3.30) следует, что

$$\begin{aligned} 1 + \mu'_1 \frac{t}{1!} + \dots + \mu'_r \frac{t^r}{r!} + \dots &= \\ &= \exp \left\{ \kappa_1 \frac{t}{1!} + \kappa_2 \frac{t^2}{2!} + \dots + \kappa_r \frac{t^r}{r!} + \dots \right\} = \\ &= \exp \left(\frac{\kappa_1 t}{1!} \right) \exp \left(\frac{\kappa_2 t^2}{2!} \right) \dots \exp \left(\frac{\kappa_r t^r}{r!} \right) \dots = \\ &= \left\{ 1 + \frac{\kappa_1 t}{1!} + \frac{\kappa_1^2 t^2}{2!} + \dots \right\} \left\{ 1 + \frac{\kappa_2 t^2}{2!} + \frac{1}{2!} \left(\frac{\kappa_2^2 t^4}{(2!)^2} \right) + \dots \right\} \dots \\ &\quad \dots \left\{ 1 + \left(\frac{\kappa_r t^r}{r!} \right) + \frac{1}{2!} \left(\frac{\kappa_r t^r}{r!} \right)^2 + \dots \right\} \dots \end{aligned}$$

Перемножая разложение экспонент и собирая члены, содержащие степень t^r , получаем

$$\mu'_r = \sum_{m=1}^r \sum \left(\frac{\kappa_{p_1}}{p_1!} \right)^{\pi_1} \left(\frac{\kappa_{p_2}}{p_2!} \right)^{\pi_2} \dots \left(\frac{\kappa_{p_m}}{p_m!} \right)^{\pi_m} \frac{r!}{\pi_1! \pi_2! \dots \pi_m!}, \quad (3.33)$$

где второе суммирование производится по всем положительным значениям π , для которых

$$p_1 \pi_1 + p_2 \pi_2 + \dots + p_m \pi_m = r. \quad (3.34)$$

Важно отметить, что утомительный процесс выписывания точных выражений для частных значений r может быть значительно упрощен. Действительно, дифференцируя (3.30) по κ_j , имеем

$$\frac{t^j}{j!} \left(1 + \mu'_1 t + \dots + \frac{\mu'_r t^r}{r!} + \dots \right) = \frac{\partial \mu'_1}{\partial \kappa_j} t + \dots + \frac{t^r}{r!} \frac{\partial \mu'_r}{\partial \kappa_j} + \dots,$$

и, следовательно, приравнивая коэффициенты при одинаковых степенях t , получаем

$$\frac{\partial \mu'_r}{\partial \kappa_j} = \binom{r}{j} \mu'_{r-j}. \quad (3.35)$$

В частности,

$$\frac{\partial \mu'_r}{\partial \kappa_1} = r \mu'_{r-1}. \quad (3.36)$$

Отсюда вытекает, что если известно выражение μ'_r через κ_j , то подобные выражения для моментов более низкого порядка могут быть получены дифференцированием *).

Вот первые десять из этих выражений для моментов относительно произвольной точки:

$$\begin{aligned}
 \mu'_1 &= \kappa_1, \\
 \mu'_2 &= \kappa_2 + \kappa_1^2, \\
 \mu'_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3, \\
 \mu'_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4, \\
 \mu'_5 &= \kappa_5 + 5\kappa_4\kappa_1 + 10\kappa_3\kappa_2 + 10\kappa_3\kappa_1^2 + 15\kappa_2^2\kappa_1 + 10\kappa_2\kappa_1^3 + \kappa_1^5, \\
 \mu'_6 &= \kappa_6 + 6\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + 15\kappa_4\kappa_1^2 + 10\kappa_3^2 + 60\kappa_3\kappa_2\kappa_1 + \\
 &\quad + 20\kappa_3\kappa_1^3 + 15\kappa_2^3 + 45\kappa_2^2\kappa_1^2 + 15\kappa_2\kappa_1^4 + \kappa_1^6, \\
 \mu'_7 &= \kappa_7 + 7\kappa_6\kappa_1 + 21\kappa_5\kappa_2 + 21\kappa_5\kappa_1^2 + 35\kappa_4\kappa_3 + 105\kappa_4\kappa_2\kappa_1 + \\
 &\quad + 35\kappa_4\kappa_1^3 + 70\kappa_3^2\kappa_1 + 105\kappa_3\kappa_2^2 + 210\kappa_3\kappa_2\kappa_1^2 + \\
 &\quad + 35\kappa_3\kappa_1^4 + 105\kappa_2^3\kappa_1 + 105\kappa_2^2\kappa_1^2 + 21\kappa_2\kappa_1^5 + \kappa_1^7, \\
 \mu'_8 &= \kappa_8 + 8\kappa_7\kappa_1 + 28\kappa_6\kappa_2 + 28\kappa_6\kappa_1^2 + 56\kappa_5\kappa_3 + 168\kappa_5\kappa_2\kappa_1 + \\
 &\quad + 56\kappa_5\kappa_1^3 + 35\kappa_4^2 + 280\kappa_4\kappa_3\kappa_1 + 210\kappa_4\kappa_2^2 + \\
 &\quad + 420\kappa_4\kappa_2\kappa_1^2 + 70\kappa_4\kappa_1^4 + 280\kappa_3^2\kappa_2 + 280\kappa_3^2\kappa_1^2 + \\
 &\quad + 840\kappa_3\kappa_2^2\kappa_1 + 560\kappa_3\kappa_2\kappa_1^3 + 56\kappa_3\kappa_1^5 + 105\kappa_2^4 + \\
 &\quad + 420\kappa_2^3\kappa_1^2 + 210\kappa_2^2\kappa_1^4 + 28\kappa_2\kappa_1^6 + \kappa_1^8, \\
 \mu'_9 &= \kappa_9 + 9\kappa_8\kappa_1 + 36\kappa_7\kappa_2 + 36\kappa_7\kappa_1^2 + 84\kappa_6\kappa_3 + 252\kappa_6\kappa_2\kappa_1 + \\
 &\quad + 84\kappa_6\kappa_1^3 + 126\kappa_5\kappa_4 + 504\kappa_5\kappa_3\kappa_1 + 378\kappa_5\kappa_2^2 + \\
 &\quad + 756\kappa_5\kappa_2\kappa_1^2 + 126\kappa_5\kappa_1^4 + 315\kappa_4^2\kappa_1 + 1260\kappa_4\kappa_3\kappa_2 + \\
 &\quad + 1260\kappa_4\kappa_3\kappa_1^2 + 1890\kappa_4\kappa_2^2\kappa_1 + 1260\kappa_4\kappa_2\kappa_1^3 + \\
 &\quad + 126\kappa_4\kappa_1^5 + 280\kappa_3^3 + 2520\kappa_3^2\kappa_2\kappa_1 + 840\kappa_3^2\kappa_1^3 + \\
 &\quad + 1260\kappa_3\kappa_2^3 + 3780\kappa_3\kappa_2^2\kappa_1 + 1260\kappa_3\kappa_2\kappa_1^4 + 84\kappa_3\kappa_1^6 + \\
 &\quad + 945\kappa_2^4\kappa_1 + 1260\kappa_2^3\kappa_1^3 + 378\kappa_2^2\kappa_1^5 + 36\kappa_2\kappa_1^7 + \kappa_1^9,
 \end{aligned} \tag{3.37}$$

*) Коэффициенты в выражениях μ через κ те же самые, что и в представлении унитарных симметрических функций (1^n) через расширенные симметрические функции, и могут быть найдены по таблицам Дэвид и Кендалла

$$\begin{aligned}
 \mu'_{10} = & \kappa_{10} + 10\kappa_3\kappa_1 + 45\kappa_8\kappa_2 + 45\kappa_8\kappa_1^2 + 120\kappa_7\kappa_3 + \\
 & + 360\kappa_7\kappa_2\kappa_1 + 120\kappa_7\kappa_1^3 + 210\kappa_6\kappa_4 + 840\kappa_6\kappa_3\kappa_1 + \\
 & + 630\kappa_6\kappa_2^2 + 1260\kappa_6\kappa_2\kappa_1^2 + 210\kappa_6\kappa_1^4 + 126\kappa_5^2 + \\
 & + 1260\kappa_5\kappa_4\kappa_1 + 2520\kappa_5\kappa_3\kappa_2 + 2520\kappa_5\kappa_3\kappa_1^2 + \\
 & + 3780\kappa_5\kappa_2^2\kappa_1 + 2520\kappa_5\kappa_2\kappa_1^3 + 252\kappa_5\kappa_1^5 + 1575\kappa_4^2\kappa_2 + \\
 & + 1575\kappa_4^2\kappa_1^2 + 2100\kappa_4\kappa_3^2 + 12\,600\kappa_4\kappa_3\kappa_2\kappa_1 + \\
 & + 4200\kappa_4\kappa_3\kappa_1^3 + 3150\kappa_4\kappa_2^3 + 9450\kappa_4\kappa_2^2\kappa_1^2 + \\
 & + 3150\kappa_4\kappa_2\kappa_1^4 + 210\kappa_4\kappa_1^6 + 2800\kappa_3^3\kappa_1 + 6300\kappa_3^2\kappa_2^2 + \\
 & + 12\,600\kappa_3^2\kappa_2\kappa_1^2 + 2100\kappa_3^2\kappa_1^4 + 12\,600\kappa_3\kappa_2^3\kappa_1 + \\
 & + 12\,600\kappa_3\kappa_2^2\kappa_1^3 + 2520\kappa_3\kappa_2\kappa_1^5 + 120\kappa_3\kappa_1^7 + 945\kappa_2^5 + \\
 & + 4725\kappa_2^4\kappa_1^2 + 3150\kappa_2^3\kappa_1^4 + 630\kappa_2^2\kappa_1^6 + 45\kappa_2\kappa_1^8 + \kappa_1^{10}.
 \end{aligned} \quad (3.37)$$

Если моменты брать относительно среднего значения ($\kappa_1 = 0$), то

$$\begin{aligned}
 \mu_2 &= \kappa_2, \\
 \mu_3 &= \kappa_3, \\
 \mu_4 &= \kappa_4 + 3\kappa_2^2, \\
 \mu_5 &= \kappa_5 + 10\kappa_3\kappa_2, \\
 \mu_6 &= \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3, \\
 \mu_7 &= \kappa_7 + 21\kappa_5\kappa_2 + 35\kappa_4\kappa_3 + 105\kappa_3\kappa_2^2, \\
 \mu_8 &= \kappa_8 + 28\kappa_6\kappa_2 + 56\kappa_5\kappa_3 + 35\kappa_4^2 + 210\kappa_4\kappa_2^2 + \\
 & \quad + 280\kappa_3^2\kappa_2 + 105\kappa_2^4, \\
 \mu_9 &= \kappa_9 + 36\kappa_7\kappa_2 + 84\kappa_6\kappa_3 + 126\kappa_5\kappa_4 + 378\kappa_5\kappa_2^2 + \\
 & \quad + 1260\kappa_4\kappa_3\kappa_2 + 280\kappa_3^3 + 1260\kappa_3\kappa_2^3, \\
 \mu_{10} &= \kappa_{10} + 45\kappa_8\kappa_2 + 120\kappa_7\kappa_3 + 210\kappa_6\kappa_4 + 630\kappa_6\kappa_2^2 + \\
 & \quad + 126\kappa_5^2 + 2520\kappa_5\kappa_3\kappa_2 + 1575\kappa_4^2\kappa_2 + 2100\kappa_4\kappa_3^2 + \\
 & \quad + 3150\kappa_4\kappa_2^3 + 6300\kappa_3^2\kappa_2^2 + 945\kappa_2^5.
 \end{aligned} \quad (3.38)$$

Обратно,

$$\frac{\kappa_1 t}{1!} + \dots + \frac{\kappa_r t^r}{r!} + \dots = \log \left(1 + \frac{\mu'_1 t}{1!} + \dots + \frac{\mu'_r t^r}{r!} + \dots \right).$$

(1949) для всех порядков до 12. С помощью этих таблиц можно также получить обратные выражения, производя домножение на соответствующие факториалы.

Разлагая логарифм и собирая, как и выше, члены с одинаковой степенью t^r , имеем

$$\kappa_r = r! \sum_{m=1}^r \sum \left(\frac{\mu'_{p_1}}{p_1!} \right)^{\pi_1} \dots \left(\frac{\mu'_{p_m}}{p_m!} \right)^{\pi_m} \frac{(-1)^{\rho-1} (\rho-1)!}{\pi_1! \dots \pi_m!}, \quad (3.39)$$

где второе суммирование распространяется на все положительные π и ρ , подчиненные условиям (3.34), и

$$\pi_1 + \pi_2 + \dots + \pi_m = \rho. \quad (3.40)$$

Вот первые десять формул, выражающих семиинварианты через моменты относительно произвольной точки:

$$\left. \begin{aligned} \kappa_1 &= \mu'_1, \\ \kappa_2 &= \mu'_2 - \mu_1'^2, \\ \kappa_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3, \\ \kappa_4 &= \mu'_4 - 4\mu'_3\mu'_1 - 3\mu_2'^2 + 12\mu'_2\mu_1'^2 - 6\mu_1'^4, \\ \kappa_5 &= \mu'_5 - 5\mu'_4\mu'_1 - 10\mu'_3\mu'_2 + 20\mu'_3\mu_1'^2 + \\ &\quad + 30\mu_2'^2\mu'_1 - 60\mu'_2\mu_1'^3 + 24\mu_1'^5, \\ \kappa_6 &= \mu'_6 - 6\mu'_5\mu'_1 - 15\mu'_4\mu'_2 + 30\mu'_4\mu_1'^2 - 10\mu_3'^2 + 120\mu'_3\mu'_2\mu'_1 - \\ &\quad - 120\mu'_3\mu_1'^3 + 30\mu_2'^3 - 270\mu_2'^2\mu_1'^2 + 360\mu'_2\mu_1'^4 - 120\mu_1'^6, \\ \kappa_7 &= \mu'_7 - 7\mu'_6\mu'_1 - 21\mu'_5\mu'_2 + 42\mu'_5\mu_1'^2 - 35\mu'_4\mu'_3 + 210\mu'_4\mu'_2\mu'_1 - \\ &\quad - 210\mu'_4\mu_1'^3 + 140\mu_3'^2\mu'_1 + 210\mu'_3\mu_2'^2 - 1260\mu'_3\mu'_2\mu_1'^2 + \\ &\quad + 840\mu'_3\mu_1'^4 - 630\mu_2'^3\mu'_1 + 2520\mu_2'^2\mu_1'^3 - \\ &\quad - 2520\mu'_2\mu_1'^5 + 720\mu_1'^7, \\ \kappa_8 &= \mu'_8 - 8\mu'_7\mu'_1 - 28\mu'_6\mu'_2 + 56\mu'_6\mu_1'^2 - 56\mu'_5\mu'_3 + 336\mu'_5\mu'_2\mu'_1 - \\ &\quad - 336\mu'_5\mu_1'^3 - 35\mu_4'^2 + 560\mu'_4\mu'_3\mu'_1 + 420\mu'_4\mu_2'^2 - \\ &\quad - 2520\mu'_4\mu'_2\mu_1'^2 + 1680\mu'_4\mu_1'^4 + 560\mu_3'^2\mu'_2 - \\ &\quad - 1680\mu_3'^2\mu_1'^2 - 5040\mu'_3\mu_2'^2\mu'_1 + 13440\mu'_3\mu'_2\mu_1'^3 - \\ &\quad - 6720\mu'_3\mu_1'^5 - 630\mu_2'^4 + 10080\mu_2'^3\mu_1'^2 - \\ &\quad - 25200\mu_2'^2\mu_1'^4 + 20160\mu'_2\mu_1'^6 - 5040\mu_1'^8, \end{aligned} \right\} \quad (3.41)$$

$$\begin{aligned}
\kappa_9 = & \mu'_9 - 9\mu'_8\mu'_1 - 36\mu'_7\mu'_2 + 72\mu'_7\mu'^2_1 - 84\mu'_6\mu'_3 + 504\mu'_6\mu'_2\mu'_1 - \\
& - 504\mu'_6\mu'^3_1 - 126\mu'_5\mu'_4 + 1008\mu'_5\mu'_3\mu'_1 + 756\mu'_5\mu'^2_2 - \\
& - 4536\mu'_5\mu'_2\mu'^2_1 + 3024\mu'_5\mu'^4_1 + 630\mu'^2_4\mu'_1 + \\
& + 2520\mu'_4\mu'_3\mu'_2 - 7560\mu'_4\mu'_3\mu'^2_1 - 11\,340\mu'_4\mu'^2_2\mu'_1 + \\
& + 30\,240\mu'_4\mu'_2\mu'^3_1 - 15\,120\mu'_4\mu'^5_1 + 560\mu'^3_3 - \\
& - 15\,120\mu'^2_3\mu'_2\mu'_1 + 20\,160\mu'^2_3\mu'^3_1 - 7\,560\mu'_3\mu'^3_2 + \\
& + 90\,720\mu'_3\mu'^2_2\mu'^2_1 - 151\,200\mu'_3\mu'_2\mu'^4_1 + 60\,480\mu'_3\mu'^6_1 + \\
& + 22\,680\mu'^4_2\mu'_1 - 151\,200\mu'^3_2\mu'^3_1 + 272\,160\mu'^2_2\mu'^5_1 - \\
& - 181\,440\mu'_2\mu'^7_1 + 40\,320\mu'^9_1, \\
\kappa_{10} = & \mu'_{10} - 10\mu'_9\mu'_1 - 45\mu'_8\mu'_2 + 90\mu'_8\mu'^2_1 - 120\mu'_7\mu'_3 + \\
& + 720\mu'_7\mu'_2\mu'_1 - 720\mu'_7\mu'^3_1 - 210\mu'_6\mu'_4 + 1680\mu'_6\mu'_3\mu'_1 + \\
& + 1260\mu'_6\mu'^2_2 - 7560\mu'_6\mu'_2\mu'^2_1 + 5040\mu'_6\mu'^4_1 - 126\mu'^2_5 + \\
& + 2520\mu'_5\mu'_4\mu'_1 + 5040\mu'_5\mu'_3\mu'_2 - 15\,120\mu'_5\mu'_3\mu'^2_1 - \\
& - 22\,680\mu'_5\mu'^2_2\mu'_1 + 60\,480\mu'_5\mu'_2\mu'^3_1 - 30\,240\mu'_5\mu'^5_1 + \\
& + 3150\mu'^2_4\mu'_2 - 9450\mu'^2_4\mu'^2_1 + 4200\mu'_4\mu'^2_3 - \\
& - 75\,600\mu'_4\mu'_3\mu'_2\mu'_1 + 100\,800\mu'_4\mu'_3\mu'^3_1 - 18\,900\mu'_4\mu'^3_2 + \\
& + 226\,800\mu'_4\mu'^2_2\mu'^2_1 - 378\,000\mu'_4\mu'_2\mu'^4_1 + 151\,200\mu'_4\mu'^6_1 - \\
& - 16\,800\mu'^3_3\mu'_1 - 37\,800\mu'^2_3\mu'^2_2 + 302\,400\mu'^2_3\mu'_2\mu'^2_1 - \\
& - 252\,000\mu'^2_3\mu'^4_1 + 302\,400\mu'_3\mu'^3_2\mu'_1 - \\
& + 1\,512\,000\mu'_3\mu'^2_2\mu'^3_1 + 1\,814\,400\mu'_3\mu'_2\mu'^5_1 - \\
& - 604\,800\mu'_3\mu'^7_1 + 22\,680\mu'^5_2 - 567\,000\mu'^4_2\mu'^2_1 + \\
& + 2\,268\,000\mu'^3_2\mu'^4_1 - 3\,175\,200\mu'^2_2\mu'^6_1 + \\
& + 1\,814\,400\mu'_2\mu'^8_1 - 362\,880\mu'^{10}_1.
\end{aligned} \tag{3.42}$$

Если моменты брать относительно среднего значения, то

$$\begin{aligned}
 \kappa_2 &= \mu_2, \\
 \kappa_3 &= \mu_3, \\
 \kappa_4 &= \mu_4 - 3\mu_2^2, \\
 \kappa_5 &= \mu_5 - 10\mu_3\mu_2, \\
 \kappa_6 &= \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3, \\
 \kappa_7 &= \mu_7 - 21\mu_5\mu_2 - 35\mu_4\mu_3 + 210\mu_3\mu_2^2, \\
 \kappa_8 &= \mu_8 - 28\mu_6\mu_2 - 56\mu_5\mu_3 - 35\mu_4^2 + 420\mu_4\mu_2^2 + \\
 &\quad + 560\mu_3^2\mu_2 - 630\mu_2^4, \\
 \kappa_9 &= \mu_9 - 36\mu_7\mu_2 - 84\mu_6\mu_3 - 126\mu_5\mu_4 + 756\mu_5\mu_2^2 + \\
 &\quad + 2520\mu_4\mu_3\mu_2 + 560\mu_3^3 - 7560\mu_3\mu_2^2, \\
 \kappa_{10} &= \mu_{10} - 45\mu_8\mu_2 - 120\mu_7\mu_3 - 210\mu_6\mu_4 + 1260\mu_6\mu_2^2 - \\
 &\quad - 126\mu_5^2 + 5040\mu_5\mu_3\mu_2 + 3150\mu_4^2\mu_2 + 4200\mu_4\mu_3^2 - \\
 &\quad - 18900\mu_4\mu_2^3 - 37800\mu_3^2\mu_2^2 + 22680\mu_2^5.
 \end{aligned} \quad (3.43)$$

Существование семиинвариантов

3.15 Формально написанное выражение (3.30) можно рассматривать как определяющее семиинварианты через моменты. В этом случае очевидно, что если моменты порядка r и ниже существуют, то существует и семиинвариант порядка r . Если же определять семиинварианты как коэффициенты при $(it)^r/r!$, $r=1, 2, \dots$ в разложении логарифма х. ф., то вывести существование r -го семиинварианта из существования r -го и низших моментов несколько сложнее.

Разлагая аналитическую функцию e^{itx} в выражении для характеристической функции, получим

$$\begin{aligned}
 \varphi(t) &= \int_{-\infty}^{\infty} \left\{ \sum_0^r \frac{(itx)^j}{j!} + O\{|tx|^{r+1}/(r+1)!\} \right\} dF = \\
 &= \sum_{j=0}^r \mu'_j (it)^j / j! + R_r, \quad (3.44)
 \end{aligned}$$

где остаточный член R_r равен $\rho_{r+1} |t|^{r+1}/(r+1)!$, $|\rho| \leq 1$. Поэтому

$$\varphi(t) = \sum_{j=0}^r \mu'_j (it)^j / j! + o(t^r).$$

При достаточно малых t существует $\log \varphi(t)$. Разлагая его, имеем

$$\log \varphi(t) = \sum_{j=0}^r \kappa_j (it)^j / j! + o(t^r), \quad (3.45)$$

где коэффициенты κ_j — семиинварианты в соответствии с самим их определением. Таким образом, если ν_{r+1} существует, то существует κ_r и все семиинварианты более низкого порядка. Можно было бы получить более сильные результаты, однако и этого достаточно для многих целей *).

Вычисление семиинвариантов

3.16 Семиинварианты не вычисляются, подобно моментам, непосредственным суммированием или интегрированием. Для их нахождения нужно или вычислить моменты и воспользоваться соотношениями (3.43), или использовать характеристическую функцию. Следующие примеры иллюстрируют применяемые методы.

Пример 3.9

В примере 3.7 мы нашли следующие значения моментов для распределения, заданного таблицей 1.7:

$$\mu'_1 = 11,020\ 850,$$

$$\mu_2 = 6,616\ 805,$$

$$\mu_3 = -0,207\ 840,$$

$$\mu_4 = 137,689\ 185.$$

Для нахождения семиинвариантов воспользуемся соотношениями (3.43); κ_2 и κ_3 совпадают соответственно с μ_2 и μ_3 , а для κ_4 получаем

$$\kappa_4 = \mu_4 - 3\mu_2^2 = 6,342\ 86.$$

Семиинвариант κ_1 совпадает с моментом μ'_1 , который в нашем случае вычислен относительно середины интервала 57 — дюймов.

Конечно, получились бы те же результаты, если бы мы воспользовались соотношениями (3.42) и моментами относительно произвольного начала отсчета.

*) В действительности полученное разложение для $\varphi(t)$ с остаточным членом $o(t^r)$ справедливо, если предполагать только, что r -й момент существует, так что из существования μ'_r следует существование всех семиинвариантов до порядка r включительно. (Прим. перев.)

Пример 3.10

Рассмотрим дискретное распределение, для которого частоты значений $0, 1, \dots, j, \dots$ суть $e^{-\lambda} \left(1, \frac{\lambda}{1!}, \dots, \frac{\lambda^j}{j!}, \dots\right)$. Характеристическая функция его равна

$$\varphi(t) = e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} e^{ijt} = e^{-\lambda} \exp(\lambda e^{it}) = \exp\{\lambda(e^{it} - 1)\}.$$

Поскольку распределение сосредоточено на неотрицательной полуоси, то все абсолютные моменты относительно нуля совпадают с обычными моментами. Имеем

$$\mu_r' = e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j j^r}{j!}.$$

При любом $r^*)$ такой ряд сходится, поэтому семинварианты всех порядков существуют и могут быть найдены из разложения $\log \varphi(t)$ в ряд по степеням t . Но

$$\log \varphi(t) = \lambda(e^{it} - 1) = \lambda \sum_{j=1}^{\infty} (it)^j / j!$$

и, следовательно,

$$\kappa_r = \lambda$$

для всех r . Таким образом, все семинварианты этого распределения равны λ .

Пример 3.11

В примере 3.4 мы нашли, по существу, характеристическую функцию распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx, \quad -\infty \leq x \leq \infty.$$

Она равна

$$\varphi(t) = \exp\left(-\frac{1}{2} \sigma^2 t^2\right),$$

и поэтому

$$\log \varphi(t) = -\frac{1}{2} \sigma^2 t^2.$$

*) Так как отношение $(n+1)$ -го члена ряда к n -му равно

$$\frac{\lambda^{n+1} (n+1)^r / \lambda^n n^r}{(n+1)! / n!} = \frac{\lambda}{n+1} \left(1 + \frac{1}{n}\right)^r = O\left(\frac{\lambda}{n}\right),$$

то ряд сходится при всех конечных λ .

Легко видеть, что для рассматриваемого распределения абсолютные моменты всех порядков, а следовательно, и семинварианты, существуют и

$$\begin{aligned}\kappa_2 &= \sigma^2, \\ \kappa_r &= 0, \quad r > 2.\end{aligned}$$

Пример 3.12

В примере 3.6 было показано, что характеристическая функция распределения

$$dF = \frac{a^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-ax} dx, \quad 0 \leq x < \infty, \quad a, \gamma > 0,$$

равна

$$\varphi(t) = \frac{1}{\left(1 - \frac{it}{a}\right)^\gamma}.$$

Легко проверить, что семинварианты всех порядков у этого распределения существуют. При этом

$$\kappa_r = \text{коэфф. при } \frac{(it)^r}{r!} \text{ в разложении } -\gamma \log\left(1 - \frac{it}{a}\right) = \gamma(r-1)! a^{-r}.$$

Пример 3.13

Рассмотрим опять распределение, изучавшееся в примере 3.3:

$$dF = \frac{k}{(1+x^2)^m} dx, \quad -\infty < x < \infty, \quad m \geq 1.$$

Характеристическая функция его

$$\varphi(t) = k \int_{-\infty}^{\infty} \frac{e^{itx}}{(1+x^2)^m} dx$$

в силу нечетности функции $\sin xt$ равна

$$k \int_{-\infty}^{\infty} \frac{\cos xt}{(1+x^2)^m} dx.$$

Этот интеграл можно вычислить, переходя в комплексную плоскость и интегрируя по контуру, состоящему из оси x , бесконечно большой полуокружности над осью x и бесконечно малой окружности вокруг точки $x=i$. В результате получается, что

$$\begin{aligned}\varphi(t) &= \frac{k\pi}{2^{2m-2}(m-1)!} e^{-|t|} \left\{ (2|t|)^{m-1} + m(m-1)(2|t|)^{m-2} + \right. \\ &\quad \left. + \frac{(m+1)m(m-1)(m-2)}{2!} (2|t|)^{m-3} + \dots + \frac{(2m-2)!}{(m-1)!} \right\}.\end{aligned}$$

Если $r < 2m - 1$, то абсолютный момент

$$v_r = k \int_{-\infty}^{\infty} \frac{|x|^r dx}{(1+x^2)^m}$$

(а следовательно, и семинвариант) порядка r существует. Однако в этом случае $\log \varphi(t)$ невозможно разложить в *бесконечный* ряд по степеням t , как это могло бы показаться, исходя из вида $\varphi(t)$; $\log \varphi(t)$ можно разложить лишь в конечный ряд с числом членов, равным числу производных функции $\varphi(t)$ при $t=0$.

Рассмотрим, для простоты, случай $m=2$. Поскольку в этом случае $k=2/\pi$, мы получаем

$$\varphi(t) = e^{-|t|} \{ |t| + 1 \},$$

$$\log \varphi(t) = -|t| + \log \{ 1 + |t| \}.$$

Если t положительно, то $\log \varphi(t)$ равен

$$-\frac{1}{2}t^2 + \frac{1}{3}t^3 - \dots,$$

а если t отрицательно, то

$$-\frac{1}{2}t^2 - \frac{1}{3}t^3 - \dots$$

Эти два выражения отличаются знаками перед нечетными степенями t . Отсюда следует, что не существует единственного разложения $\log \varphi(t)$ в ряд по степеням t в окрестности точки $t=0$. Имеется два вида таких функций, представляющих $\log \varphi(t)$, в соответствии с тем, положительно t или отрицательно.

Однако эти разложения совпадают вплоть до членов, содержащих t^2 , и первая и вторая производные от $\log \varphi(t)$ однозначно определены при $t=0$. Поэтому первый и второй семинварианты существуют:

$$\kappa_1=0, \quad \kappa_2=1.$$

Семинварианты более высоких порядков не существуют.

Факториальные семинварианты

3.17 По аналогии с соотношением между семинвариантами и моментами можно определить факториальные семинварианты $\kappa_{[r]}$ как коэффициенты при $t^r/r!$ в разложении логарифма производящей функции факториальных моментов. Так что, если $P(t)$ — производящая функция, то

$$\omega(t) = \log P(1+t) \tag{3.46}$$

будет производящей функцией факториальных семинвариантов. Подобно производящей функции факториальных моментов, функция $\omega(t)$ полезна главным образом при изучении некоторых классов дискретных распределений.

Соотношения между семиинвариантами и факториальными семиинвариантами подобны тем, которые имеются между моментами и факториальными моментами. Беря в качестве шага единицу, имеем аналогично (3.25) и (3.26)

$$\left. \begin{aligned} \kappa_{[1]} &= \kappa_1, \\ \kappa_{[2]} &= \kappa_2 - \kappa_1, \\ \kappa_{[3]} &= \kappa_3 - 3\kappa_2 + 2\kappa_1, \\ \kappa_{[4]} &= \kappa_4 - 6\kappa_3 + 11\kappa_2 - 6\kappa_1. \end{aligned} \right\} \quad (3.47)$$

Обратно,

$$\left. \begin{aligned} \kappa_2 &= \kappa_{[2]} + \kappa_{[1]}, \\ \kappa_3 &= \kappa_{[3]} + 3\kappa_{[2]} + \kappa_{[1]}, \\ \kappa_4 &= \kappa_{[4]} + 6\kappa_{[3]} + 7\kappa_{[2]} + \kappa_{[1]}. \end{aligned} \right\} \quad (3.48)$$

Пример 3.14

Для биномиального распределения имеем

$$\omega(t) = n \log(1 + pt)$$

и, следовательно,

$$\kappa_{[r]} = (-1)^{r+1} n(r-1)! p^r.$$

Поправки к группировке

3.18 Если распределение задано численно по сгруппированному данным, то его моменты несколько отличаются от моментов несгруппированного распределения вследствие того, что при группировке частоты предполагаются сосредоточенными в средних точках интервалов. При некоторых условиях возможна поправка, уменьшающая это расхождение.

Пусть $f(x)$ — непрерывная функция плотности распределения, сосредоточенного на интервале (a, b) . Разобьем этот интервал на n частей длины h каждая. Частота в j -м интервале, центр которого находится в точке $x_j = a + \left(j - \frac{1}{2}\right)h$, дается формулой

$$f_j = \int_{-\frac{1}{2}h}^{\frac{1}{2}h} f(x_j + \xi) d\xi. \quad (3.49)$$

Моменты группированного распределения, при обозначении которых здесь и ниже мы пользуемся чертой, равны

$$\bar{\mu}'_r = \sum_{j=1}^n x_j^r \int_{-\frac{1}{2}h}^{\frac{1}{2}h} f(x_j + \xi) d\xi. \quad (3.50)$$

Если $y(x)$ — функция с $2m$ производными, то согласно формуле Эйлера — Маклорена *)

$$\frac{1}{h} \int_a^b y(x) dx = \left[y \left(a + \frac{1}{2} h \right) + y \left(a + \frac{3}{2} h \right) + \dots + y \left\{ a + \left(n - \frac{1}{2} \right) h \right\} \right] - \sum_{j=1}^m \frac{h^{2j-1}}{(2j)!} B_{2j} \left(\frac{1}{2} \right) [y^{(2j-1)}(x)]_a^b - S_{2m}, \quad (3.51)$$

где $y^{(r)}$ есть r -я производная от y , $B_{2m} \left(\frac{1}{2} \right)$ — значение $(2m)$ -го полинома Бернулли в точке $\frac{1}{2}$ и S_{2m} — остаточный член, который можно представить в виде

$$S_{2m} = \frac{nh^{2m}}{(2m)!} B_{2m} \left(\frac{1}{2} \right) y^{(2m)}(a + nh\theta), \quad 0 < \theta < 1. \quad (3.52)$$

Если первые $(2m - 1)$ производных функции y на концах интервала обращаются в нуль, то, пренебрегая остатком S_{2m} и полагая в (3.51)

$$y(x) = x^r \int_{-1/2h}^{1/2h} f(x + \xi) d\xi,$$

имеем

$$\frac{1}{h} \int_a^b x^r \int_{-1/2h}^{1/2h} f(x + \xi) d\xi dx = \bar{\mu}'_r. \quad (3.53)$$

Следовательно,

$$\begin{aligned} \bar{\mu}'_r &= \frac{1}{h} \int_a^b \int_{-1/2h}^{1/2h} (u - \xi)^r f(u) du d\xi = \\ &= \frac{1}{h} \int_a^b \frac{\left(u + \frac{1}{2}h\right)^{r+1} - \left(u - \frac{1}{2}h\right)^{r+1}}{r+1} f(u) du = \\ &= \sum_{j=0}^{[1/2r]} \binom{r}{2j} \left(\frac{1}{2}h\right)^{2j} \frac{1}{2j+1} \mu'_{r-2j}, \end{aligned} \quad (3.54)$$

где $[1/2r]$ — целая часть $1/2r$. Тем самым найдено представление моментов $\bar{\mu}'_r$ через действительные моменты. Практически, од-

*) См. А. О. Гельфонд, Исчисление конечных разностей, М., 1952, гл. IV, или Г. Крамер, Математические методы статистики, М., ИП, 1948, гл. 12. (Прим. перев.)

нако, нас интересуют выражения действительных моментов через моменты $\bar{\mu}'_r$. Из (3.54) легко находим

$$\left. \begin{aligned} \mu'_1 &= \bar{\mu}'_1, \\ \mu'_2 &= \bar{\mu}'_2 - \frac{1}{12} h^2, \\ \mu'_3 &= \bar{\mu}'_3 - \frac{1}{4} \bar{\mu}'_1 h^2, \\ \mu'_4 &= \bar{\mu}'_4 - \frac{1}{2} \bar{\mu}'_2 h^2 + \frac{7}{400} h^4, \\ \mu'_5 &= \bar{\mu}'_5 - \frac{5}{6} \bar{\mu}'_3 h^2 + \frac{7}{48} \bar{\mu}'_1 h^4, \\ \mu'_6 &= \bar{\mu}'_6 - \frac{5}{4} \bar{\mu}'_4 h^2 + \frac{7}{16} \bar{\mu}'_2 h^4 - \frac{31}{1344} h_6. \end{aligned} \right\} \quad (3.55)$$

Соответствующая общая формула имеет вид

$$\mu'_r = \sum_{j=0}^r \left\{ \binom{r}{j} (2^{j-1} - 1) B_j h^j \bar{\mu}'_{r-j} \right\}, \quad (3.56)$$

где B_j — число Бернулли порядка j (Волд, 1934а).

3.19 Мы сделали целый ряд допущений при выводе этих формул: высокий порядок касания в точках a и b , конечность интервала (a, b) , возможность пренебречь остаточным членом. По абсолютному значению величина $B_{2m} \left(\frac{1}{2} \right) / (2m)!$ меньше, чем $4/(2\pi)^{2m}$, и, следовательно, $|S_{2m}|$ в (3.52) меньше, чем величина $4nh^{2m}/(2\pi)^{2m}$, умноженная на некоторое значение $y^{(2m)}(x)$ в интервале (a, b) . Степень пренебрежимости остаточным членом зависит поэтому от поведения функции $f^{(2m-1)}(x)$ в интервале (a, b) .

Практически рассмотренные поправки ведут к уточнению, если имеется касание достаточно высокого порядка. В противном случае они могут вести к еще большему расхождению. Для распределений бесконечного размаха частоты убывают вблизи концов, и с некоторой уверенностью мы можем применять поправки к тем эмпирическим распределениям (практически, конечно, они могут иметь лишь конечный размах), основная часть которых сосредоточена в середине и частоты на концах малы.

Пример 3.15

Рассмотрим распределение

$$dF = \frac{1}{B(12, 6)} x^{11} (1-x)^5 dx, \quad 0 \leq x \leq 1.$$

Это случай так называемого распределения I типа. Значения отвечающих ему частот для интервалов разбиения с шагом 0,1 могут быть найдены по таблице неполной В-функции. Они равны

Центр интервала	Частота
0,05	0,000 000 0
0,15	0,000 009 2
0,25	0,000 646 8
0,35	0,009 938 2
0,45	0,061 137 4
0,55	0,192 199 6
0,65	0,332 887 7
0,75	0,297 479 9
0,85	0,101 033 7
0,95	0,004 667 5
Полная частота	1,000 000 0

В следующей таблице приведены моменты относительно $x=0$ группированного распределения без поправок на группировку, с поправкой, а также точные моменты.

Момент	Групповой *)	Точный	Исправленный
μ'_1	0,666 662 8	0,666 666 7	0,666 662 8
μ'_2	0,456 965 5	0,456 140 4	0,456 132 2
μ'_3	0,320 952 3	0,319 298 2	0,319 285 7
μ'_4	0,230 335 1	0,228 070 2	0,228 053 2
μ'_5	0,168 512 9	0,165 869 2	0,165 848 0
μ'_6	0,125 433 2	0,122 599 0	0,122 574 0

*) По-английски групповые моменты (т. е. моменты группированного распределения) называются *raw moments* — грубые моменты. (Прим. перев.)

Точные значения вычислялись непосредственным интегрированием. Следует отметить, что порядок малости погрешностей второго и третьего исправленных моментов выше порядка членов $h^2/12$ и $h^2/4$, использованных при вычислении поправок. Главная часть погрешности четвертого исправленного момента составляет $2 \cdot 10^{-5}$ и является величиной того же порядка, что и слагаемое $7h^4/240$, использованное при построении поправки для

четвертого момента. Точно так же ошибки исправленных пятого и шестого моментов имеют тот же или более высокий порядок малости, чем некоторые из поправочных членов.

Пример 3.16

Для иллюстрации того обстоятельства, что поправки Шеппарда могут вести к еще большему расхождению, если нарушено условие касания высокого порядка, приведем следующий пример, заимствованный из работы Пэрмэна и Пирсона (1919). Приводимая ниже таблица содержит частоты нормального распределения

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

на отрезке [1,25; 5,25] при группировке с шагом 0,5.

Центр интервала	Частота
1,5	0,065 591
2,0	0,027 834
2,5	0,009 245
3,0	0,002 402
3,5	0,000 489
4,0	0,000 078
4,5	0,000 010
5,0	0,000 001
Полная частота	0,105 650

После приведения полной частоты к единице получим J-образное распределение, имеющее высокий порядок касания на правом конце, а на левом — резко обрывающееся.

В следующей таблице представлены моменты относительно среднего без поправок, с поправками Шеппарда, а также истинные моменты этого распределения.

Момент	Групповой	Точный	Исправленный
μ_2	0,158 524	0,172 222	0,137 691
μ_3	0,104 226	0,098 612	0,104 226
μ_4	0,149 090	0,156 405	0,131 697

В случае второго и четвертого момента поправки ведут к еще большему расхождению. Для четвертого момента поправка увеличивает относительную погрешность приближенного значения (округленно) с 5 до 17%. Ясно, что по крайней мере для

такого грубого группирования, как в приведенном примере, поправки Шеппарда могут совершенно себя не оправдывать *).

3.20 Соотношения (3.55) выписаны в терминах моментов относительно произвольной точки. Этой точкой может быть, в частности, среднее значение распределения. Соответственно, чтобы получить поправки относительно среднего, нужно в (3.55) отбросить штрихи и μ'_1 положить равным нулю.

Осредненные поправки

3.21 К поправкам Шеппарда ведет также другая постановка задачи. Пусть имеется распределение, сосредоточенное на неизвестном отрезке, и пусть известны частоты попаданий в заданные интервалы. Можно спросить, какие следует сделать поправки к групповым моментам, чтобы они стали более близкими к истинным моментам *в среднем*? Иными словами, каковы средние значения групповых моментов, если множество интервалов группирования располагается случайно?

Пусть X_j , $j=0, \pm 1, \pm 2$, — некоторое фиксированное множество значений x и пусть x_j изменяется от $X_j - 1/2h$ до $X_j + 1/2h$.

По определению

$$\bar{\mu}'_r = \sum_{j=-\infty}^{\infty} \left\{ x_j^r \int_{-1/2h}^{1/2h} f(x_j + \xi) d\xi \right\}.$$

Обозначая $M(\bar{\mu}'_r)$ среднее значение $\bar{\mu}'_r$ по всем x_j , меняющимся от $X_j - 1/2h$ до $X_j + 1/2h$, имеем

$$\begin{aligned} M(\bar{\mu}'_r) &= \frac{1}{h} \sum_{j=-\infty}^{\infty} \int_{X_j-1/2h}^{X_j+1/2h} x_j^r \int_{-1/2h}^{1/2h} f(x_j + \xi) d\xi dx_j = \\ &= \frac{1}{h} \int_{-\infty}^{\infty} x^r \int_{-1/2h}^{1/2h} f(x + \xi) d\xi dx. \end{aligned} \quad (3.57)$$

Соотношение (3.57) отличается от (3.53) лишь тем, что в нем вместо $\bar{\mu}'_r$ стоит $M(\bar{\mu}'_r)$. Таким образом, поправки Шеппарда применимы к осредненным групповым моментам независимо от характера поведения функции плотности на концах.

Это не дает, однако, достаточного основания для свободного применения поправок Шеппарда во всех случаях. Вместо усло-

*) В рассмотренном примере поправки Шеппарда ухудшают точность оценок не из-за грубости группирования (оно выполнено вполне корректно), а из-за того, что исследуемое распределение имеет J-образную форму. (Прим. ред.)

вий высокого порядка касания на концах и ограниченности производных, обеспечивающих применимость поправок к любому частному распределению, здесь появляется условие случайного расположения интервалов группировки на том отрезке, где сосредоточено распределение. И хотя при выполнении такого условия в каждом конкретном случае поправки могут оказаться ошибочными, однако в среднем по большому количеству случаев они будут правильными. В действительности условие о случайном расположении группировки очень часто оказывается неприменимым к J- и U-образным распределениям; например, в случае распределений доходов или смертности по возрастам почти неизбежно приходится начинать группирование в нуле.

3.22 Чтобы получить поправки для моментов относительно среднего, нельзя просто отбросить штрихи. Если обозначить среднее группированного распределения буквой m , то для среднего значения r -го момента относительно среднего получим выражение

$$M(\bar{\mu}_r) = \frac{1}{h} \int_{-\infty}^{\infty} (x - m)^r \int_{-1/2h}^{1/2h} f(x + \xi) d\xi dx,$$

где m — функция от x , так что проведенное выше в этой главе преобразование стоящего справа интеграла здесь незаконно. Точные выражения для осредненных поправок к моментам относительно среднего до сих пор не получены. Однако на основе изучения некоторых частных распределений Кендалл (1938) сделал вывод, что для всех обычных целей можно использовать соотношения (3.55), как если бы среднее было фиксированным.

3.23 Осредненные поправки можно также применять в случае дискретных распределений, когда производится группирование в большие интервалы. Эти поправки отличны от поправок в непрерывном случае. См. упражнения 3.13 и Крэйг (1936). О поправках в случае, когда условия Шеппарда не выполняются, см. Пэрман и Пирсон (1919), Сэндон (1924), Мартиш (1934) и Элдертон (1938b).

Поправки Шеппарда к факториальным моментам

3.24 Волдом (1934a) было показано, что поправки Шеппарда к факториальным моментам даются следующими формулами:

$$\left. \begin{aligned} \mu'_{[1]} &= \bar{\mu}'_{[1]}, \\ \mu'_{[2]} &= \bar{\mu}'_{[2]} - \frac{h^2}{12}, \\ \mu'_{[3]} &= \bar{\mu}'_{[3]} - \frac{h^2}{4} \bar{\mu}_{[1]} + \frac{h^3}{4}, \\ \mu'_{[4]} &= \bar{\mu}'_{[4]} - \frac{h^2}{2} \bar{\mu}'_{[2]} + h^3 \bar{\mu}'_{[1]} - \frac{71}{80} h^4, \\ \mu'_{[5]} &= \bar{\mu}'_{[5]} - \frac{5}{6} \bar{\mu}_{[3]} h^2 + \frac{5}{2} \bar{\mu}'_{[2]} h^3 - \frac{71}{16} \bar{\mu}'_{[1]} h^4 + \frac{31}{8} h^5, \\ \mu'_{[6]} &= \bar{\mu}'_{[6]} - \frac{5}{4} \bar{\mu}_{[4]} h^2 + 5 \bar{\mu}'_{[3]} h^3 - \frac{213}{16} \bar{\mu}'_{[2]} h^4 + \frac{93}{4} \bar{\mu}'_{[1]} h^5 - \frac{9129}{448} h^6. \end{aligned} \right\} \quad (3.58).$$

Общая формула имеет вид

$$\mu'_{|r|} = \sum_{j=0}^r \binom{r}{j} B_j^{(j+2)} \left(\frac{3}{2}\right) h^j \mu'_{|r-j|}, \quad (3.59)$$

где $B_j^{(j+2)} \left(\frac{3}{2}\right)$ — значение полинома Бернулли, равное при $j > 1$

$$(-1)^{j+1} \frac{(2j)!}{2^{2j}(j+1)!} \left(\frac{1}{3} + \frac{1}{5} + \dots + \frac{1}{2j-1} \right),$$

а при $j = 0, 1$ равное $B_0^{(2)} \left(\frac{3}{2}\right) = 1$, $B_1^{(3)} \left(\frac{3}{2}\right) = 0$.

Замечание о числах и полиномах Бернулли

3.25 Числа и (в несколько меньшей степени) полиномы Бернулли имеют ряд важных приложений в статистике. Число Бернулли B_j порядка j определяется как коэффициент при $t^j/j!$ в разложении $t/(e^t - 1)$, так что

$$\frac{t}{e^t - 1} = \sum_{j=0}^{\infty} \frac{B_j t^j}{j!}. \quad (3.60)$$

Непосредственное разложение левой части дает

$$B_0 = 1; \quad B_1 = -\frac{1}{2}; \quad B_{2j+1} = 0 \quad \text{при } j = 1, 2, \dots;$$

$$B_2 = \frac{1}{6}; \quad B_4 = -\frac{1}{30}; \quad B_6 = \frac{1}{42}; \quad B_8 = -\frac{1}{30}; \quad B_{10} = \frac{5}{66};$$

$$B_{12} = -\frac{691}{2730}; \quad B_{14} = \frac{7}{6}.$$

В дальнейшем нам потребуется следующее соотношение (одно из многих, содержащих числа Бернулли):

$$\log \left\{ \left(\operatorname{sh} \frac{1}{2} t \right) / \frac{1}{2} t \right\} = \sum_{j=2}^{\infty} \left(\frac{B_j t^j}{j! j} \right) = \sum_{j=1}^{\infty} \frac{B_{2j} t^{2j}}{(2j)! 2j}. \quad (3.61)$$

Оно получается интегрированием от 0 до t обеих частей следующего равенства, равносильного (3.60):

$$\frac{1}{e^t - 1} - \frac{1}{t} + \frac{1}{2} = \sum_{j=2}^{\infty} \frac{B_j t^{j-1}}{j!}.$$

Весьма часто используется также так называемый ряд Стирлинга:

$$\log \Gamma(x) = \left(x - \frac{1}{2}\right) \log x - x + \frac{1}{2} \log 2\pi + \sum_{j=1}^{\infty} \frac{B_{2j}}{2j(2j-1)x^{2j-1}}. \quad (3.62)$$

Это разложение асимптотично в том смысле, что, каково бы ни было n , для всех положительных действительных x истинное значение $\log \Gamma(x)$ лежит между суммой из n и $n+1$ членов. В частности,

$$\log \Gamma(x) = \left(x - \frac{1}{2}\right) \log x - x + \frac{1}{2} \log 2\pi + \frac{1}{12x} - \frac{1}{360x^3} + \frac{1}{1260x^5} - \frac{1}{1680x^7} + \frac{1}{1188x^9} - \dots \quad (3.63)$$

и

$$\Gamma(x) = e^{-x} x^{x-\frac{1}{2}} \sqrt{2\pi} \left\{ 1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} - \frac{571}{2488320x^4} \dots \right\}. \quad (3.64)$$

Полином Бернулли $B_j^n(x)$ порядка n степени j по x определяется равенством

$$e^{xt} t^n / (e^t - 1)^n = \sum_{j=0}^{\infty} t^j B_j^{(n)}(x) / j!. \quad (3.65)$$

Поправки Шеппарда к семиинвариантам

3.26 Как в 3.18, имеем, обозначая it буквой θ ,

$$\begin{aligned} \sum \left\{ e^{\theta x_j} \int_{-1/2h}^{1/2h} f(x_j + \xi) d\xi \right\} &= \frac{1}{h} \int_{-\infty}^{\infty} e^{\theta x} dx \int_{-1/2h}^{1/2h} f(x + \xi) d\xi = \\ &= \frac{1}{h} \int_{-1/2h}^{1/2h} e^{-\theta \xi} d\xi \int_{-\infty}^{\infty} e^{\theta x} f(x) dx = \frac{\operatorname{sh} \frac{1}{2} \theta h}{\frac{1}{2} \theta h} \int_{-\infty}^{\infty} e^{\theta x} f(x) dx. \end{aligned} \quad (3.66)$$

Левая часть представляет собой характеристическую функцию группированного распределения, а интеграл справа — истинную х. ф. Беря логарифмы от обеих частей, используя (3.61) и приравнявая коэффициенты при $\theta^r/r!$, получаем замечательно простое выражение для поправок Шеппарда к семиинвариантам:

$$\kappa_r = \bar{\kappa}_r - B_r h^r / h, \quad r > 1. \quad (3.67)$$

Так как числа Бернулли нечетного порядка равны нулю, за исключением B_1 , и первый семиинвариант совпадает со средним значением, то поправки Шеппарда не меняют семиинвариантов нечетного порядка. Для первых трех семиинвариантов четного порядка имеем

$$\left. \begin{aligned} \kappa_2 &= \bar{\kappa}_2 - \frac{h^2}{12}, \\ \kappa_4 &= \bar{\kappa}_4 + \frac{h^4}{120}, \\ \kappa_6 &= \bar{\kappa}_6 - \frac{h^6}{252}. \end{aligned} \right\} \quad (3.68)$$

Многомерные моменты и семиинварианты

3.27 Предыдущие результаты настоящей главы легко могут быть обобщены на многомерный случай. Чтобы не усложнять алгебраических выражений, мы ограничимся рассмотрением двух вариантов x_1 и x_2 . Читатель не встретит трудностей при обобщении на случай большего числа вариантов.

Двумерный момент μ'_{rs} относительно a_1 по x_1 и относительно a_2 по x_2 определяется как

$$\mu'_{rs} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - a_1)^r (x_2 - a_2)^s dF. \quad (3.69)$$

Если r или s равно нулю, то этот момент сводится к обычному одномерному моменту соответствующего маргинального распределения. Если ни r , ни s не равны нулю, то мы получаем новый тип момента — смешанный момент. Первый смешанный момент μ'_{11} играет важную роль в теории корреляции. Первый смешанный момент μ_{11} относительно средних значений x_1 и x_2 называется ковариацией.

Аналогично одномерным моментам, двумерные моменты относительно одних точек могут быть выражены через двумерные моменты относительно других точек. Если начало отсчета x_1 перенести из a_1 в b_1 , а начало отсчета x_2 — из a_2 в b_2 , то, полагая $c_1 = b_1 - a_1$, $c_2 = b_2 - a_2$, получим

$$\mu'_{rs}(a_1, a_2) = (\mu' + c_1)^r (\mu' + c_2)^s, \quad (3.70)$$

где в правой части под $(\mu')^j (\mu')^k$ понимается $\mu'_{jk}(b_1, b_2)$. Полученное равенство соответствует символическому соотношению

$$\mu'_r(a) = [\mu'(b) + c]^r$$

для одномерного случая.

Методы вычисления смешанных моментов численно заданных распределений будут рассмотрены во втором томе при изложении теории корреляции. Нахождение смешанных моментов для двумерных распределений, заданных математически, сводится к вычислению двойных сумм или двойных интегралов и не требует никаких новых статистических идей.

Пример 3.17

Рассмотрим двумерное распределение

$$dF = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right\} dx_1 dx_2, \\ -\infty \leq x_1, x_2 \leq \infty.$$

Вычислим интеграл

$$M(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{x_1 t_1 + x_2 t_2} dF.$$

Производя подстановку

$$\xi = x_1 - \sigma_1^2 t_1 - \rho\sigma_1\sigma_2 t_2,$$

$$\eta = x_2 - \rho\sigma_1\sigma_2 t_1 - \sigma_2^2 t_2,$$

находим

$$M(t_1, t_2) = \exp\left\{\frac{1}{2}(t_1^2\sigma_1^2 + 2t_1t_2\sigma_1\sigma_2\rho + t_2^2\sigma_2^2)\right\} \times \\ \times \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{\xi^2}{\sigma_1^2} - \frac{2\rho\xi\eta}{\sigma_1\sigma_2} + \frac{\eta^2}{\sigma_2^2}\right)\right\} d\xi d\eta = \\ = \exp\left\{\frac{1}{2}(t_1^2\sigma_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + t_2^2\sigma_2^2)\right\}.$$

Это соотношение позволяет вычислить все моменты, так как момент μ_{rs} равен коэффициенту при $\frac{t_1^r}{r!} \frac{t_2^s}{s!}$ в разложении $M(t_1, t_2)$. В частности,

$$\mu_{20} = \sigma_1^2, \mu_{11} = \rho\sigma_1\sigma_2, \mu_{02} = \sigma_2^2;$$

$$\mu_{30} = \mu_{21} = \mu_{12} = \mu_{03} = 0;$$

$$\mu_{40} = 3\sigma_1^4, \mu_{31} = 3\rho\sigma_1^3\sigma_2, \mu_{22} = (1 - 2\rho^2)\sigma_1^2\sigma_2^2,$$

$$\mu_{13} = 3\rho\sigma_1\sigma_2^3, \mu_{04} = 3\sigma_2^4.$$

3.28 Двумерный аналог соотношения (3.30) имеет вид

$$\begin{aligned} \exp \left\{ \frac{\kappa_{10}}{1!0!} t_1 + \frac{\kappa_{01}}{0!1!} t_2 + \dots + \frac{\kappa_{rs}}{r!s!} t_1^r t_2^s + \dots \right\} = \\ = 1 + \frac{\mu'_{10}}{1!0!} t_1 + \frac{\mu'_{01}}{0!1!} t_2 + \dots + \frac{\mu'_{rs}}{r!s!} t_1^r t_2^s + \dots \end{aligned}$$

или, символически,

$$\exp \left\{ \sum \frac{1}{p!} \kappa(t_1 + t_2)^p \right\} = \sum \frac{1}{p!} \mu'(t_1 + t_2)^p, \quad (3.71)$$

где

$$\kappa(t_1 + t_2)^p = p! \left\{ \frac{\kappa_{p0}}{p!0!} t_1^p + \frac{\kappa_{p-1,1}}{(p-1)!1!} t_1^{p-1} t_2 + \dots \right\}. \quad (3.72)$$

Пусть $\varphi(t_1, t_2)$ — характеристическая функция:

$$\varphi(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{it_1 x_1 + it_2 x_2} dF. \quad (3.73)$$

Как и в одномерном случае, если выполнены соответствующие условия существования, то

$$\varphi(t_1, t_2) = \sum_{r,s=0}^{\infty} \mu'_{rs} \frac{(it_1)^r}{r!} \frac{(it_2)^s}{s!} = \exp \left\{ \sum_{r,s=0}^{\infty} \kappa_{rs} \frac{(it_1)^r}{r!} \frac{(it_2)^s}{s!} \right\}, \quad (3.74)$$

где κ_{00} считается равным нулю.

Из этих соотношений можно вывести выражения для смешанных моментов через смешанные семинварианты и обратно. Соотношения между смешанными моментами и семинвариантами можно также получить из соответствующих одномерных соотношений посредством следующего простого формального процесса.

Рассмотрим равенство (одно из (3.37))

$$\mu'_4 = \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4. \quad (3.75)$$

Перепишем его формально в виде

$$\mu'(r^4) = \kappa(r^4) + 4\kappa(r^3)\kappa(r) + 3\{\kappa(r^2)\}^2 + 6\kappa(r^2)\{\kappa(r)\}^2 + \{\kappa(r)\}^4.$$

Считая r переменным и применяя к обеим частям оператор $s \frac{\partial}{\partial r}$, имеем

$$\begin{aligned} 4\mu'(r^3s) = 4\kappa(r^3s) + 12\kappa(r^2s)\kappa(r) + 4\kappa(r^2)\kappa(s) + \\ + 12\kappa(rs)\kappa(r^2) + 12\kappa(rs)\{\kappa(r)\}^2 + 12\kappa(r^2)\kappa(r)\kappa(s) + 4\{\kappa(r)\}^3\kappa(s). \end{aligned}$$

Деля на 4 и заменяя r и s соответствующими индексами, относящимися к первой и второй вариантам, получим выражение

μ'_{31} через семиинварианты

$$\mu'_{31} = \kappa_{31} + 3\kappa_{21}\kappa_{10} + \kappa_{30}\kappa_{01} + 3\kappa_{11}\kappa_{20} + 3\kappa_{11}\kappa_{10}^2 + 3\kappa_{20}\kappa_{10}\kappa_{01} + \kappa_{10}^3\kappa_{01}. \quad (3.76)$$

Этот процесс имеет общий характер и может быть обоснован с помощью соотношения (3.71). Он применим также для получения выражений семиинвариантов через моменты и к выводу формул, содержащих моменты относительно среднего. Например, из равенства

$$\mu_4 = \kappa_4 + 3\kappa_2^2 \quad (3.77)$$

находим, производя подобные операции, что

$$\mu_{31} = \kappa_{31} + 3\kappa_{11}\kappa_{20}. \quad (3.78)$$

Аналогично можно получить *)

$$\mu_{22} = \kappa_{22} + \kappa_{20}\kappa_{02} + 2\kappa_{11}^2. \quad (3.79)$$

3.29 Иногда бывают полезны следующие формулы:

$$\left. \begin{aligned} \mu_{11} &= \kappa_{11}; \quad \mu_{21} = \kappa_{21}; \quad \mu_{31} = \kappa_{31} + 3\kappa_{20}\kappa_{11}; \\ \mu_{22} &= \kappa_{22} + \kappa_{20}\kappa_{02} + 2\kappa_{11}^2; \quad \mu_{41} = \kappa_{41} + 4\kappa_{30}\kappa_{11} + 6\kappa_{21}\kappa_{20}; \\ \mu_{32} &= \kappa_{32} + \kappa_{30}\kappa_{02} + 6\kappa_{21}\kappa_{11} + 3\kappa_{20}\kappa_{12}; \\ \mu_{51} &= \kappa_{51} + 5\kappa_{40}\kappa_{11} + 10\kappa_{31}\kappa_{20} + 10\kappa_{30}\kappa_{21} + 15\kappa_{20}^2\kappa_{11}; \\ \mu_{42} &= \kappa_{42} + \kappa_{40}\kappa_{02} + 8\kappa_{31}\kappa_{11} + 4\kappa_{30}\kappa_{12} + 6\kappa_{22}\kappa_{20} + \\ &\quad + 6\kappa_{21}^2 + 3\kappa_{20}^2\kappa_{02} + 12\kappa_{20}\kappa_{11}^2; \\ \mu_{33} &= \kappa_{33} + 3\kappa_{31}\kappa_{02} + \kappa_{30}\kappa_{03} + 9\kappa_{22}\kappa_{11} + 9\kappa_{21}\kappa_{12} + \\ &\quad + 3\kappa_{20}\kappa_{13} + 9\kappa_{20}\kappa_{11}\kappa_{02} + 6\kappa_{11}^3. \end{aligned} \right\} \quad (3.80)$$

$$\left. \begin{aligned} \kappa_{31} &= \mu_{31} - 3\mu_{20}\mu_{11}; \quad \kappa_{22} = \mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2; \\ \kappa_{41} &= \mu_{41} - 4\mu_{30}\mu_{11} - 6\mu_{21}\mu_{20}; \\ \kappa_{32} &= \mu_{32} - \mu_{30}\mu_{02} - 6\mu_{21}\mu_{11} - 3\mu_{20}\mu_{12}; \\ \kappa_{51} &= \mu_{51} - 5\mu_{40}\mu_{11} - 10\mu_{31}\mu_{20} - 10\mu_{30}\mu_{21} + 30\mu_{20}\mu_{11}; \\ \kappa_{42} &= \mu_{42} - \mu_{40}\mu_{02} - 8\mu_{31}\mu_{11} - 4\mu_{30}\mu_{12} - 6\mu_{22}\mu_{20} - 6\mu_{21}^2 + \\ &\quad + 6\mu_{20}^2\mu_{02} + 24\mu_{20}\mu_{11}^2; \\ \kappa_{33} &= \mu_{33} - 3\mu_{31}\mu_{02} - \mu_{30}\mu_{03} - 9\mu_{22}\mu_{11} - 9\mu_{21}\mu_{12} - 3\mu_{20}\mu_{13} + \\ &\quad + 18\mu_{20}\mu_{11}\mu_{02} + 12\mu_{11}^3. \end{aligned} \right\} \quad (3.81)$$

*) Кук (1951) дала формулы для выражения μ'_{rs} через κ_{rs} и обратно при всех $r+s \leq 6$.

3.30 Волд (1934b) дал следующие выражения для поправок Шеппарда к двумерным моментам и семинвариантам (длины интервалов группировки взяты равными h_1 и h_2). Для моментов

$$\mu'_{rs} = \sum_{j=0}^r \sum_{k=0}^s h_1^j h_2^k \binom{r}{j} \binom{s}{k} (2^{1-j} - 1)(2^{1-k} - 1) B_j B_k \bar{\mu}'_{r-j, s-k}. \quad (3.82)$$

В частности,

$$\left. \begin{aligned} \mu'_{20} &= \bar{\mu}'_{20} - \frac{1}{12} h_1^2, & \mu'_{11} &= \bar{\mu}'_{11}, & \mu'_{02} &= \bar{\mu}'_{02} - \frac{1}{12} h_2^2; \\ \mu'_{30} &= \bar{\mu}'_{30} - \bar{\mu}'_{10} h_1^2/4, & \mu'_{21} &= \bar{\mu}'_{21} - \bar{\mu}'_{01} h_1^2/12; \\ \mu'_{40} &= \bar{\mu}'_{40} - \bar{\mu}'_{20} h_1^2/2 + 7h_1^4/240, & \mu'_{31} &= \bar{\mu}'_{31} - \bar{\mu}'_{11} h_1^2/4; \\ \mu'_{22} &= \bar{\mu}'_{22} - \bar{\mu}'_{20} h_2^2/12 - \bar{\mu}'_{02} h_1^2/12 + h_1^2 h_2^2/144. \end{aligned} \right\} \quad (3.83)$$

Для семинвариантов

$$\left. \begin{aligned} \kappa_{rs} &= \bar{\kappa}_{rs}, & r, s &> 0, \\ \kappa_{r0} &= \bar{\kappa}_{r0} - B_r h_1^r/r, & r &\geq 2, \\ \kappa_{0s} &= \bar{\kappa}_{0s} - B_s h_2^s/s, & s &\geq 2. \end{aligned} \right\} \quad (3.84)$$

Меры асимметрии

3.31 В главе 2 мы рассмотрели меры расположения и рассеяния. С помощью моментов мы можем теперь приступить к рассмотрению мер других качеств генеральной совокупности, в частности асимметричности.

Для симметричных распределений среднее значение, медиана и мода совпадают. Поэтому естественно в качестве меры асимметрии распределения взять отклонение среднего от моды или среднего от медианы. К. Пирсоном была предложена следующая мера:

$$Sk = \frac{\text{среднее} - \text{мода}}{\sigma}.$$

Недостатком ее является необходимость находить моду. Однако для широкого класса распределений — так называемых пирсоновских распределений (см. главу 6) — эта мера может быть выражена через первые четыре момента. Положим,

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad (3.85)$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}. \quad (3.86)$$

Можно показать, что для распределений Пирсона

$$Sk = \frac{\sqrt{\beta_1} (\beta_2 + 3)}{2(5\beta_2 - 6\beta_1 - 9)}. \quad (3.87)$$

Это выражение может быть взято в качестве определения меры асимметрии, применимого ко всякому распределению, у которого существуют все моменты вплоть до четвертого включительно.

Само отношение β_1 также является мерой асимметрии. Ясно, что если распределение симметрично, то β_1 обращается в нуль, поскольку тогда μ_3 равно нулю; величина отношения μ_3 к $\mu_2^{3/2}$ (равного *) $\sqrt{\beta_1}$ показывает степень отклонения от симметрии.

Вообще можно определить

$$\left. \begin{aligned} \beta_{2n+1} &= \frac{\mu_3 \mu_{2n+3}}{\mu_2^{n+3}}, \\ \beta_{2n} &= \frac{\mu_{2n+2}}{\mu_2^{n+1}}. \end{aligned} \right\} \quad (3.88)$$

Нельзя сказать, чтобы β_{2n+1} и β_{2n} широко использовались, но время от времени они встречаются в статистической литературе.

В некоторых отношениях более удобными, чем величины β_1 и β_2 , являются

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}}, \quad (3.89)$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\kappa_4}{\kappa_2^2}. \quad (3.90)$$

Если распределение нормировано, то γ_1 и γ_2 равны соответственно третьему и четвертому семиинвариантам.

Экссесс

3.32 Для так называемого «нормального» распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} x^2/\sigma^2} dx, \quad -\infty \leq x \leq \infty,$$

β_2 равно 3 и γ_2 — нулю. О кривых плотности, для которых $\gamma_2 = 0$, говорят, что они имеют нормальный эксцесс. Если $\gamma_2 > 0$, то говорят, что эксцесс кривой положителен, а при $\gamma_2 < 0$, что он отрицателен.

*) На самом деле здесь и далее под $\sqrt{\beta_1}$ автор понимает не арифметическое значение квадратного корня, а $\sqrt{\beta_1} \operatorname{sign} \mu_3$. (Прим. ред.)

Раньше существовало мнение, что кривые с положительным эксцессом имеют более высокую и острую вершину сравнительно с нормальной кривой, а кривые с отрицательным эксцессом — более низкую и плоскую. Это, однако, не обязательно так. Тем не менее введенные понятия полезны, но их лучше рассматривать как описывающие знак γ_2 , а не характер кривой. См. упражнения 3.20 и 3.21.

Пример 3.18

Для распределения браков в Австралии, рассмотренного в примере 3.1, групповые моменты относительно среднего при группировке в интервалы длиной три года равны

$$\bar{\mu}_2 = 7,056\ 977, \quad \bar{\mu}_3 = 36,151\ 595, \quad \bar{\mu}_4 = 408,738\ 210.$$

После применения поправок Шеппарда они переходят в

$$\mu_2 = 6,973\ 644, \quad \mu_3 = 36,151\ 595, \quad \mu_4 = 405,238\ 888.$$

По этим моментам находим

$$\begin{aligned} \beta_1 &= 3,854, & \gamma_1 &= 1,963, \\ \beta_2 &= 8,333, & \gamma_2 &= 5,333, \end{aligned}$$

что указывает на значительные асимметрию и положительный эксцесс.

Пример 3.19

С помощью формул для моментов биномиального распределения, полученных в примере 3.2, находим

$$\gamma_1 = \frac{q-p}{\sqrt{npq}}, \quad \gamma_2 = \frac{1-6pq}{npq}.$$

Отсюда следует, что γ_1 и γ_2 стремятся к нулю при $n \rightarrow \infty$. Это находится в соответствии с тем фактом (мы докажем его позднее), что нормированное биномиальное распределение сходится к нормальному при стремлении n к бесконечности.

Моменты как характеристики распределения

3.33 В последующих главах встретится множество примеров, иллюстрирующих пользу моментов и семиинвариантов при выяснении природы частотного распределения, однако уже сейчас можно сделать об этом ряд общих замечаний.

Уже отмечалось, что характеристическая функция определяет все моменты, если они существуют; в главе 4 будет доказано, что характеристическая функция определяет также и функцию распределения. Тем не менее моменты, даже если все они существуют, не определяют полностью распределения. Они определяют его однозначно лишь при некоторых дополнительных условиях. К счастью для статистиков, этим условиям удовлетворяют все встречающиеся в статистической практике распределения*). Практически, следовательно, знание моментов, когда все они существуют, эквивалентно знанию функции распределения: эквивалентно в том смысле, что *теоретически* было бы возможно все свойства распределения выразить в терминах моментов.

3.34 Можно ожидать, в частности, что если два распределения имеют некоторое число одинаковых моментов, то они в какой-то степени схожи. Если, скажем, моменты у двух распределений совпадают вплоть до порядка n и в пределе, при n , стремящемся к бесконечности, эти распределения становятся одинаковыми, то можно ожидать, что при совпадении конечного числа первых моментов распределения приблизительно одинаковы. Некоторое математическое обоснование этого так называемого принципа моментов может быть получено следующим образом.

Известно, что функция, непрерывная на конечном отрезке $[a, b]$, может быть представлена на этом отрезке как сумма равно-

номерно сходящегося ряда полиномов от x , скажем $\sum_{n=0}^{\infty} P_n(x)$,

где $P_n(x)$ — полином степени n . Предположим, что мы хотим аппроксимировать такую функцию *многочленом* вида $\sum_{n=0}^s a_n x^n$.

Коэффициенты a_n могут быть определены по принципу наименьших квадратов, т. е. из условия, чтобы интеграл

$$\int_a^b \left(f - \sum a_n x^n \right)^2 dx \quad (3.91)$$

был минимальным. Дифференцируя по a_j , имеем

$$2 \int_a^b \left(f - \sum a_n x^n \right) x^j dx = 0,$$

*) К сожалению, это не так. Например, такое важное и часто встречающееся в приложениях распределение, как логарифмически нормальное (см. 6.29), не определяется своими моментами однозначно. В утверждении авторов «все» следовало бы заменить на «почти все» (*Прим. ред.*)

или

$$\int_a^b f x^j dx = \mu'_j = \int_a^b \sum a_n x^{n+j} dx. \quad (3.92)$$

Пусть теперь два распределения имеют одинаковые моменты вплоть до порядка s . Тогда они имеют и одинаковые аппроксимации в смысле принципа наименьших квадратов, поскольку в силу (3.92) коэффициенты a_n определяются этими моментами. Более того, если f_1 отличается от $\sum a_n x^n$ не больше чем на ε_1 , а f_2 — не больше чем на ε_2 , то f_1 отличается от f_2 не больше чем на $\varepsilon_1 + \varepsilon_2$.

Подобный подход может быть использован и тогда, когда распределения не сосредоточены на конечном отрезке. В этом случае при довольно общих условиях функцию плотности можно представить в виде ряда с членами вида *) $e^{-x^2} P_n(x)$. Выводы получаются такие же.

Таким образом, распределения, которые имеют конечное число соответственно одинаковых низших моментов, в некотором смысле аппроксимируют друг друга. В дальнейшем встретится много случаев, когда мы не сможем точно определить функцию распределения, но сможем найти моменты, по крайней мере до некоторого порядка. В таком случае мы сможем аппроксимировать исследуемое распределение, найдя другое распределение известного вида с теми же первыми моментами. Практически аппроксимация такого рода часто оказывается очень хорошей, даже когда совпадают только первые три или четыре момента.

Неравенства чебышевского типа

3.35 Если известны некоторые моменты распределения, то относительно функции плотности или функции распределения можно сделать точные утверждения в терминах неравенств. Пусть F — функция распределения и пусть существует второй момент. Для любого действительного t справедливо равенство

$$\int_{-\infty}^{\infty} \left(1 - \frac{x^2}{t^2}\right) dF = \left\{ \int_{-\infty}^{-t} + \int_t^{\infty} \right\} \left(1 - \frac{x^2}{t^2}\right) dF + \int_{-t}^t \left(1 - \frac{x^2}{t^2}\right) dF.$$

*) Теорема Веры Миллер-Лебедевой (Vera Myller-Lebedeff, (1907), Math. Апп. 64, 388) утверждает, что функция может быть разложена в ряд по производным $\exp(-x^2)$. Это разложение не следует смешивать с рядом Грама — Шарлье (см. главу 6), где функция разлагается в ряд по производным $\exp\left(-\frac{1}{2}x^2\right)$.

В первых двух интегралах справа $1 - x^2/t^2 \leq 0$, следовательно,

$$1 - \frac{\mu'_2}{t^2} = \int_{-\infty}^{\infty} \left(1 - \frac{x^2}{t^2}\right) dF \leq \int_{-t}^t \left(1 - \frac{x^2}{t^2}\right) dF \leq \int_{-t}^t dF.$$

Таким образом, частота значений между $-t$ и t не меньше, чем $1 - \mu'_2/t^2$. В частности, если начало совместить со средним значением распределения, то

$$F(t) - F(-t) \geq 1 - \mu_2/t^2. \quad (3.93)$$

На языке теории вероятностей это неравенство обычно записывают в следующем виде, положив $\mu_2 = \sigma^2$, $t = \lambda\sigma$:

$$P\{|x - M(x)| \leq \lambda\sigma\} \geq 1 - 1/\lambda^2 \quad (3.94)$$

или

$$P\{|x - M(x)| \geq \lambda\sigma\} \leq 1/\lambda^2. \quad (3.95)$$

Это неравенство принадлежит Бьенэме, который опубликовал его в 1853 году, и Чебышеву, независимо открывшему его позднее. Обычно оно называется неравенством Бьенэме — Чебышева. Оно дает нижнюю границу для частоты, отвечающей интервалу длины $2\lambda\sigma$, центр которого совпадает со средним значением распределения. Будучи верным с очень большой степенью общности, это неравенство не играет особой роли при оценке частот в частных случаях, но оно неоценимо при установлении предельных закономерностей и довольно полезно, когда частоты оцениваются лишь грубо.

3.36 В настоящее время существует большое число неравенств чебышевского типа. В дальнейшем некоторые из них нам встретятся. Несколько наиболее употребительных неравенств мы приведем здесь для ссылок.

Непосредственным обобщением примененного выше метода легко показать, что если ν_r — абсолютный момент порядка r , то

$$P\{|x - M(x)|/\nu_r^{1/r} \leq \lambda\} \geq 1 - 1/\lambda^r. \quad (3.96)$$

Это неравенство, принадлежащее К. Пирсону (1919), включает неравенство Бьенэме — Чебышева как частный случай (при $r=2$). Если варианта не принимает отрицательных значений, то можно также показать (упражнение 2.15), что при всех $t > 0$

$$F(t) > 1 - \mu'_1/t. \quad (3.97)$$

Иногда (3.97) называют неравенством Маркова. Если за варианту взять $\{x - M(x)\}^2$ и положить $t = \lambda^2 \sigma^2$, то (3.97) перейдет в (3.94).

Для распределений общего вида эти неравенства не могут быть улучшены. Весьма замечательно, однако, что, накладывая лишь очень слабые ограничения на форму функции плотности, неравенства можно значительно усилить. Например, если непрерывная функция плотности имеет только одну моду m_0 и начало отсчета помещено в моду, то

$$P\{|x - m_0| \leq \lambda\tau\} \geq 1 - 4/(9\lambda^2), \quad (3.98)$$

где τ — квадратный корень из второго момента относительно m_0 . Неравенство (3.98) (в иной, правда, форме) восходит к Гауссу.

3.37 С кругом вопросов, касающихся этих неравенств, более подробно можно познакомиться по обширному обзору Годвина (1955). Предполагая известными два средних значения (например, два момента), можно получить более тонкие неравенства. Мэллоуз (1956), накладывая дальнейшие ограничения на функцию плотности, такие как существование только одной точки перегиба по каждую сторону от моды, получил очень хорошие аппроксимации самой функции распределения. Вопросы, связанные с подобными неравенствами, привлекательны с математической точки зрения и сами по себе могли бы составить содержание отдельной монографии.

УПРАЖНЕНИЯ

3.1 Показать, что первые моменты дискретного (пуассоновского) распределения, частоты которого в точках $0, 1, \dots, j, \dots$ суть

$$e^{-\lambda} \left(1, \frac{\lambda}{1!}, \dots, \frac{\lambda^j}{j!}, \dots \right),$$

выражаются формулами

$$\mu_2 = \lambda, \quad \mu_3 = \lambda, \quad \mu_4 = \lambda(1 + 3\lambda), \quad \mu_5 = \lambda(1 + 10\lambda), \quad \mu_6 = \lambda(1 + 25\lambda + 15\lambda^2)$$

3.2 Показать, что для распределения

$$dF = kx^{-p}e^{-\gamma/x} dx, \quad 0 \leq x \leq \infty, \quad \gamma > 0,$$

моменты относительно нуля при $r < p-1$ выражаются формулами

$$\mu'_r = \frac{\gamma^r \Gamma(p-r-1)}{\Gamma(p-1)}$$

и не существуют в противном случае.

3.3 Показать, что для распределения

$$dF = k \left(1 + \frac{x^2}{a^2} \right)^{-m} \exp \{ -\nu \operatorname{arctg}(x/a) \} dx, \quad -\infty \leq x \leq \infty,$$

моменты относительно нуля равны

$$\mu'_r = ka^{r+1} \int_{-\pi/2}^{\pi/2} \cos^{2m-r-2} \theta \sin^r \theta e^{-\nu \theta} d\theta$$

и, следовательно,

$$\mu'_r = \frac{a}{2m-r-1} \{ (r-1) a \mu'_{r-2} - \nu \mu'_{r-1} \}.$$

3.4 Пусть $f(x)$ — непрерывная симметричная относительно $x=0$, имеющая единственный максимум при $x=0$ функция плотности, отвечающая варианту, принимающей значения из отрезка $[-a, a]$. Показать, что

$$\mu_{2r} < \frac{a^{2r}}{2r+1}, \quad r \geq 1.$$

3.5 Показать, что для всякого симметричного распределения семиинварианты нечетного порядка (исключая μ_1) равны нулю (если они существуют).

3.6 Показать, что семиинварианты распределения

$$dF = e^{-x/\sigma} dx/\sigma, \quad 0 \leq x < \infty, \quad \sigma > 0,$$

равны

$$\mu_r = \sigma^r (r-1)!$$

3.7 Показать, что функцию e^{itx} можно разложить в бесконечный ряд

$$1 + (e^{it} - 1)x^{[1]} + (e^{it} - 1)^2 \frac{x^{[2]}}{2!} + \dots + (e^{it} - 1)^r \frac{x^{[r]}}{r!} + \dots,$$

сходящийся при $-\infty \leq x \leq \infty$ (здесь $x^{[r]} = x(x-1) \dots (x-r+1)$), и что, следовательно,

$$\mu'_{[r]} = [D^r \varphi(t)]_{t=0},$$

где $D = \frac{d}{d(e^{it})}$.

Вывести отсюда, что для биномиального распределения $[(q+p)^n]$ факториальные моменты относительно нуля равны

$$\mu'_{[r]} = n^{[r]} p^r.$$

3.8 Показать, что для распределения, частоты которого в точках $\pm 2r$ (r — целое) суть

$$e^{-2a} \left\{ \frac{a^{2r}}{0!(2r)!} + \frac{a^{2r+2}}{1!(2r+1)!} + \frac{a^{2r+4}}{2!(2r+2)!} + \dots \right\},$$

а в точках $\pm (2r+1)$ —

$$e^{-2a} \left\{ \frac{a^{2r+1}}{0!(2r+1)!} + \frac{a^{2r+3}}{1!(2r+2)!} + \frac{a^{2r+5}}{2!(2r+3)!} + \dots \right\},$$

семиинварианты нечетного порядка равны нулю, а четного порядка равны $2a$.

3.9 Показать, что

$$\mu'_r = \sum_{j=1}^r \binom{r-1}{j-1} \mu'_{r-j} \mu_j,$$

и, следовательно,

$$\mu_r = (-1)^{r-1} \begin{vmatrix} \mu'_1 & 1 & 0 & 0 \dots & 0 \\ \mu'_2 & \binom{1}{0} \mu'_1 & 1 & 0 \dots & 0 \\ \mu'_3 & \binom{2}{0} \mu'_2 & \binom{2}{1} \mu'_1 & 1 \dots & 0 \\ \dots & \dots & \dots & \dots & 1 \\ \mu'_r & \binom{r-1}{0} \mu'_{r-1} & \binom{r-1}{1} \mu'_{r-2} & \dots & \binom{r-1}{r-2} \mu'_1 \end{vmatrix}.$$

3.10 Показать, что для распределения, получающегося из распределения

$$dF = dx, \quad 0 \leq x \leq 1,$$

группированием в целом числе интервалов одинаковой длины h , поправки ко второму и четвертому моментам относительно среднего суть

$$\mu_2 = \bar{\mu}_2 + \frac{h^2}{12},$$

$$\mu_4 = \bar{\mu}_4 + \bar{\mu}_2 \frac{h^2}{2} + \frac{h^4}{80}$$

(ср. Элдертон, 1938b. Отметим, что первая поправка в точности, а вторая приблизительно равна поправке Шеппарда, взятой с *обратным знаком*).

3.11 Показать, что если ∂_p — оператор, для которого

$$\partial_p \mu'_r = \begin{cases} r! p^r \mu'_{r-p}, & r \geq p, \\ 0, & r < p, \end{cases}$$

и ∂_p дистрибутивен при действии на произведение, т. е.

$$\partial_p(AB) = B(\partial_p A) + A(\partial_p B),$$

то ∂_p аннулирует все семинварианты (рассматриваемые как функции от моментов), за исключением κ_p и

$$\partial_p \kappa_p = p!.$$

3.12 Показать, что если $f(x)$ — нечетная функция с периодом $\frac{1}{2}$, то

$$\int_0^\infty x^r x^{-\log x} f(\log x) dx = 0$$

для всех целых r . Вывести отсюда, что моменты распределений

$$dF = x^{-\log x} \{1 - \lambda \sin(4\pi \log x)\} dx, \quad 0 \leq x \leq \infty, \quad 0 \leq \lambda \leq 1,$$

одни и те же при всех значениях λ .

(Стильтьес, 1918.)

3.13 Показать, что если ненулевым частотам дискретного распределения отвечают точки, отстоящие друг от друга на расстоянии h/m , то при группировании в интервалы длиной h , по m точек в каждый, осредненные поправки к семинвариантам даются формулой

$$\kappa_r = \bar{\kappa}_r - \frac{B_r h^r}{r} \left(1 - \frac{1}{m^r}\right).$$

(Крэйг, 1936.)

3.14 Показать, что для двумерного распределения

$$dF = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2} \right\} \right] dx_1 dx_2,$$

$$-\infty \leq x_1, x_2 \leq \infty,$$

все семиинварианты κ_{rs} , $r, s > 2$ равны нулю. Показать далее, что если положить

$$\lambda_{rs} = \frac{\mu_{rs}}{\sigma_1^r \sigma_2^s},$$

то

$$\lambda_{rs} = (r+s-1) \rho \lambda_{r-1, s-1} + (r-1)(s-1)(1-\rho^2) \lambda_{r-2, s-2},$$

$$\lambda_{2r, 2s} = \frac{(2r)!(2s)!}{2^{r+s}} \sum_{j=0}^t \frac{(2\rho)^{2j}}{(r-j)!(s-j)!(2j)!},$$

$$\lambda_{2r+1, 2s+1} = \frac{(2r+1)!(2s+1)!}{2^{r+s}} \rho \sum_{j=0}^t \frac{(2\rho)^{2j}}{(r-j)!(s-j)!(2j+1)!},$$

$$\lambda_{2r, 2s+1} = \lambda_{2r+1, 2s} = 0,$$

где t — наименьшее из чисел r, s . В частности,

$$\lambda_{11} = \rho, \quad \lambda_{31} = 3\rho, \quad \lambda_{51} = 15\rho, \quad \lambda_{71} = 105\rho, \quad \lambda_{91} = 945\rho;$$

$$\lambda_{22} = (1+2\rho^2), \quad \lambda_{24} = 3(1+4\rho^2), \quad \lambda_{26} = 15(1+6\rho^2);$$

$$\lambda_{28} = 105(1+8\rho^2), \quad \lambda_{2,10} = 945(1+10\rho^2);$$

$$\lambda_{33} = 3\rho(3+2\rho^2), \quad \lambda_{35} = 15\rho(3+4\rho^2);$$

$$\lambda_{44} = 3(3+24\rho^2+8\rho^4).$$

3.15 (Неравенство Ляпунова для абсолютных моментов.) Исходя из неравенства Гельдера

$$\sum x^\alpha x^\beta \leq \sum x^\alpha \sum x^\beta, \quad \alpha + \beta = 1,$$

полагая в нем $x = u^c$, $y = u^a$, $\beta = 1 - \alpha$, $\alpha = (a-b)/(a-c)$, вывести неравенство

$$v_b^{a-c} \leq v_c^{a-b} v_a^{b-c}, \quad a \geq b \geq c \geq 0.$$

3.16 Проверить формулы для κ_{42} и κ_{33} в (3.81).

3.17 Распределение группировано в интервалы длиной h . Показать, что при движении решетки группирования вдоль оси изменения варианты групповые моменты меняются с периодом h . Далее, положив

$$\bar{\mu}'_r = \sum_{j=-\infty}^{\infty} \xi^r \int_{\xi-h/2}^{\xi+h/2} f(x) dx = A_0 + \sum_{j=1}^{\infty} A_j \sin j\theta + \sum_{j=1}^{\infty} B_j \cos j\theta,$$

где $\xi = (j + \theta/2\pi)h$, показать, что

$$A_0 = \frac{1}{h} \int_{-\infty}^{\infty} x^r \int_{-h/2}^{h/2} f(x+\xi) d\xi dx$$

и что, следовательно, это разложение ведет к поправкам Шеппарда.

Показать также, что

$$A_s = \frac{2}{h} \int_{-\infty}^{\infty} f(x) dx \int_{x-h/2}^{x+h/2} \zeta^r \sin(2\pi s \zeta/h) d\zeta$$

и аналогично для B_s .

Показать, наконец, что для распределения $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx, -\infty \leq x \leq \infty$ при $r=1$ все B_s обращаются в нуль, а $A_s = (-1)^{s+1} \frac{h}{\pi s} \exp\left(-\frac{2s^2\pi^2}{h^2}\right)$ и что, стало быть, даже для грубого группирования с $h=1$ поправка к среднему не превосходит $e^{-2\pi^2}/\pi$ (приблизительно). (Фишер, 1921a.)

3.18 (Неравенство Гаусса — Винклера.) Рассматривая $1 - [f(\bar{x}-x) + f(\bar{x}+x)]/2f(x)$, $x>0$, как функцию распределения и используя неравенство Ляпунова, показать, что абсолютные моменты относительно моды $\bar{x}$, отвечающие дифференцируемой функции плотности с единственной модой, удовлетворяют неравенству

$$\{(r+1)v'_r\}^{1/r} \leq \{(n+1)v'_n\}^{1/n}, \quad r < n.$$

Вывести отсюда, что если среднее совпадает с $\bar{x}$, то $\beta_2 \geq 1,8$.

3.19 Показать, что

$$\begin{vmatrix} \mu'_0 & \mu'_1 & \mu'_2 \\ \mu'_1 & \mu'_2 & \mu'_3 \\ \mu'_2 & \mu'_3 & \mu'_4 \end{vmatrix} > 0$$

и, следовательно, $\beta_2 > 1 + \beta_1$.

3.20 Рассмотреть два распределения с ф. п.

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

$$f_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \{1 + A(x^8 - 22x^6 + 116x^4 - 126x^2 + 3)\},$$

где A — малое положительное число. Показать, что при подходящем выборе A оба распределения имеют нулевое среднее, единичную дисперсию и только одну точку перегиба (у функций плотности) между 0 и ∞ . Показать также, что при x , близких к нулю, так же как и при больших x , $f_2 > f_1$, хотя эксцесс у f_2 отрицательный.

(Дайсон, 1943.)

3.21 Для следующих четырех унимодальных симметричных функций плотности среднее равно нулю и дисперсия — единице. Проверить, что значения μ_4 и максимумы этих плотностей соответствуют выписанным в таблице:

	Значение μ_4	Максимум плотности
(1) $\frac{1}{3\sqrt{\pi}} \left(\frac{9}{4} + x^4 \right) e^{-x^2}$	2,75	0,423
(2) $\frac{3}{2\sqrt{2\pi}} e^{-\frac{1}{2}x^2} - \frac{1}{6\sqrt{\pi}} \left(\frac{9}{4} + x^4 \right) e^{-x^2}$	3,125	0,387
(3) $\frac{1}{6\sqrt{\pi}} \left(e^{-\frac{1}{4}x^2} + 4e^{-x^2} \right)$	4,5	0,470
(4) $\frac{3\sqrt{3}}{16\sqrt{\pi}} (2 + x^2) e^{-\frac{3}{4}x^2}$	2,667	0,366

Сравнивая с соответствующими значениями для нормального распределения: $\mu_4 = 3$, максимум плотности = 0,399, показать, что как в случае, когда значение плотности в нуле больше значения в нуле нормальной плотности, так и в случае, когда она меньше, возможны и положительный и отрицательный эксцессы.

(Канлански, 1945.)

3.22 Определив по аналогии с 3.31.

$s = (\text{среднее} - \text{медиана}) / (\text{стандартное отклонение})$,

показать, что для всех распределений $-1 \leq s \leq 1$.

(Хотеллинг и Соломонс, 1932.)

3.23 Используя соотношение (3.61), показать, что прямоугольное распределение $dF = x dx/h$, $-\frac{1}{2}h \leq x \leq \frac{1}{2}h$, имеет х. ф.

$$\varphi(t) = \text{sh} \left(\frac{1}{2} it h \right) / (it h)$$

и семинварианты $\kappa_{2r-1} = 0$, $\kappa_{2r} = B_{2r} h^{2r} / (2r)$, $r = 1, 2, \dots$. Аналогично, дискретное прямоугольное распределение

$$P\{x=r\} = \frac{1}{n}, \quad r = 1, 2, \dots, n,$$

имеет х. ф.

$$\varphi(t) = \text{sh} \left(\frac{1}{2} it n \right) / \left\{ n \text{sh} \left(\frac{1}{2} it \right) \right\}$$

и семинварианты

$$\kappa_1 = \frac{1}{2} (n+1),$$

$$\kappa_{2r+1} = 0, \quad \kappa_{2r} = B_{2r} (n^{2r} - 1) / (2r), \quad r = 1, 2, \dots$$

3.24 Показать, что если варианта принимает только неотрицательные целые значения, то формула (3.29) допускает обращение:

$$f_j = \frac{1}{j!} \sum_{r=0}^{n-j} (-1)^r \mu'_{j+r} / r!.$$

Подставить в эту формулу факториальные моменты, найденные в примере 3.8, и убедиться, что получаются частоты биномиального распределения.

ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ

4.1 В предыдущей главе мы ввели характеристическую функцию

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF \quad (4.1)$$

как производящую функцию моментов. Характеристические функции (х. ф.) обладают еще целым рядом полезных и важных свойств, что и объясняет ту центральную роль, которую х. ф. играют в статистической теории. В настоящей главе будет дана сводка различных свойств х. ф. При этом для установления некоторых результатов мы будем вынуждены прибегнуть к ряду сведений из высшей математики. Читатель, интересующийся лишь статистическими приложениями, может при чтении этой главы опустить доказательства и ограничиться усвоением только самих результатов. Вообще говоря, излагаемая в этой главе теория может рассматриваться как специальный случай общей теории преобразований Фурье в силу того, что рассматриваемые нами функции не произвольны, а являются или функциями плотности, или функциями распределения. Доказательства ряда фактов в этой главе заимствованы нами из книги: P. Lévy, «Calcul des Probabilités», Gauthier — Villars, Paris, 1925 *).

4.2 Заметим сначала, что $\varphi(t)$ всегда существует, поскольку

$$\left| \int_{-\infty}^{\infty} e^{itx} dF(x) \right| \leq \int_{-\infty}^{\infty} |e^{itx}| dF = \int_{-\infty}^{\infty} dF = 1. \quad (4.2)$$

Так что интеграл в (4.1) абсолютно сходится. Далее, функция $\varphi(t)$ равномерно непрерывна по t и j раз дифференцируема под

*) Недавно Лукач (E. Lukacs, Characteristic Functions, 1960, Griffin) опубликовал изящную монографию, посвященную характеристическим функциям. (Обширный материал о характеристических функциях содержится в книге Ю. В. Линника, «Разложения вероятностных законов», Изд. ЛГУ, 1960. (Прим. перев.)

знаком интеграла, если только результирующий интеграл существует и равномерно сходится. Для этого в свою очередь достаточно существования момента ν_j . В самом деле, тогда

$$|\varphi^{(j)}(t)| = \left| \int_{-\infty}^{\infty} x^j e^{itx} dF \right| \leq \int_{-\infty}^{\infty} |x^j| dF = \nu_j. \quad (4.3)$$

Теорема обращения

4.3 Мы сейчас докажем фундаментальную теорему теории х. ф. — так называемую теорему обращения. Согласно этой теореме х. ф. однозначно определяет функцию распределения. Более точно, если $\varphi(t)$ задается формулой (4.1), то

$$F(x) - F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) \frac{1 - e^{-ixt}}{it} dt, \quad (4.4)$$

где интеграл понимается в смысле главного значения, т. е. как

$$\lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \varphi(t) \frac{1 - e^{-ixt}}{it} dt.$$

Если к тому же $F(x)$ имеет плотность $f(x)$, т. е. $dF = f(x)dx$, то

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) e^{-ixt} dt, \quad (4.5)$$

где интеграл снова надо понимать в смысле главного значения, если только он не сходится в обычном смысле. Сравнение формулы (4.5) с формулой

$$\varphi(t) = \int_{-\infty}^{\infty} f(x) e^{itx} dx \quad (4.6)$$

показывает, что между $f(x)$ и $\varphi(t)$ существует соотношение двойственности.

Уравнение (4.4), содержащее $F(0)$, можно записать в несколько иной форме. А именно,

$$F(x) = \frac{1}{2} + \frac{1}{2\pi} \int_0^{\infty} \frac{e^{itx} \varphi(-t) - e^{-itx} \varphi(t)}{it} dt. \quad (4.7)$$

Прежде чем непосредственно переходить к доказательству теоремы обращения, мы установим, что для всякого действительного числа ξ

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin t\xi}{t} dt = \begin{cases} -1, & \xi < 0, \\ 0, & \xi = 0, \\ 1, & \xi > 0. \end{cases} \quad (4.8)$$

Величина, стоящая в правой части формулы (4.8), обычно обозначается $\operatorname{sgn} \xi$ (читается «сигнум ξ »). Проще всего формулу (4.8) установить, вычисляя комплексный интеграл

$$\int \frac{e^{iz\xi}}{iz} dz$$

вдоль контура, состоящего из большой полуокружности, действительной оси и малой полуокружности около начала координат. Ниже мы будем также использовать тот факт, что для переменной u и фиксированного x

$$\int_{-\infty}^{\infty} \operatorname{sgn}(u-x) dF(u) = - \int_{-\infty}^x dF(u) + \int_x^{\infty} dF(u) = 1 - 2F(x). \quad (4.9)$$

Для каждого положительного c интеграл

$$\begin{aligned} I_c &\equiv \int_{-c}^c \frac{1-e^{-ixt}}{it} \varphi(t) dt = \int_{-c}^c \frac{1-e^{-ixt}}{it} \int_{-\infty}^{\infty} e^{itu} dF(u) dt = \\ &= \int_{-c}^c \int_{-\infty}^{\infty} \frac{e^{itu} - e^{it(u-x)}}{it} dF(u) dt = \\ &= \int_{-c}^c \int_{-\infty}^{\infty} \frac{\sin tu - \sin \{t(u-x)\}}{t} dF(u) dt \end{aligned}$$

сходится равномерно и поэтому мы можем поменять порядок интегрирования. Тогда получим

$$I_c = \int_{-\infty}^{\infty} \int_{-c}^c \left\{ \frac{\sin tu}{t} - \frac{\sin \{t(u-x)\}}{t} \right\} dt dF(u). \quad (4.10)$$

Внутренний интеграл по c непрерывен и ограничен. Поэтому при $c \rightarrow \infty$ из (4.8) и (4.9) следует, что

$$\begin{aligned} \lim_{c \rightarrow \infty} I_c &= \pi \int_{-\infty}^{\infty} \{\operatorname{sgn} u - \operatorname{sgn}(u-x)\} dF(u) = \\ &= \pi [1 - 2F(0) - \{1 - 2F(x)\}] = 2\pi [F(x) - F(0)]. \end{aligned}$$

Тем самым соотношение (4.4) установлено. Формулу (4.7) можно вывести аналогичным путем, воспользовавшись тем, что

$$\frac{1}{2\pi} \int_0^{\infty} \frac{\sin t\xi}{t} dt = \operatorname{sgn} \xi. \quad (4.11)$$

Далее, если $f(x)$ существует, то путем дифференцирования в формуле (4.4) под знаком интеграла (что законно в силу его равномерной сходимости) мы придем к формуле (4.5). Дифференцируя (4.7), мы получим также

$$f(x) = \frac{1}{2\pi} \int_0^{\infty} [e^{itx}\varphi(-t) + e^{-itx}\varphi(t)] dt. \quad (4.12)$$

4.4 Из определения х. ф. следует, что $\varphi(t)$ и $\varphi(-t)$ комплексно сопряжены. Поэтому, если $\Re(t)$ и $\Im(t)$ — действительные и мнимые части $\varphi(t)$, то из (4.4) следует, что

$$F(x) - F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\Re(t) \sin xt + \Im(t) (1 - \cos xt)}{t} dt, \quad (4.13)$$

и из (4.7)

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\Re(t) \sin xt - \Im(t) \cos xt}{t} dt. \quad (4.14)$$

Точно так же

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\Re(t) \cos xt - \Im(t) \sin xt] dt. \quad (4.15)$$

4.5 Рассмотрим теперь величину J_c , определенную формулой

$$J_c = \frac{1}{2c} \int_{-c}^c e^{-ixt} \varphi(t) dt. \quad (4.16)$$

Если $F(x)$ имеет производную $f(x)$, то из (4.5) вытекает, что

$$\lim_{c \rightarrow \infty} J_c = 2\pi \lim f(x)/2c = 0$$

и, следовательно, J_c стремится к нулю во всех тех точках, где плотность распределения непрерывна.

Если плотность распределения разрывна, то рассмотрим какую-нибудь точку разрыва. Пусть в этой точке частота равна f_j . Тогда тот вклад, который эта величина вносит в $\varphi(t)$, равен $f_j \exp(itx_j)$. Соответствующий вклад этой частоты в J_c будет

равен

$$\frac{1}{2c} \int_{-c}^c f_j \exp(itx_j - itx) dt = \frac{1}{2c} f_j \left[\frac{\exp \{it(x_j - x)\}}{i(x_j - x)} \right]_{-c}^c.$$

Если $x_j \neq x$, то ясно, что при $c \rightarrow \infty$ это выражение стремится к нулю. Но если $x_j = x$, то рассматриваемый интеграл равен просто значению

$$\frac{1}{2c} \int_{-c}^c f_j dt = f_j.$$

Поэтому J_c стремится к f_j в точке $x = x_j$. Следовательно, если J_c стремится к нулю в какой-то точке, то в ней функция распределения непрерывна. Если же это выражение стремится к некоторому положительному числу f_j , то функция распределения разрывна в этой точке и ее скачок в ней равен f_j . Это дает нам критерий, позволяющий судить о том, отвечает данная х. ф. непрерывному распределению или же нет.

4.6 В том случае, когда функция распределения сосредоточена в точках x_j ($j=1, 2, \dots$), х. ф. согласно (4.6) задается формулой

$$\varphi(t) = \sum_j f_j \exp(itx_j)$$

и

$$\begin{aligned} F(x) &= \frac{1}{2} + \frac{1}{2\pi} \int_0^\infty \frac{\sum f_j \exp \{it(x - x_j)\} - \sum f_j \exp \{-it(x - x_j)\}}{it} dt = \\ &= \frac{1}{2} + \frac{1}{\pi} \sum f_j \int_0^\infty \frac{\sin t(x - x_j)}{t} dt = \frac{1}{2} + \frac{1}{2} \sum f_j \operatorname{sgn}(x - x_j). \end{aligned} \quad (4.17)$$

Таким образом, в точках $x = x_j$ функция распределения имеет скачки, равные f_j . Заметим, однако, что для $x = x_k$

$$\begin{aligned} F(x_k) &= \frac{1}{2} + \frac{1}{2} \sum_{j=1}^{k-1} f_j - \frac{1}{2} \sum_{j=k+1}^\infty f_j = \\ &= \frac{1}{2} \left(1 - \sum_{k+1}^\infty f_j \right) + \frac{1}{2} \sum_1^{k-1} f_j = \sum_1^{k-1} f_j + \frac{1}{2} f_k. \end{aligned} \quad (4.18)$$

Поэтому значение, приписываемое $F(x)$ в точке скачка, равно частоте, накопленной до этой точки, плюс половина величины скачка в рассматриваемой точке. В силу же нашего обычного

соглашения $F(x)$ непрерывна справа, и поэтому значение $F(x_k)$ должно быть равно $\sum_1^k f_j$. Об этом несоответствии надо помнить, когда разрывные функции распределения получают путем обращения х. ф.

Пример 4.1

В примере 3.4 мы установили, что для нормального распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} e^{-x^2/2\sigma^2}, \quad -\infty \leq x \leq \infty,$$

х. ф. равна

$$\varphi(t) = \exp\left(-\frac{1}{2} t^2 \sigma^2\right).$$

Предположим теперь, что задана эта функция $\varphi(t)$ и требуется найти распределение, если таковое существует, для которого $\varphi(t)$ является характеристической функцией.

Прежде всего мы замечаем, что если распределение существует, то оно должно быть непрерывным. В самом деле,

$$J_c = \frac{1}{2c} \int_{-c}^c \exp\left\{-\frac{1}{2} t^2 \sigma^2 - itx\right\} dt$$

и интеграл по модулю меньше, чем $\int_{-c}^c \exp\left(-\frac{1}{2} t^2 \sigma^2\right) dt$, что в

свою очередь меньше интеграла $\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} t^2 \sigma^2\right) dt = \sqrt{2\pi}/\sigma$.

Поэтому всюду $J_c \rightarrow 0$.

Если теперь предположить, что у искомого распределения имеется плотность $f(x)$, то найдем

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} t^2 \sigma^2 - itx\right\} dt = \\ &= \frac{1}{2\pi} e^{-x^2/2\sigma^2} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} \left(t\sigma + \frac{ix}{\sigma}\right)^2\right\} dt. \end{aligned}$$

Полученный интеграл можно рассматривать как интеграл в комплексной плоскости, взятый вдоль прямой, параллельной действительной оси. Принимая $t\sigma + ix/\sigma = u$ за новую переменную,

найдем, что интеграл равен $\sqrt{2\pi}/\sigma$. Поэтому

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

Эта функция всюду положительна и $\int_{-\infty}^{\infty} dF = 1$. Следовательно, $f(x)$ действительно есть плотность распределения, соответствующая заданной х. ф.

Пример 4.2

Найти плотность распределения, если только она существует, для которой

$$\varphi(t) = e^{-|t|}.$$

Поскольку здесь $J_c \rightarrow 0$, то мы замечаем, что если только $\varphi(t)$ является х. ф. некоторого распределения, то оно обязательно должно быть непрерывным. Значит, если плотность $f(x)$ существует, то она выражается формулой

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-|t|} e^{-itx} dt = \frac{1}{2\pi} \left[\int_0^{\infty} e^{-t} e^{-itx} dt + \int_{-\infty}^0 e^{t} e^{-itx} dt \right] = \\ &= \frac{1}{2\pi} \int_0^{\infty} e^{-t} (e^{itx} + e^{-itx}) dt = \frac{1}{\pi} \int_0^{\infty} e^{-t} \cos tx dt. \end{aligned}$$

Дважды интегрируя по частям, находим

$$\begin{aligned} f(x) &= \frac{1}{\pi} [-e^{-t} \cos tx]_0^{\infty} - \frac{x}{\pi} \int_0^{\infty} e^{-t} \sin tx dt = \\ &= \frac{1}{\pi} + \frac{x}{\pi} [e^{-t} \sin tx]_0^{\infty} - \frac{x^2}{\pi} \int_0^{\infty} e^{-t} \cos tx dt = \frac{1}{\pi} - x^2 f(x), \end{aligned}$$

откуда

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty \leq x \leq \infty.$$

Как и в предыдущем примере, эта функция является плотностью некоторого распределения. Можно показать, что $f(x)$ действительно имеет х. ф., совпадающую с заданной,

Пример 4.3

Существует ли распределение, обладающее характеристической функцией $\varphi(t) = e^{it}$? Мы имеем

$$J_c = \frac{1}{2c} \int_{-c}^c e^{it-tx} dt = \frac{1}{2c} \int_{-c}^c e^{it(1-x)} dt.$$

Если $x \neq 1$, то интеграл равен

$$\int_{-c}^c [\cos \{(1-x)t\} + i \sin \{(1-x)t\}] dt.$$

Поскольку $\sin t$ — нечетная функция, то последний интеграл равен

$$\int_{-c}^c \cos \{(1-x)t\} dt = \left[-\frac{\sin(1-x)t}{1-x} \right]_{-c}^c.$$

Хотя с ростом c это выражение к пределу не сходится, но зато оно ограничено. Поэтому $J_c \rightarrow 0$.

Если же, однако, $x=1$, то интеграл равен просто $2c$ и поэтому $J_c=1$.

Тем самым в точке $x=1$ распределение имеет скачок, равный по величине единице. Поэтому все распределение сосредоточено в точке $x=1$. Впрочем, это можно было бы легко усмотреть и из того факта, что $\log \varphi(t) = it$, и, следовательно, второй семинвариант, совпадающий с дисперсией, равен нулю.

Пример 4.4

Какой вид имеет распределение (если таковое существует), для которого семинварианты равны $k_r = (r-1)!$? Ряд

$$\sum_{j=0}^{\infty} k_j \frac{(it)^j}{j!} = \sum \frac{(it)^j}{j}$$

абсолютно сходится для $|t| < 1$ и поэтому является производящей функцией семинвариантов (п. ф. с.) некоторого распределения, если таковое существует. Обозначим п. ф. с. $\psi(t)$. Тогда

$$\psi(t) = \sum \frac{(it)^j}{j} = -\log(1-it),$$

и поэтому $\varphi(t) = \frac{1}{1-it}$.

Если плотность распределения существует, то она выражается формулой

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ixt}}{1-it} dt.$$

Этот интеграл можно вычислить, интегрируя функцию $\frac{e^{-ixz}}{1-iz}$ вдоль контура, состоящего из действительной оси и полуокружности бесконечного радиуса, лежащей в нижней полуплоскости. Интеграл вдоль действительной оси сводится к интегралу (с обратным знаком), который нам требуется отыскать. Рассмотрим теперь полуокружность радиуса R и положим $z = R(\cos \theta + i \sin \theta)$. В результате интегральное выражение станет равным

$$\frac{\exp(-ixR \cos \theta + xR \sin \theta)}{1 - iR \cos \theta + R \sin \theta},$$

где $\pi \leq \theta \leq 2\pi$ и, следовательно, $\sin \theta$ отрицателен. Поэтому если x положительно, то предыдущее выражение по модулю меньше, чем

$$\frac{e^{-xR |\sin \theta|}}{R},$$

что стремится к нулю при $R \rightarrow \infty$.

Функция $\frac{e^{-ixz}}{1-iz}$ внутри области интегрирования имеет полюс в точке $z = -i$, а соответствующий вычет равен ie^{-x} . Следовательно,

$$f(x) = \frac{1}{2\pi} 2\pi e^{-x} = e^{-x}, \quad 0 \leq x < \infty.$$

Вообще же, если $k_r = p(p-1)!$, $p > 0$, то аналогичным образом можно установить, что вычет функции $\frac{e^{-ixz}}{(1-iz)^p}$ равен $\frac{i^p x^{p-1} e^{-x}}{\Gamma(p)}$, откуда

$$f(x) = \frac{e^{-x} x^{p-1}}{\Gamma(p)}, \quad 0 \leq x < \infty, \quad p > 0.$$

Пример 4.5

Каково распределение, если таковое существует, для которого все нечетные семинварианты равны нулю, а все четные равны некоторой константе, скажем $2a$?

Для п. ф. с. имеем

$$\psi(t) = 2a \left\{ \frac{(it)^2}{2!} + \frac{(it)^4}{4!} + \dots \right\}.$$

Этот ряд сходится и $\psi(t) = 2a \cos(t-1)$. Следовательно, $\varphi(t) = e^{2a(\cos t - 1)}$.

Непосредственно проинтегрировать интеграл

$\int_{-\infty}^{\infty} e^{2a(\cos t - 1)} e^{-itx} dt$ не удастся. Поэтому исследуем вопрос о непрерывности функции распределения.

Имеем

$$J_c = \frac{e^{-2a}}{2c} \int_{-c}^c e^{2a \cos t} e^{-ixt} dt = \frac{e^{-2a}}{2c} \int_{-c}^c \sum_{j=0}^{\infty} \frac{(2a)^j}{j!} \cos^j t e^{-ixt} dt.$$

Входящий в подынтегральное выражение ряд равномерно сходится и поэтому

$$\begin{aligned} J_c &= \frac{e^{-2a}}{2c} \sum_{j=0}^{\infty} \int_{-c}^c \frac{(2a)^j}{j!} \cos^j t e^{-ixt} dt = \\ &= \frac{e^{-2a}}{2c} \sum_{j=0}^{\infty} \int_{-c}^c \frac{(2a)^j}{j!} \cos^j t \cos xt dt \end{aligned}$$

в силу того, что $\sin xt$ — нечетная функция.

Рассмотрим теперь интеграл $\int_{-\infty}^{\infty} 2^j \cos^j t \cos xt dt$. В силу хорошо известного разложения

$$\begin{aligned} 2^j \cos^j t \cos xt &= \frac{1}{2} (e^{it} + e^{-it})^j (e^{ixt} + e^{-ixt}) = \\ &= \frac{1}{2} (e^{ixt} + e^{-ixt}) \left\{ e^{ijt} + \binom{j}{1} e^{(j-2)it} + \dots + e^{-ijt} \right\}. \end{aligned}$$

Нас в этом выражении интересует только постоянный член, поскольку все остальные слагаемые вносят в J_c конечный вклад. Коэффициент при e^0 равен 0, если x нецелое. Если же x является целым числом, он равен

$$\frac{1}{2} \left\{ \binom{j}{\frac{1}{2}(j-x)} + \binom{j}{\frac{1}{2}(j+x)} \right\} = \binom{j}{\frac{1}{2}(j-x)}.$$

Поэтому $J_c \rightarrow 0$, если x — нецелое число. В случае целых x

$$J_c \rightarrow e^{-2a} \sum_{j=x}^{\infty} \frac{a^j}{j!} \binom{j}{\frac{1}{2}(j-x)}.$$

Следовательно, если x четно, скажем $2r$, то в точке $x = \pm 2r$ частота будет равна

$$e^{-2a} \left\{ \frac{a^{2r}}{(2r)!} + \frac{a^{2r+2}}{(2r+2)!} \binom{2r+2}{1} + \frac{a^{2r+4}}{(2r+4)!} \binom{2r+4}{2} + \dots \right\} = \\ = e^{-2a} \left\{ \frac{a^{2r}}{(2r)!} + \frac{a^{2r+2}}{(2r+1)!1!} + \frac{a^{2r+4}}{(2r+2)!2!} + \dots \right\}.$$

Если же $x = \pm (2r+1)$, то в этой точке частота равна

$$e^{-2a} \left\{ \frac{a^{2r+1}}{(2r+1)!} + \frac{a^{2r+3}}{(2r+2)!1!} + \frac{a^{2r+5}}{(2r+3)!2!} + \dots \right\}.$$

Можно убедиться, что эти частоты в сумме дают единицу и, следовательно, ими полностью описывается распределение.

Условия, при которых некоторая функция будет характеристической

4.7 Каждая неотрицательная интегрируемая по Лебегу функция может являться (с точностью до константы) плотностью некоторого распределения. Точно так же, всякая неубывающая функция, изменяющаяся от 0 до 1, может быть функцией распределения. Условия же, при выполнении которых данная функция является характеристической функцией некоторого распределения, много более ограничительны.

Заметим сначала, что необходимое и достаточное условие для того, чтобы функция $\varphi(t)$ была характеристической, состоит в том, чтобы функция

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(t) \frac{1 - e^{-ixt}}{it} dt$$

(с точностью до аддитивной константы $F(0)$) являлась функцией распределения. Однако на практике этот критерий не очень полезен, так как трудно проверить.

Вспоминая определение х. ф. $\varphi(t)$ как $\int e^{itx} dF(x)$, мы видим, что необходимое условие для того, чтобы $\varphi(t)$ была х. ф., состоит в следующем:

- а) $\varphi(t)$ непрерывна по t ,
- б) $\varphi(t)$ определена на каждом конечном t -интервале,
- в) $\varphi(0) = 1$,
- г) $\varphi(t) = \overline{\varphi(-t)}$,

$$\text{д) } |\varphi(t)| \leq \int_{-\infty}^{\infty} |e^{itx}| dF \leq 1 = \varphi(0).$$

4.8 Крамёром (1937) была доказана общая теорема, согласно которой для того, чтобы ограниченная и непрерывная функция $\varphi(t)$ была х. ф., необходимо и достаточно, чтобы $\varphi(0) = 1$ и интеграл

$$\int_0^A \int_0^A \varphi(t-u) \exp\{ix(t-u)\} dt du$$

был действительным и неотрицательным для всех действительных x и всех $A > 0$.

Следует также отметить ряд специальных теорем, посвященных тому же вопросу: когда некоторая функция является х. ф.? Так, например, Марцинкевич (1938) доказал, что $\exp P(t)$, где $P(t)$ — полином по t , не может быть х. ф., за исключением тех случаев, когда этот полином первой или второй степени. Лукач и Сас (1952) нашли условия, при которых рациональная функция является х. ф.

Предельные свойства распределений и характеристических функций

4.9 Мы сейчас рассмотрим ряд результатов, касающихся предельных свойств последовательностей х. ф. и ф. р. Пусть задана последовательность ф. р. $F_n(x)$, зависящая от неограниченно растущего параметра n . Каждой функции $F_n(x)$ соответствует х. ф. $\varphi_n(t)$. Вопрос, которым мы интересуемся, состоит в следующем: если $F_n(x)$ сходится к $G(x)$, то сходится ли последовательность $\varphi_n(t)$ к некоторому пределу $\varphi(t)$ и будет ли $\varphi(t)$ х. ф. для $G(x)$? И обратно, если φ_n стремится к пределу φ , то сходится ли последовательность F_n к некоторому пределу G и является ли G ф. р., для которой φ есть х. ф.? Ниже при довольно общих условиях мы дадим положительные ответы на поставленные вопросы.

4.10 Прежде всего остановимся на вопросе о том, в каком смысле следует понимать сходимость одной функции распределения к другой. Если все рассматриваемые функции непрерывны, то последовательность $F_n(x)$ называется сходящейся $G(x)$, если для каждого $\varepsilon > 0$ существует такое n_0 , что $|F_n - G| < \varepsilon$ для всех $n > n_0$ и всех x . Если же у рассматриваемых распределений имеются точки разрывов, то последовательность F_n называется сходящейся к G , если указанная выше сходимость имеет место в каждой точке непрерывности $G(x)$ *). Такое определение сходимости оказывается удобным и вполне достаточным для большинства статистических приложений.

4.11 Приведем некоторые результаты о сходимости ф. р. Первый из них состоит в том, что если последовательность ф. р. $\{F_n(x)\}$ сходится к ф. р. $G(x)$ во всех точках непрерывности последней, то эта сходимость равномерна на каждом замкнутом интервале непрерывности $G(x)$.

Пусть $G(x)$ непрерывна на отрезке $[a, b]$. Разобьем этот отрезок на конечное число частей точками $a = \xi_1, \xi_2, \dots, \xi_n = b$ так, чтобы на каждом частичном интервале приращение функции $G(x)$ не превышало ε . Выберем n_0 настолько большим, чтобы

*) Обычно этот вид сходимости называют сходимостью в основном или слабой сходимостью. (Прим. перев.)

$|F_n(\xi_j) - G(\xi_j)| < \varepsilon$, $j=1, \dots, h$; $n > n_0$. Тогда для каждого n и любого x из отрезка $[a, b]$ найдется такое r , что x лежит в (ξ_r, ξ_{r+1}) и

$$\begin{aligned} G(x) - 2\varepsilon &\leq G(\xi_r) - \varepsilon < F_n(\xi_r) \leq F_n(x) \leq \\ &\leq F_n(\xi_{r+1}) < G(\xi_{r+1}) + \varepsilon \leq G(x) + 2\varepsilon, \end{aligned} \quad (4.19)$$

и, следовательно,

$$|F_n(x) - G(x)| < 2\varepsilon.$$

Это справедливо для всех точек отрезка $[a, b]$, что и доказывает искомую равномерную сходимость.

В частности, если $G(x)$ непрерывна всюду, то сходимость (в основном) будет равномерной на всей прямой $-\infty \leq x \leq \infty$.

4.12 Вторая теорема (называемая теоремой Монтеля — Хелли) состоит в следующем: если последовательность $\{F_n(x)\}$ состоит из монотонных и ограниченных функций (что имеет место для ф. р.), то тогда можно выбрать такую подпоследовательность $\{F_{n_i}(x)\}$, которая сходится к некоторой монотонно неубывающей функции $G(x)$ (которая может и не быть ф. р.).

Рассмотрим некоторую последовательность точек $x_1, x_2, \dots$. Известно, что из каждого ограниченного множества чисел можно выбрать сходящуюся последовательность. Поэтому из последовательности $\{F_n(x_1)\}$ мы можем выбрать сходящуюся подпоследовательность $\{F_{n_1}(x_1)\}$. Из $\{F_{n_1}\}$ в свою очередь можно выбрать подпоследовательность $\{F_{n_2}\}$, которая будет сходиться как в точке x_1 , так и в x_2 . Продолжим этот процесс неограниченно и в каждой из подпоследовательностей $\{F_{n_i}\}$ выберем i -ю функцию. Назовем ее G_i . Рассмотрим теперь последовательность $G_1, G_2, \dots$. В силу сказанного последовательность $\{G_n(x)\}$ сходится в каждой из точек $x_1, x_2, \dots$ (Мы использовали здесь хорошо известный диагональный процесс Вейерштрасса.) Если с самого начала в качестве этих точек брать все рациональные точки, то тогда можно утверждать, что последовательность $G_n(x)$ сходится в каждой рациональной точке. Поскольку $G_n(a) \leq G_n(x) \leq G_n(b)$ для каждого $a \leq x \leq b$, то мы видим, что если $G_n(a)$ и $G_n(b)$ сходятся, то предельные значения $G_n(x)$ лежат между этими предельными значениями, скажем $G(a)$ и $G(b)$.

Тогда функция $u(x) = \lim_{n \rightarrow \infty} G_n(x)$ (где x уже теперь произвольно) определена, является неубывающей и, следовательно, имеет не более чем счетное число точек разрыва. Если в точке x функция $u(x)$ непрерывна, то возьмем y и z такими, чтобы $y < x < z$ и $u(z) - u(y) < \varepsilon$. Тогда, если a и b — рациональные точки, удовлетворяющие условию $y < a < x < b < z$, то $u(y) \leq G(a) \leq G(b) \leq u(z)$. Однако, поскольку все предельные значения функ-

ции $G_n(x)$ лежат между $G(a)$ и $G(b)$, то они заключены и между $u(y)$ и $u(z)$. Следовательно, в силу того, что ε может быть сделано сколь угодно малым, мы видим, что $G_n(x)$ стремится к $u(x)$ в каждой точке непрерывности $u(x)$. Наконец, используя диагональный процесс, мы сможем выбрать последовательность, которая будет также сходиться и в точках разрыва функции $u(x)$. Тем самым теорема доказана.

4.13 Вовсе не очевидно, что если значения каждой из функций последовательности $\{F_n(x)\}$ лежат между 0 и 1, то их предел $G(x)$ будет достигать этих крайних значений. Противоречащий пример можно получить следующим образом. Пусть

$$F_n(x) = \begin{cases} 0, & x < -n, \\ \frac{1}{2}, & -n \leq x \leq n, \\ 1, & x > n. \end{cases} \quad (4.20)$$

Предел последовательности $F_n(x)$ равен $1/2$ для всех x . Достаточное (но не необходимое) условие для того, чтобы предел $G(x)$ был функцией распределения, было дано Кендаллом и Рао (1950). Состоит оно в следующем. Если $F_n(x)$ имеет второй момент $\mu'_{2(n)} \leq A < \infty$ для всех $n > n_0$, то $G(x)$ является функцией распределения.

В самом деле, в силу неравенства Бьенэме — Чебышева (3.35)

$$F_n(-X) + 1 - F_n(X) \leq A/X^2, \quad n > n_0,$$

и, следовательно, для достаточно больших X

$$F_n(-X) + 1 - F_n(X) < \varepsilon,$$

откуда

$$G(-X) + 1 - G(X) \leq \varepsilon, \quad (4.21)$$

если только функция $G(X)$ непрерывна в точках $\pm X$, что можно предположить, не нарушая общности. Поэтому

$$G(-\infty) = 0, \quad G(\infty) = 1.$$

Первая предельная теорема

4.14 Докажем следующую теорему. Если последовательность функций распределения $\{F_n\}$ стремится к непрерывной ф. р. G , тогда соответствующая последовательность х. ф. $\{\varphi_n\}$ стремится к φ равномерно в каждом конечном t -интервале, где φ — х. ф. распределения G^*).

*) Теорема справедлива и без предположения о непрерывности функции $G(x)$. (Прим. перев.)

Итак, нам надо установить, что для заданного $\varepsilon > 0$ и фиксированного T -интервала $(-T, T)$ существует n_0 , не зависящее от значений t , принадлежащих этому интервалу, такое, что

$$|\varphi(t) - \varphi_n(t)| = \left| \int_{-\infty}^{\infty} e^{itx} (dG - dF_n) \right| < \varepsilon, \quad n > n_0.$$

Пусть X и $-X$ — две точки непрерывности функции $G(x)$. При этом X можно считать достаточно большим. Пусть

$$\int_{-\infty}^{\infty} e^{itx} (dG - dF_n) = \int_{-X}^X e^{itx} (dG - dF_n) + \int_{x > |X|} e^{itx} (dG - dF_n). \quad (4.22)$$

Имеем

$$\left| \int_{x > |X|} e^{itx} dG \right| \leq \int_{x > |X|} dG \leq 1 - G(X) + G(-X),$$

и, следовательно, беря X достаточно большим, мы можем сделать этот интеграл меньше, чем $\varepsilon/6$.

Аналогично

$$\left| \int_{x > |X|} e^{itx} dF_n \right| \leq 1 - F_n(X) + F_n(-X),$$

и поскольку F_n сходится к G (равномерно в интервале $|x| \leq X$), то, полагая X большим, можно считать этот интеграл меньше $\varepsilon/3$. Таким образом,

$$\left| \int_{x > |X|} e^{itx} (dG - dF_n) \right| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{6} = \frac{\varepsilon}{2}.$$

Рассмотрим теперь

$$\int_{-X}^X e^{itx} (dG - dF_n). \quad (4.23)$$

Это выражение является пределом сумм

$$\sum e^{itx_j} [G(\xi_{j+1}) - G(\xi_j)] - [F_n(\xi_{j+1}) - F_n(\xi_j)], \quad (4.24)$$

где ξ_j и ξ_{j+1} — граничные точки интервалов разбиения и x_j — произвольная точка из интервала (ξ_j, ξ_{j+1}) .

Разность между этой суммой и предельным значением может быть сделана меньше $\varepsilon/4$, если только интервал разбиения выбрать достаточно малым. В самом деле, пусть длины интервалов разбиения меньше η . Тогда по теореме о среднем разность между e^{itx_j} и e^{itx} (x принадлежит интервалу (ξ_j, ξ_{j+1})) меньше

$\eta|t|$. Поэтому разность между (4.23) и (4.24) по модулю меньше

$$\eta T \left| \sum [G(\xi_{j+1}) - G(\xi_j)] - \{F_n(\xi_{j+1}) - F_n(\xi_j)\} \right| < 2\eta T,$$

что в свою очередь меньше $\varepsilon/4$, когда $\eta < \varepsilon/(8T)$.

Сумма же (4.24) меньше $\varepsilon/4$ для некоторого $n > n_0$, ибо она состоит из конечного числа членов, каждый из которых стремится к нулю. Следовательно, (4.23) меньше $\varepsilon/2$, откуда

$$|\varphi(t) - \varphi_n(t)| < \varepsilon, \quad n > n_0.$$

Обращение первой предельной теоремы

4.15 Обратная теорема играет еще более важную роль. Пусть $\{\varphi_n\}$ — последовательность х. ф., отвечающих последовательности ф. р. $\{F_n\}$. Тогда, если $\varphi_n(t)$ сходится с $\varphi(t)$ для всех действительных значений t и эта сходимость равномерна *) в некотором конечном интервале $|t| < a$, то последовательность $\{F_n\}$ сходится к ф. р. G и φ является х. ф. для G .

Установим сначала один вспомогательный результат, состоящий в том, что если G есть ф. р. с х. ф. φ , то для всех действительных значений ξ и всех $h > 0$

$$\frac{1}{h} \int_{\xi}^{\xi+h} G(u) du - \frac{1}{h} \int_{\xi-h}^{\xi} G(u) du = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\frac{\sin t}{t} \right)^2 e^{-\frac{2it\xi}{h}} \varphi\left(\frac{2t}{h}\right) dt. \quad (4.25)$$

Действительно, положим

$$H(x) = \frac{1}{h} \int_x^{x+h} G(u) du.$$

Эта функция является непрерывной ф. р. и ее х. ф. равна

$$\int_{-\infty}^{\infty} e^{itx} dH = \frac{1}{h} \int_{-\infty}^{\infty} e^{itx} \{G(x+h) - G(x)\} dx.$$

Интегрируя последний интеграл по частям, убеждаемся, что он равен

$$\begin{aligned} \frac{1}{h} \left[\frac{\{G(x+h) - G(x)\} e^{itx}}{it} \right]_{-\infty}^{\infty} - \frac{1}{ith} \int_{-\infty}^{\infty} e^{itx} \{dG(x+h) - dG(x)\} = \\ = -\frac{1}{ith} \int_{-\infty}^{\infty} \{e^{it(x-h)} dG(x) - e^{itx} dG(x)\} = \varphi(t) \frac{1 - e^{-ith}}{ith}. \end{aligned}$$

*) Или, что эквивалентно, если $\varphi(t)$ непрерывна в точке $t=0$ или если $\varphi(t)$ есть х. ф.

Подставляя $H(x)$ в (4.4), найдем

$$\frac{1}{h} \int_{\xi+h}^{\xi+2h} G(u) du - \frac{1}{h} \int_{\xi}^{\xi+h} G(u) du = \frac{1}{2\pi h} \int_{-\infty}^{\infty} \left(\frac{1-e^{-it}}{it} \right)^2 e^{-it\xi} \varphi(t) dt,$$

откуда, заменяя ξ на $\xi+h$, после несложных преобразований придем к соотношению (4.25).

Возвращаясь теперь к теореме, которую надо доказать, заметим, что нам достаточно показать, что если $\varphi_n \rightarrow \varphi$ равномерно в некотором интервале $|t| < a$, то $\{F_n\}$ сходится к некоторой функции распределения G в каждой точке непрерывности последней. Если только это будет установлено, то тогда из первой предельной теоремы будет следовать, что φ является х.ф. для G и φ_n сходятся к φ равномерно в каждом конечном t -интервале.

Как показано в 4.12, из заданной последовательности $\{F_n\}$ мы можем выбрать некоторую подпоследовательность $\{F_{n'}\}$ такую, что $\{F_{n'}\}$ сходится к неубывающей функции G в каждой точке ее непрерывности.

Выберем такую подпоследовательность. Ясно, что $0 \leq G \leq 1$, причем функция G может считаться всюду непрерывной справа. Функция $G(x)$ будет ф.р., если $G(+\infty) - G(-\infty) = 1$. Покажем, что это условие в действительности выполнено. Полагая в (4.25) $\xi = 0$, получим

$$\frac{1}{h} \int_0^h F_{n'}(u) du - \frac{1}{h} \int_{-h}^0 F_{n'}(u) du = \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\frac{\sin t}{t} \right)^2 \varphi_{n'} \left(\frac{2t}{h} \right) dt.$$

Согласно предположению, φ_n сходятся равномерно к φ для $|t| < a$. Следовательно, этим же свойством обладает и последовательность $\varphi_{n'}$. При этом легко видеть, что в предшествующей формуле интеграл справа равномерно сходится. Поэтому при заданном $\varepsilon > 0$ можно найти такое h_0 , что для $h > h_0$

$$\left(\frac{1}{h} \int_0^h - \frac{1}{h} \int_{-h}^0 \right) G(u) du = \frac{1}{\pi} \int_{-\frac{1}{2}ah}^{\frac{1}{2}ah} \left(\frac{\sin t}{t} \right)^2 \varphi \left(\frac{2t}{h} \right) dt + \eta,$$

где $|\eta| < \varepsilon$. Пусть $h \rightarrow \infty$. Поскольку G является неубывающей функцией, левая часть стремится к $G(+\infty) - G(-\infty)$. Правая же часть в силу равномерной сходимости $\varphi_{n'}$ и, следовательно, в силу непрерывности φ в окрестности точки $t=0$ стремится

к интегралу

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \left(\frac{\sin t}{t} \right)^2 dt,$$

который равен единице.

Следовательно, предел последовательности $\{F_n\}$ является ф. р., причем ее х. ф. равна ф.

Но поскольку любая подпоследовательность из φ_n сходится к ф, то любая сходящаяся подпоследовательность из $\{F_n\}$ сходится к G . Следовательно, сама последовательность $\{F_n\}$ сходится к G в каждой ее точке непрерывности. Тем самым теорема доказана.

Пример 4.6

В примере 3.5 было, по существу, установлено, что биномиальное распределение имеет х. ф.

$$\varphi(t) = (q + pe^{it})^n.$$

Частота, отвечающая точке $x=j$, равна $\binom{n}{j} q^{n-j} p^j$. Она больше частоты в точке $x=j+1$, если

$$\binom{n}{j} q^{n-j} p^j > \binom{n}{j+1} q^{n-j-1} p^{j+1}$$

или если $j > pn - q$.

Для больших n максимум частоты расположен в окрестности точки $j = pn$ и, значит, он эквивалентен частоте

$$\binom{n}{pn} q^{qn} p^{pn}.$$

По формуле Стирлинга эта частота в свою очередь эквивалентна выражению

$$\frac{n^n e^{-n} \sqrt{2\pi n} q^{qn} p^{pn}}{(pn)^{pn} e^{-pn} \sqrt{2\pi pn} (qn)^{qn} e^{-qn} \sqrt{2\pi qn}} \sim \frac{1}{\sqrt{2\pi pqn}},$$

которое при $n \rightarrow \infty$ стремится к нулю.

Поэтому каждая частота в биномиальном распределении стремится с ростом n к нулю и предельного распределения не существует.

Рассмотрим, однако, это распределение в нормированной форме. Полагая $y = (x - \mu'_1)/\sigma$, мы получим

$$\varphi_x(t) = \int_{-\infty}^{\infty} e^{itx} dF(x) = \int_{-\infty}^{\infty} \exp\{it(\sigma y + \mu'_1)\} dF(y) = e^{it\mu'_1} \varphi_y(\sigma t),$$

откуда

$$\varphi_y(t) = \exp(-it\mu'_1/\sigma) \varphi_x(t/\sigma).$$

Таким образом, при этом преобразовании x в y $\varphi(t)$ преобразуется таким образом, что t заменяется на t/σ и умножается на $\exp(-it\mu'_1/\sigma)$. Для биномиального распределения $\mu'_1 = np$, $\mu_2 = npq$, и поэтому х.ф. биномиального распределения в нормированной форме имеет вид

$$\exp\left\{-\frac{itnp}{(npq)^{1/2}}\right\} \left\{q + p \exp \frac{it}{(npq)^{1/2}}\right\}^n,$$

откуда ($0 \leq |t| \leq 1$)

$$\begin{aligned} \log \varphi &= \frac{-itnp}{(npq)^{1/2}} + n \log \left\{1 + p \left(\exp \frac{it}{(npq)^{1/2}} - 1\right)\right\} = \\ &= \frac{-itnp}{(npq)^{1/2}} + n \log \left\{1 + p \frac{it}{(npq)^{1/2}} - \frac{pt^2}{2npq} + \frac{p\theta t^3}{6(npq)^{3/2}}\right\} = \\ &= n \left\{\frac{-pt^2}{2npq} - \frac{1}{2} \frac{(pit)^2}{npq}\right\} + O(t^3 n^{-1/2}) = -\frac{1}{2} t^2 + O(t^3 n^{-1/2}). \end{aligned}$$

Тем самым для каждого конечного t $\log \varphi$ стремится при $n \rightarrow \infty$ к $\left(-\frac{1}{2} t^2\right)$ и, следовательно,

$$\varphi(t) \rightarrow e^{-\frac{t^2}{2}}.$$

Итак, биномиальное распределение в нормированной форме стремится к распределению, х.ф. которого равна $e^{-t^2/2}$, т. е. к нормальному распределению

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2} dx, \quad -\infty \leq x \leq \infty.$$

Многомерные характеристические функции

4.16 Характеристическая функция двумерного распределения $F(x_1, x_2)$ определяется по формуле

$$\varphi(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{it_1 x_1 + it_2 x_2} dF(x_1, x_2), \quad (4.26)$$

и вообще х.ф. многомерного распределения $F(x_1, \dots, x_n)$ по определению равна

$$\varphi(t_1, \dots, t_n) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{it_1 x_1 + \dots + it_n x_n} dF(x_1, \dots, x_n). \quad (4.27)$$

Если $x_1, \dots, x_n$ независимы, то

$$\begin{aligned}\varphi(t_1, \dots, t_n) &= \int_{-\infty}^{\infty} e^{it_1 x_1} dF_1(x_1) \dots \int_{-\infty}^{\infty} e^{it_n x_n} dF_n(x_n) = \\ &= \varphi_1(t_1) \varphi_2(t_2) \dots \varphi_n(t_n), \quad (4.28)\end{aligned}$$

где φ_r — х. ф. распределения F_r . И аналогично для логарифма х. ф. (л. х. ф.) ψ получим

$$\psi(t_1, \dots, t_n) = \sum_{j=1}^n \log \varphi_j(t_j). \quad (4.29)$$

Тем самым х. ф. совместного распределения независимых вариантов равна произведению их х. ф., а л. х. ф. равен сумме л. х. ф. В выборочной теории этот результат играет чрезвычайно важную роль.

4.17 Обобщая (4.4), можно получить формулу

$$\begin{aligned}F(x_1, x_2, \dots, x_n) - F(0, 0, \dots, 0) &= \\ &= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1 - e^{ix_1 t_1}}{it_1} \dots \frac{1 - e^{ix_n t_n}}{it_n} \varphi(t_1, t_2, \dots, t_n) dt_1 \dots dt_n.\end{aligned} \quad (4.30)$$

Участвующие здесь интегралы следует понимать в смысле главного значения, т. е. как $\lim_{c \rightarrow \infty} \int_{-c}^c \dots \int_{-c}^c$.

Доказательство формулы (4.30) осуществляется так же, как в одномерном случае. Отметим лишь только, что вместо соотношения (4.10), например, в двумерном случае будем иметь следующее:

$$\begin{aligned}I_c &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-c}^c \int_{-c}^c \left\{ \frac{\sin t_1 u_1}{t_1} - \frac{\sin t_1 (u_1 - x_1)}{t_1} \right\} \times \\ &\quad \times \left\{ \frac{\sin t_2 u_2}{t_2} - \frac{\sin t_2 (u_2 - x_2)}{t_2} \right\} dt_1 dt_2 dF(u_1, u_2)\end{aligned}$$

и

$$\begin{aligned}\lim I_c &= \lim \int_{-\infty}^{\infty} \int_{-c}^c \left\{ \frac{\sin t_1 u_1}{t_1} - \frac{\sin t_1 (u_1 - x_1)}{t_1} \right\} dt_1 (2\pi) \times \\ &\quad \times \{dF(u_1, x_2) - dF(u_1, 0)\} = (2\pi)^2 \{F(x_1, x_2) - F(0, 0)\} \quad (4.31)\end{aligned}$$

и т. д. Как и в одномерном случае, х. ф. однозначно определяет ф. р., и соответствующие аналоги предельных теорем, рассмот-

ренных выше, остаются в силе. Из однозначности соответствия х.ф. и ф.р., в частности, следует справедливость результата, обратного к результату, сформулированному в 4.16: если х.ф. совместного распределения вариант равны произведению их х.ф., то эти варианты независимы.

4.18 Если мы имеем распределение $F(x)$ и некоторую функцию $y=y(x)$ от варианты x , то х.ф. для y определяется формулой

$$\varphi_y(t) = \int_{-\infty}^{\infty} \exp(ity) dF(x). \quad (4.32)$$

Путем ее обращения можно получить распределение y . Например, если $g(y)$ — плотность распределения y , то

$$g(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-ity) \varphi_y(t) dt. \quad (4.33)$$

Аналогичный результат справедлив и для функций g от нескольких переменных.

Проблема моментов

4.19 Рассмотрим вопрос о том, насколько моменты (в предположении их существования) определяют однозначным образом распределение. Чтобы показать, что два разных распределения могут обладать одними и теми же моментами, приведем такой пример. Рассмотрим интеграл

$$\int_0^{\infty} t^{p-1} e^{-qt} dt = \Gamma(p)/q^p, \quad p > 0, \quad \Re(q) > 0.$$

Положим $p = (n+1)/\lambda$; пусть n — неотрицательное целое число; $0 < \lambda < \frac{1}{2}$; $q = \alpha + i\beta$; $\beta/\alpha = \operatorname{tg} \lambda\pi$; $x^\lambda = t$. После их подстановки в интеграл мы найдем, учитывая действительность $(1 + i \operatorname{tg} \lambda\pi)^{(n+1)/\lambda}$, что мнимая часть интеграла равна нулю и что

$$\int_0^{\infty} x^n \exp(-\alpha x^\lambda) \sin(\beta x^\lambda) dx = 0. \quad (4.34)$$

Следовательно, распределения с плотностями

$$f(x) = k \exp(-\alpha x^\lambda) \{1 + \varepsilon \sin(\beta x^\lambda)\}, \quad 0 \leq x < \infty, \\ \alpha > 0, \quad 0 < \lambda < \frac{1}{2}, \quad |\varepsilon| < 1 \quad (4.35)$$

обладают для всех $\varepsilon (|\varepsilon| < 1)$ одинаковыми моментами. Эти распределения сосредоточены только на положительной полуоси. Нетрудно построить примеры распределений, сосредоточенных на всей прямой и имеющих одни и те же моменты. С этой целью положим $p = (2n + 1)/\rho$, $q = \alpha + i\beta$, $\beta/\alpha = \operatorname{tg} \frac{1}{2} \rho \pi$, $x^\rho = t$, $\rho = 2s/(s + 1)$, где s — положительное целое число. Тогда у семейства плотностей

$$f(x) = k \exp \{-\alpha |x|^\rho\} \{1 + \varepsilon \cos(\alpha |x|^\rho)\}, \quad -\infty \leq x \leq \infty, \quad (4.36)$$

все моменты совпадают.

4.20 В полной общности проблема моментов ставится следующим образом: имеются константы $c_0, c_1, c_2, \dots$. Требуется указать, при каких условиях они могут являться моментами некоторого распределения. Для статистических целей этот вопрос не представляет особого интереса*). Нас будет интересовать вопрос о единственности проблемы моментов: задано некоторое множество констант, являющихся моментами некоторого распределения. Спрашивается, могут ли они служить моментами другого распределения?

Прежде чем переходить к решению этой проблемы, заметим, что ее стоит рассматривать только тогда, когда существуют абсолютные моменты всех порядков. В самом деле, нетрудно видеть, что существуют заведомо различные распределения, для которых ограниченное число моментов конечны и совпадают, тогда как все остальные моменты бесконечны. К тому же, как известно, если некоторый момент четного порядка существует, то существуют и все моменты низшего порядка. В частности, если μ'_{2r} существует, то существуют и все абсолютные моменты до порядка $2r$. Поэтому мы будем рассматривать только тот случай, когда существуют абсолютные моменты всех порядков.

4.21 Мы начнем с доказательства теоремы о том, что моменты однозначно определяют распределение, если ряд $\sum_{j=0}^{\infty} v_j t^j / j!$ сходится для некоторого действительного $t \neq 0$. Для краткости в этом и следующих пунктах мы будем записывать v без штриха, хотя результат остается верным и для моментов, взятых относительно произвольной точки.

Характеристическая функция непрерывна по t , и ее производные в точке $t=0$ существуют, если существуют моменты.

*) По поводу общих результатов см. работы Стильтьеса (1918) и Гамбурга (1920).

Тогда в окрестности точки $t=0$

$$\varphi(t) = \sum_{j=0}^r (it)^j \mu'_j / j! + R_r, \quad (4.37)$$

где R_r по модулю меньше, чем $\frac{v_{r+1} |t|^{r+1}}{(r+1)!}$. Поэтому, если $\sum v_j t^j / j!$ сходится, то $v_j t^j / j!$ стремится к нулю и, следовательно, $\varphi(t)$ представима в виде ряда $\sum_{j=0}^{\infty} (it)^j \mu'_j / j!$, если только он сходится. Однако этот ряд мажорируется рядом $\sum v_j t^j / j!$ и, следовательно, абсолютно сходится, если сходится этот последний. Итак,

$$\varphi(t) = \sum_{j=0}^{\infty} (it)^j \mu'_j / j!, \quad (4.38)$$

и поэтому $\varphi(t)$ однозначно определена в окрестности $t=0$. В окрестности же $t=t_0$

$$\varphi(t) = \sum_{j=0}^r \left\{ \frac{t^j (t-t_0)^j}{j!} \int_{-\infty}^{\infty} x^j e^{it_0 x} dF \right\} + R'_r,$$

и модуль коэффициента при $(t-t_0)^j / j!$ не больше, чем v_j . Следовательно, функция $\varphi(t)$ представляется всюду сходящимся рядом Тейлора и совпадает с суммой этого ряда. Таким образом, $\varphi(t)$ можно аналитически продолжить из окрестности точки $t=t_0$ на любой конечный t -интервал. Поэтому $\varphi(t)$ всюду однозначно определена. Но $\varphi(t)$ полностью определяет функцию распределения и, следовательно, последняя определена однозначно.

4.22 Несколько простых, но эффективных результатов можно получить теперь в виде следствий.

(а) Моменты однозначно определяют распределение, если

$$\limsup_{n \rightarrow \infty} (v_n^{1/n}) / n < \infty.$$

Это следует из того, что ряды с общим членом $v_n t^n / n!$ сходятся, если

$$\limsup_{n \rightarrow \infty} (v_n t^n / n!)^{1/n} < 1.$$

Заменяя факториал по формуле Стирлинга, мы видим, что предыдущее неравенство выполнено, если

$$\limsup_{n \rightarrow \infty} (v_n^{1/n}) / n < k/t,$$

где k — некоторая постоянная. Это неравенство выполнено при некотором ненулевом t , если $\limsup_{n \rightarrow \infty} (v_n^{1/n})/n < \infty$.

(б) Достаточным условием для единственности решения проблемы моментов является также условие

$$\limsup_{n \rightarrow \infty} \frac{(\mu'_{2n})^{1/2n}}{2n} < \infty,$$

не содержащее абсолютных моментов. Нетрудно видеть, что это условие и условие (а) эквивалентны, откуда и следует достаточность сформулированного условия.

(в) Моменты однозначно определяют распределение, если только известно, что они являются моментами распределений, сосредоточенных на конечном интервале (скажем, длины h). Действительно, беря моменты относительно левого конца интервала, на котором сосредоточены распределения, получим $v_r \leq h^r$, и, следовательно, $(v_n^{1/n})/n \leq h/n$, что стремится к нулю.

4.23 Упомянем еще о двух критериях. Первый из них был дан Карлеманом (1925): моменты однозначно определяют распределение (сосредоточенное на $(-\infty, \infty)$), если

$$\sum_{j=0}^{\infty} \frac{1}{(\mu_{2j})^{1/2j}} = \infty. \quad (4.39)$$

В случае распределений, сосредоточенных на $(0, \infty)$, достаточно требовать, чтобы

$$\sum_{j=0}^{\infty} \frac{1}{(\mu_j)^{1/2j}} = \infty. \quad (4.40)$$

Второй критерий состоит в следующем. Распределение, имеющее плотность $f(x)$, определяется моментами однозначно, если (в случае $-\infty \leq x \leq \infty$)

$$f(x) < M|x|^{\beta-1} \exp(-\alpha|x|^\lambda) \text{ для } |x| > x_0; \quad M, \beta, \alpha, > 0; \lambda \geq 1, \quad (4.41)$$

или если (в случае $0 \leq x \leq \infty$)

$$f(x) < M|x|^{\beta-1} \exp(-\alpha|x|^\lambda) \text{ для } |x| > x_0; \quad M, \beta, \alpha, > 0; \lambda \geq \frac{1}{2}. \quad (4.42)$$

Этот критерий был впервые предложен Стильтьесом. Его легко получить из критерия Карлемана.

Интересно отметить, что если для некоторого x_0

$$f(x) > \exp(-\alpha|x|^\lambda), \quad \alpha > 0, \quad x > x_0, \quad (4.43)$$

то проблема моментов необходимо неопределенная (как обычно, $\lambda < 1/2$ для $(0, \infty)$ и $\lambda < 1$ для $(-\infty, \infty)$). Это следует из рассмотрения примеров распределений (4.35) и (4.36). В самом деле, наряду с $f(x)$ теми же моментами обладает и распределение, плотность которого получена путем добавления к $f(x)$ соответствующих членов из (4.35) и (4.36) с нулевыми моментами.

Пример 4.7

Моменты нормального распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} \exp(-x^2/2\sigma^2), \quad -\infty \leq x \leq \infty,$$

задаются (см. пример 3.4) формулами

$$\mu_{2r+1} = 0, \quad \mu_{2r} = \frac{(2r)!}{2^r r!} \sigma^{2r}.$$

Из формулы Стирлинга, примененной к факториалам в $(\mu_{2n}^{1/2n})/2n$, следует, что

$$\frac{\sigma}{2n \sqrt{2}} \left[\frac{e^{-2n} (2n)^{2n} \sqrt{4\pi n}}{e^{-n} n^n \sqrt{2\pi n}} \right]^{1/2n} \sim \frac{\sigma}{(2en)^{1/2}}.$$

Тем самым верхний предел последовательности $(\mu_{2n}^{1/2n})/2n$ равен нулю и распределение однозначно определяется своими моментами.

4.24 Крамер и Волд (1936) распространили критерий Карлемана на случай многомерных распределений. Рассмотрим моменты маргинальных (одномерных) распределений $\mu'_{r00} \dots$, $\mu'_{0r0} \dots$, $\mu'_{00r} \dots$ и т. д. Пусть сумма этих величин равна λ_r . Критерии (4.39) и (4.40) останутся в силе, если μ_r заменить на λ_r ; например, распределение, сосредоточенное на всей прямой, однозначно определяется по моментам, если

$$\sum_{j=0}^{\infty} \frac{1}{(\lambda_{2j})^{1/2j}} = \infty. \quad (4.44)$$

4.25 Если у распределения существует момент порядка r , то его можно получить из r -й производной характеристической функции при $t=0$. Поэтому, если $\varphi(t)$ можно разложить в бесконечный ряд Тейлора, то он должен быть не чем иным, как рядом $\sum (it)^j \mu'_j/j!$. Если этот ряд не сходится, то $\varphi(t)$ не может быть представлена рядом Тейлора. В случае существования только r моментов $\varphi(t)$ можно выразить асимптотически в виде

частной суммы с остаточным членом

$$\varphi(t) = \sum_{j=0}^r (it)^j \mu'_j / j! + R.$$

В том случае, когда ряд $\sum (it)^j \mu'_j / j!$ не сходится, возникает ряд трудностей. Так, известно, что существует бесконечное число функций, обладающих одними и теми же асимптотическими разложениями (с конечным числом членов). Например, асимптотические разложения для функций $\alpha(t)$ и $\alpha(t) + kt^{-\log t}$ совпадают. Поэтому неудивительно, что в случае расходимости ряда $\sum (it)^j \mu'_j / j!$ может существовать более одного распределения с одними и теми же моментами.

Но это, конечно, не означает, что в случае расходимости ряда $\sum (it)^j \mu'_j / j!$ необходимо существует более чем одно распределение. Может случиться, например, что функции, чьи моменты совпадают с моментами данной функции распределения, сами не будут функциями распределения. Например, на некотором участке они могут быть отрицательными. В примере, указанном выше, функция $t^{-\log t}$ не может быть характеристической, поскольку она не удовлетворяет хорошо известному необходимому условию для характеристических функций, состоящему в том, что $\varphi(t)$ и $\varphi(-t)$ должны быть комплексно сопряжены.

4.26 Теперь мы перейдем к так называемой второй предельной теореме, касающейся вопроса о том, как последовательность функций распределения $\{F_n(x)\}$ сходится к пределу, если соответствующая последовательность моментов $\{\mu_j(n), j \geq 1\}$ сходится к $\{\mu_j, j \geq 1\}$? (Для упрощения записи штрихи у моментов опускаются.) Мы будем следовать доказательству, предложенному Кендаллом и Рао (1950).

Прежде чем к нему переходить, установим один вспомогательный результат. Если μ_{2m} существует, то для всех действительных t

$$\varphi(t) = \sum_{j=0}^{2m-1} (it)^j \mu_j / j! + \rho(it)^{2m} \mu_{2m} / (2m)!; \quad (4.45)$$

ρ здесь таково, что $|\rho| < 1$ и $\lim_{t \rightarrow 0} \rho = 1$.

Для случая $m=0$ эта формула тривиальна. Пусть $m>0$; для всех действительных x и t справедливо равенство

$$e^{ixt} = \sum_{j=1}^n \frac{(ixt)^j}{j!} + \rho' \frac{(ixt)^{n+1}}{(n+1)!}, \quad |\rho'| \leq 1. \quad (4.46)$$

Подставляя это разложение в интеграл, определяющий $\varphi(t)$, мы получим при $n=2m-1$ равенство (4.45). Для t и μ_{2m} , равных нулю, мы можем взять $\rho=1$; для других значений, рассматривая разность выражений (4.46), взятых с n и $n+1$ членами, мы найдем

$$|\rho' - 1| \leq |xt|/(n+2). \quad (4.47)$$

Поскольку

$$|\rho - 1| \mu_{2m} = \int_{-\infty}^{\infty} (\rho' - 1) x^{2m} dF,$$

то

$$|\rho - 1| \mu_{2m} < \varepsilon + \int_{-X}^X |\rho' - 1| x^{2m} dF$$

для любого ε , если X достаточно велико (в противном случае момент μ_{2m} не был бы конечным). Но в силу (4.47) при фиксированном X , взяв достаточно малое t , интеграл можно сделать сколь угодно малым. Поэтому $|\rho - 1| \mu_{2m} < 2\varepsilon$ и $\rho \rightarrow 1$ при $t \rightarrow 0$.

4.27 Вторым результатом (являющийся обратным к предыдущему), который также будет нам необходим, состоит в следующем. Если для некоторого положительного целого m

$$\varphi(t) = \sum_{j=0}^{2m} (it)^j \lambda_j / j! + o(t^{2m}), \quad (4.48)$$

когда действительное $t \rightarrow 0$, то первые $2m$ моментов существуют и $\mu_r = \lambda_r$, $r=0, 1, \dots, 2m$.

Пусть δ^2 обозначает операцию взятия центральной разности с шагом $2h$, т. е.

$$\delta^2 A(t) = A(t+2h) - 2A(t) + A(t-2h).$$

Тогда, если остаточный член в (4.48) есть $t^{2m}B(t)$, где $B(t) \rightarrow 0$ ($t \rightarrow 0$), то несложное вычисление показывает, что при $t=0$

$$\left(\frac{\delta^2}{4h^2}\right)^m t^{2m}B(t) = \sum_{j=0}^{2m} \binom{2m}{j} (-1)^j (m-j)^{2m} B(2mh-2jh).$$

Поэтому для $|B(t)| < \varepsilon$

$$\left(\frac{\delta^2}{4h^2}\right)^m t^{2m}B(t) \leq \sum_{j=0}^{2m} \binom{2m}{j} |m-j|^{2m} \varepsilon, \quad (4.49)$$

когда $|h| \leq \eta/2m$ и $|t| \leq \eta$. Следовательно, если соотношение (4.48) верно, то

$$\left[\lim_{h \rightarrow 0} \left(\frac{\delta^2}{4h^2}\right)^m t^{2m}B(t) \right]_{t=0} = 0. \quad (4.50)$$

Но оператор $\lim_{h \rightarrow 0} \left(\frac{\delta^2}{4h^2} \right)^m$ (обозначим его D^{2m}) есть не что иное, как оператор взятия $(2m)$ -й производной. Применяя этот оператор к (4.48) и учитывая (4.50), мы найдем, что при $t=0$

$$D^{2m}\varphi(t) = (-1)^m \lambda_{2m}. \quad (4.51)$$

Но $\delta^2 e^{ixt} = -(2 \sin xh)^2 e^{ixt}$, поэтому

$$\lim_{h \rightarrow 0} \int_{-\infty}^{\infty} \left(\frac{\sin xh}{xh} \right)^{2m} x^{2m} dF = \lambda_{2m}. \quad (4.52)$$

Из равномерной сходимости к единице последовательности $(\sin xh/xh)^{2m}$ мы видим, что для любого конечного интервала (a, b)

$$\int_a^b x^{2m} dF \leq \lambda_{2m},$$

и поэтому момент порядка $2m$ конечен. Следовательно, из (4.45) при $t \rightarrow 0$ следует, что

$$\varphi(t) = \sum_{j=0}^{2m-1} (it)^j \mu_j / j! + \frac{(it)^{2m} \mu_{2m}}{(2m)!} [1 + o(1)]. \quad (4.53)$$

Сравнение этого разложения с (4.48) показывает, что

$$\mu_r = \lambda_r, \quad r = 0, 1, \dots, 2m.$$

Из результатов этого раздела следует, что если при $t=0$ существует $(2m)$ -я производная х. ф., то $(2m)$ -й момент также существует. Мы уже видели в 3.5, что в случае существования моментов существуют и соответствующие производные у х. ф. в нуле (результат, обратный результату, установленному выше для четных моментов). Для нечетных моментов существование производной при $t=0$ еще не влечет за собой существования соответствующего момента. Питмэн (Pitman, 1956, Ann. Math. Statist. 27, 1156) показал, что необходимое и достаточное условие существования соответствующей производной состоит в том (1), что существует в смысле главного значения момент μ'_2 и (2), что

$$\lim_{x \rightarrow \infty} x^r \{F(-x) + 1 - F(x)\} = 0.$$

4.28 Результаты предшествующих двух разделов необходимы для установления (при некоторых условиях) единственности асимптотических разложений х.ф. Мы сейчас докажем некоторые теоремы о последовательностях распределений и моментов.

Пусть распределения $F_n(x)$, являющиеся n -ми членами последовательности $\{F_n\}$, обладают конечным моментом j -го

порядка, скажем $\mu_j(n)$, для всех $n > n_0$. Пусть

$$\lim_{n \rightarrow \infty} \mu_j(n) = \lambda_j$$

для каждого j и пусть $\{F_n(x)\}$ сходится к предельной функции $G(x)$ во всех точках непрерывности последней. (Функция $G(x)$ необходимо ограничена, монотонно возрастает и может считаться непрерывной справа.) Тогда $G(x)$ — функция распределения, обладающая моментами всех порядков; $\{\lambda_j\}$ — последовательность моментов; λ_j есть j -й момент $G(x)$.

Тот факт, что $G(x)$ — ф. р., следует из 4.13, ибо вторые моменты образуют сходящуюся и, следовательно, ограниченную последовательность. Поэтому согласно первой предельной теореме (см. 4.14) последовательность х. ф. $\{\varphi_n(t)\}$ сходится к х. ф. распределения $G(x)$, скажем к $\varphi(t)$. Тогда для $n > n_0$

$$\varphi_n(t) = \sum_{j=0}^{2m-1} (it)^j \mu_j(n)/j! + \rho(it)^{2m} \mu_{2m}(n)/(2m)!, \quad (4.54)$$

где последний член в правой части стремится к пределу, поскольку все остальные члены имеют пределы. Поэтому

$$\varphi(t) = \sum_{j=0}^{2m-1} (it)^j \lambda_j/j! + R(it)^{2m} \lambda_{2m}/(2m)!, \quad |R| \leq 1. \quad (4.55)$$

Это верно для всех m . Используя методы 4.26, легко можно показать, что $R \rightarrow 1$ при $t \rightarrow 0$. Отсюда и из результатов 4.27 вытекает справедливость нашей теоремы.

Пример 4.8

Заметим, что результат 4.28 может быть верен и в том случае, когда F_n не обладают моментами всех порядков. Например, распределение

$$dF = \frac{k dt}{(1 + t^2/v)^{\frac{1}{2}(v+1)}}, \quad -\infty \leq t \leq \infty, \quad v > 1,$$

обладает моментами только до порядка $(v-1)$ включительно. Но (см. пример 3.3)

$$\mu_{2r} = \frac{\Gamma\left(\frac{1}{2}v - r\right) \Gamma\left(r + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}v\right) \Gamma\left(\frac{1}{2}\right)} v^r, \quad (4.56)$$

откуда, применяя формулу Стирлинга к Γ -функции, получим, что

$$\mu_{2r} \rightarrow \frac{(2r)!}{2^r r!}, \quad (4.57)$$

т. е. имеет место сходимость к моментам нормированного нормального распределения (пример 3.4). Таким образом, если предельное распределение существует, оно необходимо будет нормальным, все моменты которого существуют.

Непосредственно можно проверить также, что последовательность плотностей рассмотренных выше распределений при $v \rightarrow \infty$ сходится к плотности нормального закона, т. е. $k \exp\left(-\frac{1}{2} t^2\right)$.

Вторая предельная теорема

4.29 Пусть $\{F_n(x)\}$ сходится к $G(x)$ и пусть моменты $\mu_j(n)$ существуют для $n > n_0$ и для всех $j \geq 0$. Пусть, далее, $\mu_j(n)$ ограничены сверху некоторой константой A_j . Тогда все моменты λ_j распределения $G(x)$ существуют и $\mu_j(n) \rightarrow \lambda_j$ при $n \rightarrow \infty$.

Из первой предельной теоремы вытекает, что последовательность х. ф. $\{\varphi_n(t)\}$ сходится к $\varphi(t)$, являющейся х. ф. распределения $G(x)$. Далее

$$\varphi_n(t) = 1 + it\mu_1(n) + O(t^2), \quad n > n_0,$$

где константа, входящая в $O(t^2)$, может быть взята не зависящей от n и t . Поэтому если m — некоторое значение $> n_0$, то

$$|\mu_1(n) - \mu_1(m)| \leq A_1 |t| + \delta_{m,n},$$

где $\delta_{m,n} \rightarrow 0$, когда m и n независимым образом стремятся к бесконечности. Поскольку это верно для каждого ненулевого t , то

$$\lim_{m, n \rightarrow \infty} |\mu_1(n) - \mu_1(m)| = 0,$$

и, следовательно, существует константа λ_1 такая, что

$$\lim_{n \rightarrow \infty} \mu_1(n) = \lambda_1. \quad (4.58)$$

Тем же самым методом можно установить существование предела и для $\mu_2(n)$. Действительно, из (4.54) в силу существования шестого момента

$$\varphi_n(t) = 1 + it\mu_1(n) + (it)^2 \mu_2(n)/2! + O(t^3).$$

Поэтому

$$|\mu_2(n) - \mu_2(m)| \leq \frac{2}{3} A_6^{\frac{1}{2}} |t| + \delta'_{m,n},$$

откуда, как и раньше, следует существование предела

$$\lim_{n \rightarrow \infty} \mu_2(n) = \lambda_2. \quad (4.59)$$

Продолжая эти рассуждения далее, мы установим существование моментной последовательности $\{\lambda_j\}$. Из теоремы 4.28 следует,

что эти моменты в действительности являются моментами распределения $G(x)$.

4.30 Для многих целей оказывается полезным следующее обращение второй предельной теоремы: пусть моменты $\mu_j(n)$ существуют и пределы $\lim_{n \rightarrow \infty} \mu_j(n) = \lambda_j$ являются моментами не- которого распределения $G(x)$, для которого решение *проблемы моментов единственно*. Тогда $\{F_n\}$ сходятся к $G(x)$ во всех точках непрерывности функции $G(x)$.

В последовательности ф.р. $\{F_n\}$ мы всегда можем выделить подпоследовательность, сходящуюся к некоторой функции распределения. Из теоремы 4.28 следует, что функция распределения с моментами λ_j единственна. Поэтому эта ф.р. совпадает с $G(x)$.

Предположим теперь, что в некоторой точке непрерывности, скажем в точке $x=a$,

$$\alpha \equiv \limsup F_n(a) \neq G(a).$$

Тогда можно выбрать подпоследовательность, сходящуюся к α в точке $x=a$, и подпоследовательность, сходящуюся и к α и к $G(a)$, поскольку $G(x)$ непрерывна в точке $x=a$. Этого противоречия можно избежать лишь тогда, когда

$$\limsup F_n(a) = G(a),$$

и аналогично только тогда, когда

$$\liminf F_n(a) = G(a),$$

что и доказывает обратную теорему.

Пример 4.9

Дискретное распределение с частотами $e^{-\lambda} \lambda^j / j!$ в точках $x=j$ ($j=0, 1, \dots$) имеет х.ф. (пример 3.10)

$$\varphi(t) = \exp \{ \lambda (e^{it} - 1) \},$$

и, следовательно, все его семиинварианты равны λ . Очевидно, не существует других распределений с такими же семиинвариантами, поскольку ряд $\sum \kappa_j (it)^j / j!$ сходится и его сумма равна $\lambda(e^{it} - 1)$. Таким образом, ф.р. и х.ф. определяются однозначно.

Пусть $\lambda \rightarrow \infty$. Частота в точке x_j , равная $e^{-\lambda} \lambda^j / j!$, стремится к нулю, и поэтому последовательность таких распределений не сходится ни к какому пределу. Это, между прочим, согласуется и с поведением семиинвариантов, которые стремятся к бесконечности.

Рассмотрим, однако, нормированное распределение. Тогда

$$\kappa_1 = 0, \quad \kappa_r = \lambda / \kappa_2^{\frac{1}{2}r} = \lambda^{-(\frac{1}{2}r-1)}, \quad r \geq 2.$$

Таким образом, при $\lambda \rightarrow \infty$ $\kappa_r \rightarrow 0$ для всех $r > 2$. И следовательно, семинварианты при $\lambda \rightarrow \infty$ сходятся к семинвариантам нормального распределения

$$dF = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx, \quad -\infty \leq x \leq \infty.$$

Мы знаем, что это распределение полностью определяется своими моментами (пример 4.7). Мы также знаем, что семинварианты определяют однозначно моменты и наоборот. Так что из сходимости семинвариантов следует и сходимость моментов. Следовательно, здесь применима обратная форма второй предельной теоремы, и мы видим, что это дискретное распределение в нормированной форме действительно сходится к нормальному.

4.31 В связи с той двойственностью, которая существует между x . ф. и ф. р., стоит отметить, что поведение одной из этих функций в окрестности нуля связано с поведением другой на бесконечности. В самом деле, m -я производная $\varphi(t)$ при $t=0$ пропорциональна m -му моменту, существование которого зависит от поведения $x^m f(x)$ на бесконечности. Обратно, из (4.5) мы видим, что m -я производная $f(x)$ в случае ее существования дается интегралом

$$\int_{-\infty}^{\infty} (-it)^m \varphi(t) e^{-ixt} dt / 2\pi, \text{ модуль которого не больше, чем интеграл}$$

$$\int_{-\infty}^{\infty} t^m \varphi(t) dt / 2\pi. \text{ Сходимость последнего интеграла зависит от поведения } t^m \varphi(t) \text{ на бесконечности.}$$

4.32 Естественнo задать вопрос, существуют ли другие преобразования типа Фурье от функций распределения, обладающие свойствами, присущими x . ф.? Ответ на этот вопрос отрицателен. Лукачем (1952) была доказана следующая теорема. Пусть $\kappa(x, t)$ — комплексная функция, определенная для всех действительных x и t , ограниченная и измеримая по t . Определим

$$\varphi(t) = \int_{-\infty}^{\infty} \kappa(x, t) dF(x),$$

и пусть выполнены следующие два условия:

$$(1) \quad \varphi_1(t) \equiv \varphi_2(t) \text{ тогда и только тогда, когда } F_1(x) = F_2(x);$$

$$(2) \text{ если } F(x) = \int_{-\infty}^{\infty} F_1(x-y) dF_2(y), \text{ то } \varphi(t) = \varphi_1(t) \varphi_2(t).$$

Тогда $\kappa(x, t)$ имеет вид $\exp\{itA(x)\}$, где $A(x)$ — действительная функция, принимающая значения из всюду плотного множества на числовой прямой. Замечательным здесь является то обстоятельство, что в условиях этой теоремы не требуется выполнения ни теоремы обращения, ни предельных теорем. Тем самым можно сказать, что преобразование Фурье автоматически обладает этими полезными и желаемыми свойствами.

УПРАЖНЕНИЯ

4.1 Показать, что если плотность распределения $f(x)$ есть симметричная функция, то х. ф. $\varphi(t)$ — четная функция, т. е. $\varphi(t) = \varphi(-t)$, и, следовательно, действительная. Обратно, если $\varphi(t)$ есть действительная х. ф. распределения, обладающего плотностью, то последняя симметрична.

4.2 Показать, что функция

$\varphi(t) = \left(\frac{e^{it} - 1}{it} \right)^n$ (n — целое положительное) является х. ф. распределения

$$f(x) = \frac{1}{(n-1)!} \sum_{j=0}^{[x]} (-1)^j \binom{n}{j} (x-j)^{n-1}.$$

4.3 Показать, что распределение с х. ф. $1/(1+t^2)$ имеет плотность $e^{-|x|}$, $-\infty \leq x \leq \infty$, а распределение с х. ф. $\cos \frac{1}{2} \pi t / (1-t^2)$ обладает плотностью

$$\frac{1}{2} \cos x, \quad -\frac{1}{2} \pi \leq x \leq \frac{1}{2} \pi.$$

4.4 Если для некоторого распределения

$$\kappa_r = b a^r,$$

где $a, b > 0$, то это распределение дискретно, сосредоточено в точках $0, a, \dots, ra$, ... с частотами, равными $e^{-b} b^r / r!$.

4.5 Показать, что функция $\exp(-t^a)$ не может быть характеристической за исключением случая $a=2$.

4.6 Показать, что существует только одно распределение с моментами

$$\mu'_r = \Gamma(v+r)/\Gamma(v)$$

и что это распределение задается формулой

$$dF = \frac{1}{\Gamma(v)} e^{-x} x^{v-1} dx, \quad 0 \leq x \leq \infty.$$

4.7 Показать, что распределение

$$dF = \frac{\sqrt{2}}{\pi} \frac{dx}{1+x^4}, \quad 0 \leq x \leq \infty,$$

имеет х. ф.

$$\varphi(t) = \sqrt{2} \exp(-|t|/\sqrt{2}) \sin\left(|t|/\sqrt{2} + \frac{1}{4} \pi\right).$$

4.8 Показать, что распределение

$$dF = \frac{nb^{2n-1} \sin(\pi/2n)}{\pi(b^{2n} + x^{2n})} dx, \quad -\infty \leq x \leq \infty, \quad b > 0, \quad n - \text{целое}$$

имеет х. ф.

$$\begin{aligned} \varphi(t) = & \sum_{s=0}^{n-1} \exp\left(-b|t| \sin \frac{2s+1}{2n} \pi\right) \times \\ & \times \sin\left(\frac{2s+1}{2n} \pi + b|t| \cos \frac{2s+1}{2n} \pi\right) \sin \frac{\pi}{2n}. \end{aligned}$$

4.9 Показать, что х. ф. распределения

$$dF = \frac{dx}{\operatorname{ch} \pi x}, \quad -\infty \leq x \leq \infty,$$

равна

$$\varphi(t) = \operatorname{sch} \frac{1}{2} t.$$

4.10 Показать, что $(1+t^{2k})^{-1}$ не может быть х. ф. при $k > 1$; $(1-t^2)^{-1}$ также не может быть х. ф.

4.11 Используя теорему Марцинкевича (4.8), показать, что если все семинварианты κ_r некоторого распределения обращаются в нуль при всех $r \geq r_0 > 2$, тогда все семинварианты равны нулю для всех $r > 2$ и, следовательно, рассматриваемое распределение нормально.

4.12 Показать, что распределение

$$F_n(x) = \left(1 - \frac{1}{n}\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}u^2} du + \frac{1}{2n} \{1 + \operatorname{sgn}(x-n)\}$$

стремится к нормальному при $n \rightarrow \infty$, но что высшие моменты стремятся к бесконечности.

(Рао и Кендалл, 1950.)

4.13 Теорема Вейерштрасса устанавливает, что каждая функция, непрерывная на отрезке $[a, b]$, может быть представлена равномерно сходящимся рядом из полиномов $\sum_{n=0}^{\infty} P_n(x)$, где $P_n(x)$ — полином степени n . Вывести из

нее, что если две непрерывные плотности f_1 и f_2 , обращающиеся в нуль вне конечного отрезка $[a, b]$, имеют одни и те же моменты всех порядков, то

$$\int_a^b (f_1 - f_2)^2 dx = 0,$$

и, следовательно, моменты определяют распределение единственным образом, если оно сосредоточено на конечном интервале и имеет непрерывную плотность.

4.14 Если θ — неотрицательная функция от x и

$$\alpha(t) = \int_{-\infty}^{\infty} \theta^t(x) dF(x),$$

то плотность распределения для θ дается формулой

$$f(\theta) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \theta^{-t-1} \alpha(t) dt.$$

4.15 Показать, что если характеристическая функция $\varphi(t)$ обладает производными до второго порядка включительно, то

$$\left| \left(\frac{d\varphi}{dt} \right)_{t=0}^2 \right| \leq \left| \left(\frac{d^2\varphi}{dt^2} \right)_{t=0} \right|.$$

Обобщить этот результат,

4.16 Пусть λ — варианта с х. ф. $\varphi(t)$ и y — другая варианта такая, что при каждом фиксированном значении λ ее распределение задается распределением Пуассона из упражнения 3.1. Показать, что х. ф. y равна

$$\varphi\{(e^{it} - 1)/i\}.$$

4.17 Показать, что распределение

$$dF = \frac{dx_1 dx_2}{2\pi(1-\rho^2)^{1/2}} \exp\left\{\frac{-1}{2(1-\rho^2)}(x_1^2 - 2\rho x_1 x_2 + x_2^2)\right\},$$

$$-\infty \leq x_1, x_2 \leq \infty,$$

однозначно определяется своими моментами.

4.18 Если варианта распределена по нормальному закону

$$dF = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}x^2/\sigma^2}, \quad -\infty \leq x \leq \infty,$$

то х. ф. ее квадрата выражается формулой

$$\varphi(t) = (1 - 2it\sigma^2)^{-1}.$$

4.19 Если варианты x_1 и x_2 распределены так, как в упражнении 4.17, то совместная х. ф. величин x_1^2 и x_2^2 равна

$$\{(1 - 2it_1)(1 - 2it_2) + 4\rho^2 t_1 t_2\}^{-1}.$$

4.20 Некоторое распределение сосредоточено в целых точках отрезка $\left[-\frac{1}{2}n(n-1), \frac{1}{2}n(n-1)\right]$ и имеет производящую функцию

$$P(t) = \prod_{j=1}^n \frac{t^j - t^{-j}}{j(t - t^{-1})}.$$

Рассматривая логарифм производящей функции и используя следующее соотношение для чисел Бернулли:

$$|B_{2k}| = \frac{2(2k)!}{(2\pi)^{2k}} \sum_{j=1}^{\infty} \left(\frac{1}{j^{2k}}\right),$$

показать, что в нормированной форме это распределение стремится к нормальному.

4.21 Показать, что х. ф. «логистического» распределения

$$dF = \frac{dx}{4 \operatorname{ch}^2\left(\frac{1}{2}x\right)} = \frac{\exp(-x) dx}{\{1 + \exp(-x)\}^2}, \quad -\infty \leq x \leq \infty,$$

дается формулой

$$\varphi(t) = \Gamma(1 - it)\Gamma(1 + it)$$

и что его дисперсия равна $\pi^2/3$.

4.22 Пусть x — варианта, принимающая целочисленные значения. Показать, что в этом случае пределы интегрирования в формуле обращения (4.5) можно заменить на $-\pi$ и π . Использовать этот факт для обращения х. ф. пуассоновской варианты (см. пример 3.10).

СТАНДАРТНЫЕ РАСПРЕДЕЛЕНИЯ — (1)

5.1 Существует ряд распределений, которые занимают центральное положение в статистике как с теоретической, так и с практической точки зрения. В этой и следующей главах мы рассмотрим их свойства, оставляя иллюстрацию и обсуждение вопросов их статистических применений до последующих разделов книги. Мы кратко укажем, однако, каким образом они возникают, предвосхищая тем самым некоторые из идей, подробно развиваемых в дальнейшем. Это не повредит логической стройности изложения и придаст ему известную конкретность, без которой оно может стать несколько абстрактным.

Биномиальное распределение

5.2 Пусть имеется большая генеральная совокупность элементов, каждый из которых наделен либо качеством P , либо дополнительным качеством Q ($\neq P$); например, совокупность голубоглазых и неголубоглазых мужчин. Предположим, что элементы генеральной совокупности, обладающие качеством P , составляют p -ю часть, а обладающие качеством Q составляет q -ю часть, так что $p+q=1$. Если производится случайная выборка объема N , то мы ожидаем, что в среднем в ней будет Np элементов с качеством P и Nq — с качеством Q . Мы можем, стало быть, символически записать расположение элементов в соответствии с их качеством как

$$N(p+q).$$

Предположим теперь, что выбираются N пар элементов. Возможны следующие пары PP , PQ , QP и QQ . Из Np пар, в которых первый элемент обладает качеством P , в среднем p -я часть будет состоять из пар с P на втором месте и q -я часть — из пар, у которых на втором месте Q . Аналогичное заключение справедливо для Nq пар, первые элементы которых обладают качеством Q . Следовательно, символически пары могут быть расположены

следующим образом:

$$Np(p+q) + Nq(p+q) = N(p+q)^2.$$

Вообще, если извлекаются N множеств, по n элементов в каждом, то получится следующее расположение: $N(p+q)^n$, т. е. доля случаев, в которых встречается j раз P и $n-j$ раз Q , равна $\binom{n}{j} p^j q^{n-j}$ ($\binom{n}{j} p^j q^{n-j}$ — член в разложении $(p+q)^n$, содержащий $p^j q^{n-j}$). Мы приходим, таким образом, к рассмотрению биномиального распределения*), задаваемого разложением выражения

$$f = (p+q)^n. \quad (5.1)$$

Это распределение дискретно. Оно отвечает варианту, меняющейся от n до 0 и равной числу элементов со свойством P в множестве из n элементов. Часто более удобно считать, что значения варианты изменяются в возрастающем порядке от 0 до n . В этом случае распределение будет задаваться разложением

$$f = (q+p)^n. \quad (5.2)$$

5.3 В практике встречаются распределения, весьма близкие к биномиальному, особенно при таких искусственных экспериментах, как подбрасывание монеты или кости. Таблица 5.1 содержит данные, принадлежащие Уэлдону, который подбросил 12 костей 26 306 раз и зафиксировал результат каждого бросания. Этот эксперимент эквивалентен извлечению выборок, по 12 элементов в каждой, из большой генеральной совокупности. Выпадение пятерки или шестерки на любой кости рассматривалось как наличие качества P , или, иначе выражаясь, как «успех».

Если бы кости были идеальны (что в действительности бывает редко), то доля p успехов была бы равна $\frac{1}{3}$ и соответствующее биномиальное выражение в форме (5.2) имело бы вид $(\frac{2}{3} + \frac{1}{3})^{12}$. В рассматриваемом нами эксперименте кости не были вполне безупречными, поскольку доля выпадений 5 и 6 составила 0,3377. Беря это число в качестве p , мы можем символически записать распределение относительных частот как $(0,6623 + 0,3377)^{12}$. Умножив это выражение на полную частоту 26 306, по формуле бинома Ньютона найдем теоретические частоты. Они представлены в третьем столбце таблицы 5.1. Из таблицы видно, что соответствие с наблюдаемыми частотами весьма неплохое.

*) Иной подход к биномиальному распределению см., например, в книге: Б. В. Гнеденко, Курс теории вероятностей, 1954. (Прим. перев.)

Таблица 5.1

Частотное распределение результатов 26 306 бросаний 12 костей; успехом считается выпадение 5 или 6

Число успехов	Наблюденная частота	Теоретическая частота, отвечающая биномиальному распределению с $p=0,3377$	Число успехов	Наблюденная частота	Теоретическая частота, отвечающая биномиальному распределению с $p=0,3377$
0	185	187	6	3 067	3 043
1	1 149	1 146	7	1 331	1 330
2	3 265	3 215	8	403	424
3	5 475	5 465	9	105	96
4	6 114	6 269	10 и более	18	16
5	5 194	5 115	Полная частота	26 306	26 306

5.4 Мы уже вычисляли моменты и факториальные моменты биномиального распределения в примерах 3.2 и 3.8. Характеристическая функция этого распределения выражается формулой

$$\varphi(t) = \sum_{j=0}^n \binom{n}{j} q^{n-j} p^j e^{ijt} = (q + pe^{it})^n. \quad (5.3)$$

Логарифмируя и раскладывая экспоненту в ряд, получим следующее выражение для производящей функции семиинвариантов ($\theta = it$):

$$\psi(t) = n \log \left(1 + p\theta + \frac{1}{2} p\theta^2 + \frac{1}{6} p\theta^3 + \frac{1}{24} p\theta^4 + \dots \right).$$

Разлагая логарифм и объединяя члены с одинаковыми степенями θ , находим

$$\begin{aligned} \kappa_1 = \mu'_1 &= np, & \kappa_2 = \mu_2 &= np(1-p) = npq, \\ \kappa_3 = \mu_3 &= np(1-p)(1-2p) = npq(q-p), \\ \kappa_4 &= np(1-p)(1-6p+6p^2) = npq(1-6pq) \text{ и т. д.,} \end{aligned} \quad (5.4)$$

откуда

$$\mu_4 = 3n^2 p^2 q^2 + npq(1-6pq), \quad (5.5)$$

$$\gamma_1 = \mu_3 / \mu_2^{3/2} = \frac{q-p}{(npq)^{1/2}} = \sqrt{\beta_1}, \quad (5.6)$$

$$\gamma_2 = \kappa_4 / \kappa_2^2 = \frac{1-6pq}{npq} = \beta_2 - 3. \quad (5.7)$$

5.5 Дальнейшие формулы требуются редко и, если в них есть необходимость, могут быть получены из интересных рекуррентных соотношений, связывающих моменты биномиального распределения.

Характеристическая функция централизованного биномиального распределения с началом отсчета, перенесенным в среднее значение, имеет вид

$$\varphi(t) = e^{-np\theta}(q + pe^\theta)^n, \quad \theta = it. \quad (5.8)$$

Дифференцируя по θ , имеем

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{\mu_j \theta^{j-1}}{(j-1)!} &= -np e^{-np\theta} (q + pe^\theta)^n + n e^{-np\theta} (q + pe^\theta)^{n-1} p e^\theta = \\ &= -np \sum_{j=0}^{\infty} \frac{\mu_j \theta^j}{j!} + \frac{np e^\theta}{q + p e^\theta} \sum_{j=0}^{\infty} \frac{\mu_j \theta^j}{j!}, \end{aligned}$$

откуда, после несложного преобразования, получаем

$$(q + p e^\theta) \left\{ \sum_{j=1}^{\infty} \frac{\mu_j \theta^{j-1}}{(j-1)!} \right\} - npq (e^\theta - 1) \sum_{j=0}^{\infty} \frac{\mu_j \theta^j}{j!} = 0.$$

Приравнявая коэффициенты при θ^{r-1} , находим следующее выражение для момента r -го порядка относительно среднего через такие же моменты более низкого порядка:

$$\mu_r = npq \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_j - p \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_{j+1}. \quad (5.9)$$

Далее, согласно определению

$$\mu_r = \sum_{j=0}^n (j - np)^r \binom{n}{j} (1-p)^{n-j} p^j.$$

Дифференцируя по p , имеем

$$\begin{aligned} \frac{d\mu_r}{dp} &= -rn \sum (j - np)^{r-1} \binom{n}{j} (1-p)^{n-j} p^j - \\ &\quad - \sum (j - np)^r \binom{n}{j} q^{n-j-1} (n-j) p^j + \\ &\quad + \sum (j - np)^r \binom{n}{j} q^{n-j} p^{j-1}. \end{aligned}$$

Первый член правой части равен $-rn\mu_{r-1}$. Сумма остальных двух, как легко убедиться, равна

$$\frac{1}{pq} \sum (j - np)^{r+1} \binom{n}{j} q^{n-j} p^j = \frac{1}{pq} \mu_{r+1}.$$

Следовательно,

$$\mu_{r+1} = pq \left(nr\mu_{r-1} + \frac{d\mu_r}{dp} \right). \quad (5.10)$$

Например, $\mu_1=0$, $\mu_2=npq=np(1-p)$ и поэтому

$$\mu_3=pq(n-2np)=npq(q-p),$$

что совпадает с результатом (5.4).

Формулы для факториальных моментов относительно нуля принимают особенно простую форму. Нами уже было найдено (пример 3.8), что

$$\mu'_{[r]} = p^r n^{[r]}. \quad (5.11)$$

5.6 Если $p=q$, то биномиальное распределение, очевидно, симметрично. При $p \neq q$ оно несимметрично. Но в обоих случаях оно унимодально, если только pn не мало. Действительно, превышение частоты значения $r+1$ над частотой значения r означает, что

$$\binom{n}{r+1} q^{n-r-1} p^{r+1} > \binom{n}{r} q^{n-r} p^r,$$

иначе

$$\frac{n!}{(n-r)! r!} \frac{(n-r-1)! (r+1)!}{n!} < \frac{p}{q},$$

или

$$\frac{r+1}{n-r} < \frac{p}{q},$$

т. е.

$$\frac{r+1}{n+1} < p. \quad (5.12)$$

Следовательно, частота возрастает до тех пор, пока не станет $(r+1) > p(n+1)$, а затем она убывает. Несколько типичных биномиальных распределений представлены в таблице 5.2.

5.7 Частоты биномиального распределения можно вычислить прямо по формулам $\binom{n}{j} q^{n-j} p^j$; при малых значениях n вычисление ведется непосредственно, при больших значениях n удобно воспользоваться таблицей для $\log n!$. Прямое вычисление функции распределения суммированием частот длинно и утомительно; для их отыскания проще использовать приводимую ниже формулу и таблицы неполной В-функции. В общей формуле Тейлора с остаточным членом в интегральной форме

$$f(a+h) = \sum_{j=0}^{r-1} \frac{h^j}{j!} f^{(j)}(a) + \int_0^1 \frac{h^r (1-t)^{r-1}}{(r-1)!} f^{(r)}(a+th) dt \quad (5.13)$$

Таблица 5.2

Частоты биномиального распределения для $n = 20$ и $p = 0,1$ (0,1)
0,5 (все частоты в таблице увеличены в 10 000 раз)

Число успе- хов	$p=0,1$ $q=0,9$	$p=0,2$ $q=0,8$	$p=0,3$ $q=0,7$	$p=0,4$ $q=0,6$	$p=0,5$ $q=0,5$
0	1 216	115	8	—	—
1	2 702	576	68	5	—
2	2 852	1 369	278	31	2
3	1 901	2 054	716	123	11
4	898	2 182	1 304	350	46
5	319	1 746	1 789	746	148
6	89	1 091	1 916	1 244	370
7	20	545	1 643	1 659	739
8	4	222	1 144	1 797	1 201
9	1	74	654	1 597	1 602
10	—	20	308	1 171	1 762
11	—	5	120	710	1 602
12	—	1	39	355	1 201
13	—	—	10	146	739
14	—	—	2	49	370
15	—	—	—	13	148
16	—	—	—	3	46
17	—	—	—	—	11
18	—	—	—	—	2
19	—	—	—	—	—
20	—	—	—	—	—

положим $a=q$, $h=p$, $f(a+h) = (q+p)^n$. Тогда получим

$$(q+p)^n = \sum_{j=0}^{r-1} \binom{n}{j} q^{n-j} p^j + R_r,$$

где

$$R_r = \int_0^1 \frac{p^r (1-t)^{r-1}}{(r-1)!} \frac{n!}{(n-r)!} (q+pt)^{n-r} dt. \quad (5.14)$$

Подстановкой $t=1-x/p$ этот интеграл сводится к

$$\begin{aligned} R_r &= \frac{n!}{(r-1)!(n-r)!} \int_0^p x^{r-1} (1-x)^{n-r} dx = \\ &= \frac{\Gamma(n+1)}{\Gamma(r)\Gamma(n-r+1)} B_p(r, n-r+1) = \\ &= \frac{B_p(r, n-r+1)}{B(r, n-r+1)} = I_p(r, n-r+1). \end{aligned} \quad (5.15)$$

Согласно хорошо известному свойству В-функции сумма первых r членов равна, следовательно,

$$(q + p)^n - R_r = 1 - I_p(r, n - r + 1) = I_q(n - r + 1, r). \quad (5.16)$$

Аналогично, сумма первых $r+1$ членов есть $I_q(n - r, r + 1)$, так что для $(r+1)$ -го члена имеет место формула

$$\binom{n}{r} q^{n-r} p^r = I_q(n - r, r + 1) - I_q(n - r + 1, r). \quad (5.17)$$

Пример 5.1

При $n=20$, $r=11$, $p=0,4$ по таблицам найдено, что $R_{11} = I_{0,4}(11, 10) = 0,1275212$. Значение, получающееся суммированием последних шести чисел соответствующего столбца в таблице 5.2, равно 0,1276. Расхождение в последней цифре происходит из-за округления. Согласно таблицам $R_{12} = I_{0,4}(12, 9) = 0,0565264$. Отсюда получаем частоту значения 11 (11 «успехов») $R_{12} - R_{11} = 0,0710$, что совпадает с соответствующей частотой, указанной в таблице 5.2.

Таблицы биномиального распределения

Имеется ряд подробных таблиц биномиального распределения.

(а) *Biometrika Tables*. Приводятся значения индивидуальных частот биномиального распределения с точностью до пятого десятичного знака*) для $p=0,01, 0,02(0,02)0,1(0,01)0,5$ и $n=5(5)30$.

(б) *Tables of the Binomial Probability Distribution* (National Bureau of Standards Applied Mathematics Series 6, Washington, 1950). Даны значения индивидуальных частот и функции распределения с точностью до семи десятичных знаков для $p=0,01(0,01)0,50$ и $n=2(1)49$.

(в) H. G. Romig. 50—100 *Binomial Tables* (Wiley, New York, 1953). Эти таблицы дополняют таблицы (б) до значений $p=50(5)100$ с точностью до шести знаков.

(г) *Tables of the Cumulative Binomial Probabilities* (U. S. Army Ordnance Corps Pamphlet ORD P20—1, Washington, 1952). Даны значения функции распределения для $p=0,01(0,01)0,50$ и $n=1(1)150$ с точностью до семи знаков.

(д) *Tables of the Cumulative Binomial Probability Distribution* (Harvard University Press, 1955). Даны значения функции

*) Эта таблица помещена также в сборнике *Таблицы математической статистики*, стр. 104 и 346. (Прим. ред.)

распределения для $p=0,01$ (0,01) 0,50, $p=\frac{1}{16}\left(\frac{1}{16}\right)\frac{1}{2}$, $p=\frac{1}{12}\left(\frac{1}{12}\right)\frac{1}{2}$ и $n=1$ (1) 50 (2) 100 (10) 200 (20) 500 (50) 1000.

(е) В таблицах Миллера (1954) приводятся значения $\binom{n}{r}$ для $2 \leq r \leq \frac{1}{2}n \leq 100$; $r=2$ (1) 12, $n \leq 500$; $r=2$ (1) 11, $n=500$ (1) 1000; $r=2$ (1) 5, $n=1000$ (1) 2000; $r=2, 3$, $n=2000$ (1) 5000.

В примере 4.6 мы видели, что биномиальное распределение при возрастании n стремится к «нормальному» распределению. Раффом (1956) показано, что если $np^{3/2} > 1,07$, то ошибка при использовании нормальной функции распределения вместо биномиальной не превосходит 0,05 для всех r .

Пуассоновское распределение

5.8 Встречаются иногда случаи, когда доля p «успехов» в генеральной совокупности очень мала. Можно предположить число n достаточно большим так, чтобы величина np не была малой при малых p . Этот путь приводит нас к рассмотрению предельной формы биномиального распределения при $p \rightarrow 0$, когда np остается конечным и равным, скажем, λ .

При этих условиях

$$\begin{aligned} \binom{n}{r} p^r q^{n-r} &= \frac{n!}{(n-r)! r!} \frac{\lambda^r}{n^r} \left(1 - \frac{\lambda}{n}\right)^{n-r} \sim \\ &\sim \frac{\sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}}{\sqrt{2\pi} (n-r)^{n-r+\frac{1}{2}} e^{-n+r} n^r} \frac{\lambda^r}{r!} e^{-\lambda} \sim \\ &\sim \frac{1}{\left(1 - \frac{r}{n}\right)^n e^r} \frac{\lambda^r}{r!} e^{-\lambda} \sim \frac{\lambda^r}{r!} e^{-\lambda}. \end{aligned}$$

Следовательно, частоты рассматриваемых биномиальных распределений в пределе имеют вид

$$e^{-\lambda} \left(1, \frac{\lambda}{1!}, \dots, \frac{\lambda^j}{j!}, \dots\right). \quad (5.18)$$

Распределение, для которого (5.18) представляет собой последовательные частоты в точках $0, 1, 2, \dots$, называется пуассоновским. Впервые оно было рассмотрено Пуассоном в 1837 году.

В терминах характеристических функций имеем (полагая $\theta = it$)

$$\begin{aligned} \varphi(t) &= \lim_{n \rightarrow \infty} (q + pe^{\theta})^n = \\ &= \lim_{n \rightarrow \infty} \left\{1 + \frac{\lambda}{n} (e^{\theta} - 1)\right\}^n = \exp \{\lambda (e^{\theta} - 1)\}. \end{aligned} \quad (5.19)$$

Выше, в примере 3.10, уже было доказано, что (5.19) есть х. ф. распределения (5.18).

Соответствующая производящая функция семиинвариантов равна, стало быть,

$$\psi(t) = \lambda(e^t - 1) = \lambda \sum \theta^j / j!$$

Следовательно, все семиинварианты пуассоновского распределения равны λ . Отсюда

$$\left. \begin{aligned} \mu'_1 &= \lambda, \\ \mu_2 &= \lambda, \\ \mu_3 &= \lambda, \\ \mu_4 &= \lambda + 3\lambda^2. \end{aligned} \right\} \quad (5.20)$$

Если в (5.9) и (5.10) перейти к пределу при $n \rightarrow \infty$, то получим

$$\mu_r = \lambda \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_j \quad (5.21)$$

и

$$\mu_{r+1} = r\lambda\mu_{r-1} + \frac{\lambda d\mu_r}{d\lambda}. \quad (5.22)$$

5.9 Сравнение последовательных членов $\lambda^r/r!$ и $\lambda^{r+1}/(r+1)!$ показывает, что частоты возрастают до тех пор, пока $r+1 < \lambda$, а затем убывают к нулю. При малых λ полигоны частот очень асимметричны, и при $\lambda < 1$ они J-образны; при возрастании λ они становятся почти унимодальными и симметричными.

Для нахождения суммы первых r частот пуассоновского распределения можно воспользоваться способом, подобным примененному в 5.7. Согласно (5.13) r -й остаточный член равен

$$R_r = \frac{\lambda^r}{\Gamma(r)} \int_0^1 e^{-\lambda(1-t)} (1-t)^{r-1} dt = \frac{\Gamma_\lambda(r)}{\Gamma(r)} = I\left(\frac{\lambda}{\sqrt{r}}, r-1\right). \quad (5.23)$$

Мы применили здесь обозначение из пирсоновских таблиц неполной Г-функции. С аргументом, использованным в этих таблицах, работать в данном случае трудно*), и хотя формулу

*) Существуют таблицы (В. И. Пагурова, *Таблицы неполной гамма-функции*, М., Изд-во Вычислит. центра АН СССР, 1963), в которых указаны значения $R_r = I(\lambda, r)$ (с семью десятичными знаками) непосредственно для аргументов λ и r . Эту функцию не следует смешивать с функцией, введенной К. Пирсоном, так как

пагуровская $I(\lambda, r)$ = пирсоновская $I\left(\frac{\lambda}{\sqrt{r}}, r-1\right)$. (Прим. ред.)

(5.23) можно применять для суммирования частот пуассоновского распределения, вычислять $e^{-\lambda} \lambda^r / r!$ легче непосредственно, чем с помощью выражения, аналогичного (5.17):

$$I = \left(\frac{\lambda}{\sqrt{r}}, r-1 \right) - I \left(\frac{\lambda}{\sqrt{r+1}}, r \right).$$

В *Biometrika Tables for Statisticians* приводятся значения *) членов $e^{-\lambda} \lambda^r / r!$ (с точностью до шести знаков) для $\lambda = 0,1(0,1)15$, $r = 0(1)k$, где k — то наибольшее значение r , для которого функция $e^{-\lambda} \lambda^r / r!$ при заданном λ с точностью до шести знаков не равна нулю. Суммы членов можно находить по таблицам χ^2 -распределения в *Biometrika Tables for Statisticians*, которые позво-

ляют эффективно получать $\sum_{j=0}^{r-1} e^{-\lambda} \lambda^j / j!$ с точностью до пяти знаков для $\lambda = 0,0005(0,0005)0,005(0,005)0,05(0,05)1,0(0,1)5,0(0,25)10,0(0,5)20(1)60$. Эти таблицы сумм могут быть также использованы как дополнение к таблицам отдельных членов, особенно при $0 \leq \lambda \leq 1$. Молина (1942) опубликовал таблицы, в которых даны значения отдельных членов и сумм с точностью до шести-семи знаков для $\lambda = 0,001(0,001)0,010(0,010)0,30(0,10)15(1)100$.

5.10 Рассмотрим теперь одно обобщение биномиального и пуассоновского распределений. В 5.2 наш подход был основан на извлечении множеств по n элементов из одной и той же генеральной совокупности. Предположим теперь, что каждое множество объема n состоит из элементов, извлеченных из n различных генеральных совокупностей, которым соответствуют доли $p_1, p_2, \dots, p_n$. В этом случае частоты будут получаться из символической формулы

$$(p_1 + q_1)(p_2 + q_2) \dots (p_n + q_n) = \prod_{j=1}^n (p_j + q_j). \quad (5.24)$$

Если все p одинаковы, то мы будем иметь, конечно, биномиальное распределение.

Характеристическая функция этого распределения выражается формулой

$$\varphi(t) = \prod (q_j + p_j e^{it}).$$

*) Эта таблица помещена в сборнике *Таблицы математической статистики*, стр. 110 и 360. Там же (стр. 21 и 202) даны более подробные, чем в *Biometrika Tables*, пятизначные таблицы χ^2 -распределения, позволяющие находить значения функции пуассоновского распределения. (Прим. ред.)

Следовательно,

$$\begin{aligned}\Psi(t) &= \sum \log(q_j + p_j e^{\theta}) = \\ &= \sum \log\left(1 + p_j \theta + \frac{1}{2} p_j \theta^2 + \dots\right) = \\ &= \theta \sum p_j + \frac{1}{2} \theta^2 \sum (p_j - p_j^2) + \dots\end{aligned}$$

и

$$\left. \begin{aligned}\mu'_1 &= \sum p_j, \\ \mu_2 &= \mu_2 = \sum p_j q_j.\end{aligned} \right\} \quad (5.25)$$

Обозначая буквой $\bar{p}$ арифметическое среднее всех p_j , имеем

$$\begin{aligned}\mu'_1 &= n\bar{p}, \\ \mu_2 &= \sum pq = \sum p - \sum p^2 = \\ &= \sum p - \frac{1}{n} \left(\sum p\right)^2 - \left\{ \sum p^2 - \frac{1}{n} \left(\sum p\right)^2 \right\} = \\ &= n\bar{p} - n\bar{p}^2 - n\sigma_p^2 = n\bar{p}q - n\sigma_p^2,\end{aligned} \quad (5.26)$$

где σ_p^2 — дисперсия p .

Сравнение этих результатов с соответствующими результатами для биномиального распределения показывает, что дисперсия распределения, задаваемого (5.24), *меньше*, чем дисперсия биномиального распределения с тем же средним $\bar{p}$, на n раз взятую дисперсию «варианты» p .

Подобным образом для пуассоновского распределения в аналогичном случае имеем

$$\mu'_1 = \bar{\lambda} \quad (5.27)$$

и

$$\mu_2 = \bar{\lambda} - \sigma_{\lambda}^2/n = \bar{\lambda} + O(n^{-1}),$$

где $\bar{\lambda}$ — среднее значение λ_j .

Дисперсия распределения, задаваемого (5.24), приблизительно, стало быть, равна дисперсии пуассоновского распределения, несмотря на различие p в разных генеральных совокупностях, лишь бы дисперсия «варианты» λ была мала по сравнению с n ; так будет, если все p малы.

5.11 Рассмотрим теперь случай, когда множества, содержащие по n элементов, извлекаются из k различных генеральных совокупностей, доли «успехов» в которых $p_1, \dots, p_k$. (В предыдущем случае каждое множество объема n состояло из элементов,

извлеченных из n различных генеральных совокупностей. Теперь мы предполагаем, что элементы каждого множества извлекаются только из одной генеральной совокупности, но разные множества — из разных генеральных совокупностей.) Частоты в этом случае определяются из символической формулы

$$\frac{1}{k} \sum_{j=1}^k (q_j + p_j)^n. \quad (5.28)$$

Моменты этого распределения представляют собой, очевидно, средние значения моментов составляющих распределений $(p_j + q_j)^n$. В частности, из (5.4) получаем

$$\begin{aligned} \mu'_1 &= \frac{1}{k} \sum n p, \\ \mu'_2 &= \frac{1}{k} \left\{ \sum n p + \sum n(n-1) p^2 \right\}. \end{aligned}$$

Обозначая, как и выше, буквой $\bar{p}$ среднее различных значений p , имеем

$$\begin{aligned} \mu'_1 &= n \bar{p}, \\ \mu_2 &= n \bar{p} + \frac{1}{k} \sum n(n-1) p^2 - n^2 \bar{p}^2 = \\ &= n \bar{p} \bar{q} + \frac{1}{k} \sum n(n-1) p^2 - n(n-1) \bar{p}^2 = \\ &= n \bar{p} \bar{q} + n(n-1) \left(\frac{1}{k} \sum p^2 - \bar{p}^2 \right) = n \bar{p} \bar{q} + n(n-1) \sigma_p^2. \end{aligned} \quad (5.29)$$

Дисперсия здесь больше дисперсии биномиального распределения со средним $\bar{p}$ на величину, равную $n(n-1) \sigma_p^2$, где σ_p^2 — дисперсия «варианты» p .

Переходя к пределу, для пуассоновского распределения имеем

$$\left. \begin{aligned} \mu'_1 &= \bar{\lambda}, \\ \mu_2 &= \bar{\lambda} + \sigma_{\lambda}^2. \end{aligned} \right\} \quad (5.30)$$

И здесь тоже дисперсия распределения отлична от дисперсии пуассоновского распределения со средним $\bar{\lambda}$.

5.12 Мы рассмотрим подобные вопросы более детально, когда будем заниматься теорией выборочного метода, однако несколько замечаний могут оказаться полезными и теперь. В практике часто бывает так, что данные, полученные по выборкам из генеральных совокупностей, охватывающих расширенную область или продолжительный отрезок времени, не соответствуют прос-

тому биномиальному типу. Например, предположим, что голубые глаза у индивидуума рассматриваются как успех, и возьмем несколько выборок по n элементов в каждой из разных областей Соединенного Королевства. Весьма вероятно, что доли успехов в этих выборках не отвечают одному и тому же биномиальному распределению. Дисперсия, вычисленная по известному n и среднему наблюдаемому p , быть может, окажется меньше дисперсии, наблюдаемой среди различных выборок. В таком случае в силу (5.29) следует сделать вывод, что доля успеха p меняется от области к области, и поэтому дисперсия p положительна. При этом мы предполагаем, что колебания выборочных значений p не могут быть объяснены на основе лишь одних выборочных флуктуаций. Более подробно эти вопросы будут обсуждаться в последующих главах.

5.13 Тот же эффект обнаруживается в случае распределений, которые на первый взгляд, казалось бы, должны быть пуассоновскими. Например, самоубийство — редкое событие, и можно было бы ожидать, что в последовательности больших выборок, скажем, среди населения Соединенного Королевства в последовательные годы, частоты самоубийств должны подчиняться пуассоновскому распределению. Однако это не обязательно так, поскольку различные члены совокупности в различной степени подвержены возможности самоубийства и склонность к самоубийству может меняться от года к году; например, она больше в годы кризисов. Такое непостоянство степени риска характерно и для несчастных случаев на производстве, при изучении которых без достаточного основания ранее использовали пуассоновское распределение. Второй столбец таблицы 5.3 показывает частоту несчастных случаев среди женщин, работавших на

Таблица 5.3

Несчастные случаи, происшедшие с 647 женщинами, работавшими на производстве снарядов, в течение пяти недель (Гринвуд и Юл, 1920)

Число несчастных случаев	Наблюдаемая частота	Пуассоновское распределение с тем же средним	Распределение (5.33)
0	447	406	442
1	132	189	140
2	42	45	45
3	21	7	14
4	3	1	5
5 и больше	2	0,1	2
Полная частота	647	648	648

производстве снарядов. Пуассоновское распределение, представленное в третьем столбце, дает весьма посредственное приближение. Одна из возможных причин этого состоит в неодинаковой степени подверженности индивидуумов несчастным случаям.

{ В качестве рабочей гипотезы будем считать (Гринвуд и Юл, 1920), что генеральная совокупность состоит из элементов, в различной степени подверженных несчастным случаям. Это различие характеризуется различными значениями параметра λ пуассоновского распределения; предположим еще, что распределение параметра λ дается формулой *Гамма - распределение* ;

$$dF = \frac{c^p}{\Gamma(p)} e^{-c\lambda} \lambda^{p-1} d\lambda, \quad 0 \leq \lambda \leq \infty, \quad c > 0. \quad (5.31)$$

Частота осуществления j успехов *) равна при этом

$$\int_0^{\infty} \frac{c^p}{\Gamma(p)} e^{-c\lambda} \lambda^{p-1} e^{-\lambda} \frac{\lambda^j}{j!} d\lambda$$

и представляет собой коэффициент при t^j в разложении интеграла

$$\frac{c^p}{\Gamma(p)} \int_0^{\infty} e^{-c\lambda} \lambda^{p-1} e^{-\lambda + \lambda t} d\lambda,$$

который вычисляется подстановкой $(c+1-t)\lambda = u$ и принимает вид

$$\frac{c^p}{(c+1-t)^p} = \left(\frac{c}{c+1}\right)^p \left(1 - \frac{t}{c+1}\right)^{-p}. \quad (5.32)$$

Частоты 0, 1, 2, ... успехов равны, следовательно,

$$\left(\frac{c}{c+1}\right)^p \left\{ 1, \frac{p}{c+1}, \frac{p}{2!} \frac{(p+1)}{(c+1)^2}, \dots \right\}. \quad (5.33)$$

Легко подсчитать среднее этого распределения

$$\begin{aligned} \mu'_1 &= \left(\frac{c}{c+1}\right)^p \left\{ \frac{p}{c+1} + \frac{p(p+1)}{(c+1)^2} + \dots \right\} = \\ &= \left(\frac{c}{c+1}\right)^p \frac{p}{c+1} \left(\frac{c+1}{c}\right)^{p+1} = \frac{p}{c}. \end{aligned} \quad (5.34)$$

Аналогично

$$\mu'_2 = \frac{p}{c} \left(\frac{p+c+1}{c} \right),$$

*) Рассматривая данные таблицы 5.3, было бы более уместно слово «успех» заменить словом «неудача». (Прим. ред.)

так что

$$\mu_2 = \frac{p(c+1)}{c^2} = \frac{p}{c} + \frac{p}{c^2}. \quad (5.35)$$

Если левые части (5.34) и (5.35) заменить соответственно средним и дисперсией, оцененными по данным таблицы 5.3, то получатся два уравнения, которые однозначно определяют p и c , а значит, и распределение (5.33). Частоты этого распределения приведены в четвертом столбце таблицы 5.3. Очевидно, что они значительно лучше согласуются с результатами наблюдений, чем частоты пуассоновского распределения.

Отрицательное биномиальное распределение

5.14 Распределение, заданное формулой (5.33), обладает тем интересным свойством, что оно порождается биномом с *отрицательным индексом*. В подходе, рассмотренном в 5.2, индекс необходимо положителен. Нередко, однако, встречаются случаи, когда данные соответствуют отрицательному индексу. Мы укажем один весьма простой процесс, приводящий к такому распределению, когда будем рассматривать теорию последовательного выбора. Другой классический пример, принадлежащий Юлу (1910), исторически является первым примером, обратившим внимание статистиков на этот тип распределений. Он не содержит в себе никаких произвольных допущений относительно характера распределения, таких, например, как условие (5.31).

Пусть имеется генеральная совокупность индивидуумов, подверженная последовательным вспышкам болезни, и предположим, что каждая вспышка поражает (случайным образом) p -ю долю генеральной совокупности. После n вспышек доли индивидуумов, болевших 0, 1, ... раз, будут выражаться членами разложения бинома $(q+p)^n$. Если после r заболеваний индивидуум погибает, то доля индивидуумов, переживших n вспышек, равна сумме первых r членов этого разложения. Доля индивидуумов, умерших при n -й вспышке, образуется теми индивидуумами, которые болели $r-1$ раз в течение первых $n-1$ вспышек и заболели снова во время n -й вспышки. Она равна

$$\binom{n-1}{r-1} q^{n-r} p^{r-1} \times p = \binom{n-1}{r-1} q^{n-r} p^r.$$

Таким образом, поскольку смерть не наступает до r -го заболевания, доли смертельных исходов при r -й, $(r+1)$ -й, ... вспышках суть

$$p^r \left\{ 1, r q \frac{r(r+1)}{2!} q^2, \dots \right\};$$

инными словами, они задаются последовательными членами разложения бинома с отрицательным индексом $p^r(1-q)^{-r}$. Было обнаружено, что этому закону подчиняются данные экспериментов по уничтожению бактерий дезинфекцией.

5.15 Х. ф. отрицательного биномиального распределения, заданного разложением $p^n(1-q)^{-n}$, имеет вид (0, как обычно, обозначает it)

$$\varphi(t) = p^n(1 - qe^{\theta})^{-n}. \quad (5.36)$$

Следовательно, производящая функция семиинвариантов выражается формулой

$$\psi(t) = n \log p - n \log(1 - qe^{\theta}) = -n \log \left\{ 1 - \frac{q}{p}(e^{\theta} - 1) \right\}. \quad (5.37)$$

Разлагая $\psi(t)$ в ряд и приводя подобные члены, находим

$$\left. \begin{aligned} \kappa_1 &= \frac{nq}{p}, & \kappa_2 &= \frac{nq}{p^2}, & \kappa_3 &= \frac{nq(1+q)}{p^3}, \\ \kappa_4 &= \frac{nq(1+4q+q^2)}{p^4}. \end{aligned} \right\} \quad (5.38)$$

Эти выражения могут быть получены из равенств (5.4) заменой p на q/p и q на $1/p$.

П. ф. ф. м. имеет вид

$$\omega(t) = p^n \{1 - q(1+t)\}^{-n} = (1 - qt/p)^{-n}. \quad (5.39)$$

Отсюда находим факториальные моменты относительно нуля

$$\mu'_{[r]} = \left(\frac{q}{p}\right)^r (n+r-1)!^r. \quad (5.40)$$

Следует отметить, что для отрицательного биномиального распределения $\mu_2 > \mu'_1$, в то время как для обычного биномиального распределения $\mu_2 < \mu'_1$, а для пуассоновского распределения $\mu_2 = \mu'_1$. Предварительный подсчет величины μ_2/μ'_1 часто оказывается полезным при определении, какому из этих трех типов лучше всего соответствуют имеющиеся данные.

Логарифмическое распределение

5.16 Существует одно интересное распределение, являющееся предельным для отрицательного биномиального распределения. Пусть q — вероятность наличия некоторого признака и p — вероятность его отсутствия в отдельном испытании. При некоторых обстоятельствах постоянное отсутствие признака во всех испытаниях может оказаться невозможным событием. Мы приходим тогда к рассмотрению отрицательного биномиального распреде-

ления с отсутствующим нулевым классом, т. е. к распределению, частоты которого пропорциональны величинам

$$p^n \left\{ nq, \frac{n(n+1)}{2!} q^2, \dots \right\}. \quad (5.41)$$

Поскольку полная частота равна $1 - p^n$, то получаем следующее распределение:

$$\frac{np^n}{1-p^n} \left\{ q, \frac{n+1}{2!} q^2, \frac{(n+1)(n+2)}{3!} q^3, \dots \right\}. \quad (5.42)$$

Пусть теперь n стремится к нулю. Применяя к стоящему перед скобками множителю правило Лопиталья, имеем

$$\lim \frac{n}{p^n - 1} = \lim \frac{n}{\exp(-n \log p) - 1} = \frac{1}{-\log p}.$$

Следовательно, предельное распределение имеет вид

$$\frac{-1}{\log(1-q)} \left\{ q, \frac{q^2}{2}, \frac{q^3}{3}, \dots \right\}. \quad (5.43)$$

Производящая функция его равна

$$P(t) = \frac{1}{\log(1-qt)} \log(1-qt). \quad (5.44)$$

Это распределение было введено Р. А. Фишером и др. (1943). Путь, который нас к нему привел, представляется несколько искусственным, поскольку отрицательное биномиальное распределение было определено для целых индексов n , а здесь мы устремляем n к нулю. Основания для этого процесса станут более ясными, когда будет обсуждаться выборочная теория. Действительно, сумма N независимых наблюдений, подчиняющихся одному и тому же отрицательному биномиальному распределению с индексом n , сама подчиняется этому же распределению, но с индексом Nn (этот факт мы приводим здесь без доказательства). Порождающую логарифмическое распределение схему можно рассматривать как переход к пределу, когда N стремится к бесконечности, а Nn — к нулю. Пример распределения логарифмического типа дается в таблице 5.4. Его следует рассматривать как иллюстрацию хорошего согласия опытных и теоретических данных, а не как совокупность данных, полученных посредством только что указанной схемы.

5.17 Заменяя в (5.44) t на e^θ , находим, что характеристическая функция логарифмического распределения выражается формулой

$$\varphi(t) = 1 + \alpha \log \left\{ 1 - \frac{q}{p} \theta - \frac{q}{p} \frac{\theta^2}{2!} - \frac{q}{p} \frac{\theta^3}{3!} - \frac{q}{p} \frac{\theta^4}{4!} - \dots \right\},$$

Таблица 5.4

Распределение бабочек в Малайе (теоретические частоты вычислены по логарифмическому распределению).

Данные Фишера, Корбе и Уилльямса (Fisher, Corbet, Williams, Jour. An. Ecology, 1943, 12, 42)

Номер вида	Теоретическая частота	Наблюденная частота	Номер вида	Теоретическая частота	Наблюденная частота
1	135,05	118	13	10,02	6
2	67,33	74	14	9,28	12
3	44,75	44	15	8,63	6
4	33,46	24	16	8,07	9
5	26,69	29	17	7,57	9
6	22,17	22	18	7,13	6
7	18,95	20	19	6,74	10
8	16,53	19	20	6,38	10
9	14,65	20	21	6,06	11
10	13,14	15	22	5,77	5
11	11,91	12	23	5,50	3
12	10,89	14	24	5,25	3

где $\alpha = 1/\log(1 - q)$. Пользуясь аналогией между этим выражением и (5.37), получаем следующие формулы для моментов (но не симинвариантов) относительно нуля:

$$\begin{aligned}\mu'_1 &= -\frac{\alpha q}{p}, \quad \mu'_2 = -\frac{\alpha q}{p^2}, \quad \mu'_3 = -\frac{\alpha q(1+q)}{p^3}, \\ \mu'_4 &= -\frac{\alpha q(1+4q+q^2)}{p^4}.\end{aligned}\quad (5.45)$$

Отсюда

$$\left. \begin{aligned}\mu_2 &= -\frac{\alpha q(1+\alpha q)}{p^2}, \\ \mu_3 &= -\frac{\alpha q(1+q+3\alpha q+2\alpha^2 q^2)}{p^3}, \\ \mu_4 &= -\frac{\alpha q\{1+4q+q^2+4\alpha q(1+q)+6\alpha^2 q^2+3\alpha^3 q^3\}}{p^4}.\end{aligned}\right\} \quad (5.46)$$

Гипергеометрическое распределение

5.18 Рассмотрим теперь следующее обобщение схемы, изложенной в 5.2. Пусть из генеральной совокупности, состоящей из N элементов, извлекается выборка объема n (N не обязательно велико). Частота появления выборки, содержащей r элементов с качеством P и $n-r$ элементов с качеством Q , равна

$$\begin{aligned}\binom{n}{r} \frac{Np(Np-1) \dots (Np-r+1) Nq(Nq-1) \dots (Nq-n+r+1)}{N(N-1) \dots (N-n+1)} &= \\ &= \binom{n}{r} \frac{(Np)^{[r]} (Nq)^{[n-r]}}{N^{[n]}}.\end{aligned}\quad (5.47)$$

Действительно, существуют $\binom{N}{n}$ различных выборок объема n , причем r элементов с качеством P могут быть выбраны $\binom{Np}{r}$ различными способами, а $n - r$ элементов с качеством Q могут быть выбраны $\binom{Nq}{n-r}$ различными способами (предполагается, конечно, что $r \leq Np$, $n - r \leq Nq$). Выражение (5.47) есть не что иное, как иная запись отношения $\binom{Np}{r} \binom{Nq}{n-r} / \binom{N}{n}$.

Мы приходим к рассмотрению дискретного распределения с общим членом вида

$$f_j = \frac{1}{N^{[n]}} \binom{n}{j} (Np)^{[j]} (Nq)^{[n-j]}, \quad j = 0, 1, \dots, n; \quad Np \geq n \leq Nq. \quad (5.48)$$

Аналогия с биномиальным распределением очевидна. При $N \rightarrow \infty$ распределение (5.48) стремится к биномиальному.

Производящая функция, соответствующая (5.48),

$$P(t) = \frac{1}{N^{[n]}} \sum_{j=0}^n \left\{ \binom{n}{j} (Np)^{[j]} (Nq)^{[n-j]} t^j \right\}, \quad (5.49)$$

равна

$$\frac{(Nq)^{[n]}}{N^{[n]}} \sum \frac{(Np)^{[j]} n^{[j]}}{(Nq - n + j)^{[j]}} \frac{t^j}{j!}.$$

Иными словами, она представляет собой гипергеометрическую функцию

$$\frac{(Nq)^{[n]}}{N^{[n]}} F(\alpha, \beta; \gamma, t),$$

где

$$\alpha = -n, \quad \beta = -Np, \quad \gamma = Nq - n + 1. \quad (5.50)$$

Поэтому распределение (5.48) называется гипергеометрическим. Функция $F(\alpha, \beta; \gamma, t)$ представляется рядом

$$F(\alpha, \beta; \gamma, t) = 1 + \frac{\alpha\beta}{\gamma} \frac{t}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)} \frac{t^2}{2!} + \dots$$

Хорошо известно, что она удовлетворяет дифференциальному уравнению

$$t(1-t) \frac{d^2 F}{dt^2} + \{\gamma - (\alpha + \beta + 1)t\} \frac{dF}{dt} - \alpha\beta F = 0, \quad (5.51)$$

в чем, впрочем, легко можно убедиться и непосредственно.

Если в (5.49) положить $t = e^\theta$, то получим характеристическую функцию гипергеометрического распределения. Делая эту

замену в (5.51), после нескольких преобразований и подстановки вместо α , β , γ их значений по формулам (5.50), находим

$$(1 - e^\theta) \left\{ \frac{d^2\varphi}{d\theta^2} - (n + Np) \frac{d\varphi}{d\theta} + nNp\varphi \right\} - Nnp\varphi + N \frac{d\varphi}{d\theta} = 0. \quad (5.52)$$

Так как $\varphi = \sum \mu'_j \theta^j / j!$, то, приравнявая к нулю коэффициент при θ^0 в левой части (5.52), получаем

$$\left. \begin{aligned} -Nnp + N\mu'_1 &= 0, \\ \mu'_1 &= np. \end{aligned} \right\} \quad (5.53)$$

Таким образом, среднее значение гипергеометрического распределения не зависит от N и совпадает со средним соответствующего биномиального распределения.

Чтобы найти моменты относительно среднего, заменим в (5.52) φ на $e^{np\theta}\varphi$. Тогда получим

$$(1 - e^\theta) \left[\frac{d^2\varphi}{d\theta^2} + \frac{d\varphi}{d\theta} \{n(p - q) - Np\} + (N - n)nq\varphi \right] + N \frac{d\varphi}{d\theta} = 0. \quad (5.54)$$

Отсюда, рассматривая коэффициенты при θ , θ^2 , θ^3 , находим

$$\left. \begin{aligned} \mu_2 &= \frac{npq(N - n)}{N - 1}, \\ \mu_3 &= \frac{npq(q - p)(N - n)(N - 2n)}{(N - 1)(N - 2)}, \\ \mu_4 &= \frac{npq(N - n)}{(N - 1)(N - 2)(N - 3)} \times \\ &\quad \times [N(N + 1) - 6n(N - n) + \\ &\quad + 3pq\{N^2(n - 2) - Nn^2 + 6n(N - n)\}]. \end{aligned} \right\} \quad (5.55)$$

Вообще, если E означает операцию увеличения порядка момента на единицу, например $E\mu_r = \mu_{r+1}$, то

$$N\mu_{r+1} = \{(1 + E)^r - E^r\} [\mu_2 - \{Np + n(q - p)\} \mu_1 + + \{npq(N - n)\} \mu_0]. \quad (5.56)$$

Как и следовало ожидать, при $N \rightarrow \infty$ эти значения стремятся к соответствующим значениям для биномиального распределения.

5.19 Пример гипергеометрического распределения, возникающего в практике, приведен в таблице 5.5, показывающей частоту появления на руках карт определенной масти при игре в вист. Здесь $N=52$ — число карт в колоде, $n=13$ и $p=1/4$. Соот-

ветствующие частоты вычисляются, следовательно, по формуле

$$\frac{1}{52^{13}} \binom{13}{j} 13! 39^{13-j}.$$

Они приведены в третьем столбце таблицы. Согласие наблюдаемых частот с теоретическими достаточно хорошее.

Т а б л и ц а 5.5

Распределение 3400 «первых рук» при игре в вист в соответствии с числом козырей на руке (К. Пирсон (1924а))

Число козырей на руке	Наблюденная частота	Частота гипергеометрического распределения	Число козырей на руке	Наблюденная частота	Частота гипергеометрического распределения
0	35	43,5	5	444	424,0
1	290	272,2	6	115	141,3
2	696	700,0	7	21	30,0
3	937	973,5	8	11	4,0
4	851	811,3	9 и больше	0	0,2
			Полная частота	3400	3400

Либерман и Оуэн (G. J. Lieberman and D. B. Owen, 1961, Tables of the Hypergeometric Probability Distribution, Stanford U. P.) опубликовали значения ф. р. и частот для $N=1(1)50(10)100$ и всех n , а также для $N=200(100)2000$ и некоторых n . Уайз (1954) нашел быстро сходящееся разложение для ф. р. в терминах неполной бета-функции. Сэндифорд (Sandiford, 1960, *J. Amer. Statist. Ass.* 55, 718) показал, что хорошую аппроксимацию дает биномиальное распределение с тем же средним и дисперсией, что и у гипергеометрического распределения *).

Нормальное распределение

5.20 Мы уже отмечали в примерах 4.6 и 4.9, что биномиальное и пуассоновское распределения, выраженные в нормированной форме, стремятся к распределению

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx, \quad -\infty \leq x \leq \infty. \quad (5.57)$$

*) Еще более точной является аппроксимация, учитывающая не два, а три момента (Большев, 1964, Теория вероят. и ее примен. IX, 687), (Прим. ред.)

Это распределение есть частный случай более общего распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu_1')^2 \right\} dx, \quad -\infty \leq x \leq \infty, \quad \sigma > 0, \quad (5.58)$$

называемого нормальным*). Нормальное распределение является наиболее важным теоретическим распределением в статистике. Выражение (5.57) представляет собой нормированное нормальное распределение. Напомним некоторые из его свойств, указанных в рассмотренных выше примерах.

Легко можно показать, что х. ф. распределения (5.58) выражается формулой

$$\varphi(t) = \exp \left(it\mu_1' - \frac{1}{2} t^2 \sigma^2 \right). \quad (5.59)$$

Отсюда

$$\left. \begin{aligned} \mu_{2r} &= \frac{(2r)!}{2^r r!} \sigma^{2r}, \\ \mu_{2r+1} &= 0, \quad r \geq 1. \end{aligned} \right\} \quad (5.60)$$

Производящая функция семинвариантов имеет вид

$$\psi(t) = it\mu_1' - \frac{1}{2} t^2 \sigma^2,$$

так что

$$\kappa_2 = \sigma^2, \quad \kappa_r = 0, \quad r > 2. \quad (5.61)$$

Имеем также $\beta_2 = 3$, $\gamma_2 = 0$; это объясняет выбор стандарта, принятого для нормального эксцесса (см. 3.32).

5.21 Поскольку функция плотности нормированного нормального распределения будет встречаться очень часто, мы введем для нее специальное обозначение. Именно, положим

$$\alpha(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2}. \quad (5.62)$$

Тогда функция распределения будет выражаться интегралами

$$F(x) = \int_{-\infty}^x \alpha(y) dy = \frac{1}{2} + \int_0^x \alpha(y) dy. \quad (5.63)$$

*) Термин «нормальное распределение», введенный К. Пирсоном, сейчас почти общепринят среди английских авторов (как и в советской литературе. — *Перев.*). Авторы из других стран называют его по-разному: второй закон Лапласа, лапласовское распределение, гауссовское распределение, распределение Лапласа — Гаусса, распределение Гаусса — Лапласа. В качестве аппроксимации к биномиальному распределению оно рассматривалось Муавром еще в 1733 году, однако Муавр не изучал его свойств.

Разлагая $\alpha(x)$ в ряд по степеням x и интегрируя почленно, получим

$$F(x) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left(x - \frac{x^3}{2 \cdot 3} + \frac{x^5}{2! 2^2 \cdot 5} - \dots \right). \quad (5.64)$$

Это разложение, однако, не очень полезно для вычисления $F(x)$, так как, за исключением малых x , ряд сходится слишком медленно.

5.22 Для вычисления $F(x)$ при больших x можно использовать разложение $F(x)$ в асимптотический ряд. Заметим, что

$$\frac{\exp\left\{-\frac{1}{2}x^2(1+u^2)\right\}}{1+u^2} = \int_{\frac{1}{2}x^2}^{\infty} \exp\{-v(1+u^2)\} dv.$$

Возьмем интеграл по u от обеих частей в пределах от $-\infty$ до ∞ и переменим в правой части порядок интегрирования (это возможно в силу равномерной сходимости). Получим

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\exp\left\{-\frac{1}{2}x^2(1+u^2)\right\}}{1+u^2} du &= \int_{\frac{1}{2}x^2}^{\infty} \int_{-\infty}^{\infty} \exp\{-v(1+u^2)\} du dv = \\ &= \sqrt{2\pi} \int_{\frac{1}{2}x^2}^{\infty} \frac{1}{\sqrt{2v}} e^{-v} dv = \sqrt{2\pi} \int_x^{\infty} e^{-\frac{1}{2}u^2} du. \end{aligned}$$

Следовательно,

$$\begin{aligned} 1 - F(x) &= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{1}{2}u^2} du = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp\left\{-\frac{1}{2}x^2(1+u^2)\right\}}{1+u^2} du = \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{1}{2}x^2u^2\right)}{1+u^2} du. \end{aligned} \quad (5.65)$$

Отношение «хвоста» распределения $1 - F(x)$ к $\alpha(x)$ называют отношением Миллса *) и обозначают $R(x)$. Таким образом,

$$R(x) = \{1 - F(x)\} / \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} = e^{\frac{1}{2}x^2} \int_x^{\infty} e^{-\frac{1}{2}u^2} du, \quad (5.66)$$

*) Оно было табулировано Миллсом в 1926 году, однако в действительности рассматривалось другими авторами и до него.

и поэтому согласно (5.65)

$$R(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{1}{2} x^2 u^2\right)}{1+u^2} du. \quad (5.67)$$

Заменив $\frac{1}{2} x^2 u^2$ на t , получим

$$R(x) = \frac{x}{2\sqrt{\pi}} \int_0^{\infty} \frac{e^{-t} t^{-\frac{1}{2}}}{t + \frac{1}{2} x^2} dt.$$

Разлагая знаменатель и вычисляя по формулам для гамма-функции получающиеся интегралы, находим

$$\begin{aligned} R(x) &= \frac{x}{2\sqrt{\pi}} \int_0^{\infty} \frac{2}{x^2} \left(1 - \frac{2t}{x^2} + \frac{4t^2}{x^4} - \dots\right) e^{-t} t^{-\frac{1}{2}} dt = \\ &= \frac{1}{x} - \frac{1}{x^3} + \frac{1 \cdot 3}{x^5} - \dots + (-1)^j \frac{1 \cdot 3 \cdot 5 \dots (2j-1)}{x^{2j+1}} + R_j(x), \end{aligned} \quad (5.68)$$

где

$$|R_j(x)| = \frac{2x^{-(2j+1)}}{2\sqrt{\pi}} \int_0^{\infty} \frac{t^{2j+2} e^{-t} t^{-\frac{1}{2}}}{t + x^2} dt.$$

Построенный ряд не сходится, но

$$|R_j(x)| < \frac{2x^{-(2j+1)}}{2\sqrt{\pi}} \int_0^{\infty} t^{2j+\frac{1}{2}} e^{-t} dt < x^{-(2j+1)} 1 \cdot 3 \cdot 5 \dots (2j-1), \quad (5.69)$$

т. е. остаточный член $R_j(x)$ при всех j по абсолютной величине меньше последнего не включенного в него члена, так что ряд этот асимптотический. При больших x полученное разложение достаточно эффективно.

5.23 Наиболее полезным инструментом для вычисления $F(x)$, или эквивалентно $R(x)$, являются непрерывные дроби. Вот одно из разложений в непрерывную дробь, предложенное Лапласом в 1805 году. Положим

$$z(t) = xR\{x(1-t)\} = \frac{x}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\exp\left\{-\frac{1}{2} x^2 u^2 (1-t)^2\right\}}{1+u^2} du. \quad (5.70)$$

Тогда

$$\frac{1}{x} \frac{dz}{dt} = -x \{(1-t)z - 1\}. \quad (5.71)$$

Следовательно, если z разлагается в ряд по степеням t ,

$$z = \sum_{j=0}^{\infty} y_j t^j,$$

то, подставляя это разложение в (5.71) и идентифицируя коэффициенты при одинаковых степенях, получим

$$\frac{1}{x^2} (r+1) y_{r+1} + y_r - y_{r-1} = 0. \quad (5.72)$$

Отсюда

$$\frac{y_r}{y_{r-1}} = \frac{1}{1 + \frac{r+1}{x^2} \frac{y_{r+1}}{y_r}}.$$

Последовательное применение этой формулы дает

$$\frac{y_1}{y_0} = \frac{x}{x+} \frac{2}{x+} \frac{3}{x+} \cdots \frac{n}{x+} \cdots$$

Из (5.70) имеем также

$$y_0 = xR(x), \quad y_1/x^2 = 1 - y_0.$$

Комбинируя найденные формулы, получаем

$$R(x) = y_0/x = \frac{1}{x+} \frac{1}{x+} \frac{2}{x+} \frac{3}{x+} \cdots \frac{n}{x+} \cdots \quad (5.73)$$

Это выражение было использовано Шеппардом (1939, посмертное издание) при вычислении стандартных таблиц нормального распределения, описанных в 5.25. В конце этого тома мы приводим таблицы, достаточные для иллюстрации теории и примеров, содержащихся в этой книге.

5.24 Шентон (1954) предложил разложение в иную непрерывную дробь, сходящуюся значительно быстрее лапласовской. Положив

$$\bar{R}(x) = \int_0^x \alpha(y) dy / \alpha(x) = \frac{1}{2\alpha(x)} - R(x),$$

Шентон получил

$$\bar{R}(x) = \frac{x}{1-} \frac{x^2}{3+} \frac{2x^2}{5-} \frac{3x^2}{7+} \cdots \quad (5.74)$$

Таблицы нормального распределения

5.25 (а) Таблицы Шеппарда содержат, помимо другого материала, значения отношения Миллса с точностью до 12 знаков с шагом 0,01 по x и с точностью до 24 знаков с шагом по x , равным 0,1. Кроме того, в них приводятся натуральные логарифмы хвоста распределения с точностью до 16 знаков с шагом 0,1 по x .

(б) В Biometrika Tables имеется основанная на работе Шеппарда таблица, содержащая значения функции плотности и функции распределения нормированного распределения с точностью до семи знаков для $x=0(0,01)4,50$ и с точностью до 10 знаков для $x=4,50(0,01)6,00$. Кроме этого, там имеется вспомогательная таблица до $x=500$, а также обратные таблицы, позволяющие по значению функции распределения находить квантиль и значение функции плотности.

(в) В Tables of Normal Probability Functions *) (National Bureau of Standards, Applied Mathematics Series 23, Washington, 1953) приводятся значения функции плотности и функции $\{2F(x)-1\}$, где F — функция распределения, для $x=0(0,0001)1(0,001)7,800$ с точностью до 15 знаков, а также значения функции плотности и функции $2\{1-F(x)\}$ с точностью до семи значащих цифр для $x=6(0,01)10$. Кроме того, там имеются вспомогательные таблицы для больших значений x .

(г) В The Kelley Statistical Tables *) (Harvard U. P., 1948) содержится восьмизначная таблица квантилей и функции плотности по значениям $0,0001(0,0001)0,9999$ функции нормированного распределения. Более точное табулирование производится для 10 меньших и 10 больших значений, доходящих до $0,000\ 000\ 001$ и $0,999\ 999\ 999$.

5.26 Вид функции плотности нормированного нормального распределения

$$\alpha(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

показан на рис. 5.1. Функция эта симметрична, не обращается в нуль и, монотонно убывая, очень быстро приближается к нулю при возрастании x ; в точках $x=\pm 1$ у нее имеются перегибы.

Среднее отклонение для нормированного нормального распределения равно

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| e^{-\frac{1}{2}x^2} dx = \sqrt{\frac{2}{\pi}} \int_0^{\infty} x e^{-\frac{1}{2}x^2} dx = \sqrt{\frac{2}{\pi}} = 0,79788, \quad (5.75)$$

а дисперсия равна, конечно, единице, поскольку распределение нормировано. С помощью таблиц легко можно найти, что квантили отстоят от нуля на $0,67448975$.

5.27 Данные таблицы 1.7 о распределении мужчин по росту могут служить примером распределения из практики, очень

*) Таблицы (в) и (г) изданы в СССР с переводом пояснений на русский язык: «Таблицы вероятностных функций», т. II (серия «Библиотека математич. таблиц», вып. 3), М., Изд-во Вычислит. центра АН СССР, 1959; Т. Л. Келли, «Статистические таблицы», М., Изд-во Вычислит. центра АН СССР, 1965.

Подробные таблицы для отношения Миллса, плотности, функции и квантилей нормального распределения помещены в книгах: «Таблицы математической статистики» и «Сборник статистических таблиц». (Прим. ред.)

близкого к нормальному. Таблица 5.6 содержит наблюдаемые частоты и частоты нормального распределения, среднее и стандартное отклонение которого те же, что и у наблюдаемого распределения (67,46 и 2,56 дюймов соответственно).

Соответствие, очевидно, довольно хорошее. Следует, однако, отметить, что, в то время как теоретическое распределение неограниченно, наблюдаемое распределение ограничено, поскольку рост не может быть отрицательным. В нашем случае относительная частота нормального распределения вне интервала 57—77 дюймов столь мала, что указанное обстоятельство несущественно, но в общем случае, когда ограниченное распределение аппроксимируется неограниченным, нужно помнить, что соответствие вблизи хвостов может не быть хорошим.

5.28 Поскольку нормальное распределение можно рассматривать как предельное для биномиального распределения, то

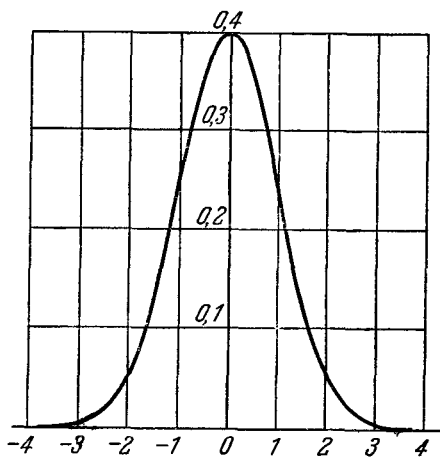


Рис. 5.1. Плотность нормального рас-

$$\text{пределения } y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

Таблица 5.6

Частоты распределения 8 585 мужчин по росту (таблица 1.7) в сравнении с теоретическими частотами нормального распределения с теми же средним и дисперсией

Рост (дюймы)	Наблюденная частота	Теоретическая частота	Рост (дюймы)	Наблюденная частота	Теоретическая частота
57—	2	1	68—	1 230	1 234
58—	4	3	69—	1 063	989
59—	14	11	70—	646	682
60—	41	33	71—	392	405
61—	83	88	72—	202	207
62—	169	200	73—	79	91
63—	394	395	74—	32	34
64—	669	569	75—	16	11
65—	990	976	76—	5	3
66—	1 223	1 227	77—	2	1
67—	1 329	1 326	Полная частота	8 585	6 586

естественно поставить вопрос о формах предельных распределений, если такие существуют, для гипергеометрического распределения. Разность между двумя последовательными частотами гипергеометрического распределения (см. (5.48)) есть

$$\begin{aligned} \frac{1}{N^{[n]}} \frac{n!}{r!(n-r-1)!} (Np)^{[r]} (Nq)^{[n-r-1]} \left\{ \frac{Np-r}{r+1} - \frac{Nq-n+r+1}{n-r} \right\} = \\ = \frac{1}{N^{[n]}} \frac{n!}{r![n-r]!} (Np)^{[r]} (Nq)^{[n-r]} \left\{ \frac{Nnp-Nq+n-1-r(N+2)}{(r+1)(Nq-n+r+1)} \right\}. \end{aligned}$$

Отношение этой разности к r -й частоте равно

$$\frac{\Delta y_r}{y_r} = \frac{A + Br}{C + Dr + Er^2},$$

где $A, \dots, E$ — константы. Когда распределение нормировано, Δy_r в пределе есть приращение, отвечающее малому увеличению r . Мы приходим, таким образом, к рассмотрению дифференциального уравнения

$$\frac{df}{f} = \frac{A + Bx}{C + Dx + Ex^2} dx, \quad (5.76)$$

определяющего функцию плотности.

Графики плотностей, отвечающих этому уравнению, образуют так называемое семейство кривых К. Пирсона и будут рассмотрены с несколько иной точки зрения в следующей главе.

Некоторые обобщения

5.29 Распределения, изученные в этой главе, допускают различные обобщения. Одно из них — нормальное распределение в многомерном случае — настолько важно, что в дальнейшем ему будет посвящена отдельная глава. Настоящую главу мы закончим кратким рассмотрением двух обобщений биномиального распределения.

Мультиномиальное распределение

5.30 Предположим, что элементы генеральной совокупности классифицируются не в соответствии с наличием или отсутствием качества P , как в 5.2, а разбиваются на $k+1$ классов в соответствии с наличием качеств $P_1, P_2, \dots, P_k$ или отсутствием их всех — P_0 . Так, например, при $k=3$ P_1 может обозначать наличие голубых глаз, P_2 — серых, P_3 — коричневых, а P_0 — глаза какого-либо другого цвета или вообще отсутствие глаз. Если $p_1, p_2, \dots, p_k, p_0$ обозначают соответственно относительные количества (пропорции) элементов каждого класса, то, об-

общая рассуждения 5.2, легко видеть, что частоты для множества из n элементов задаются разложением мультиномиального выражения

$$(p_0 + p_1 + p_2 + \dots + p_k)^n. \quad (5.77)$$

Иначе говоря, частота множества, имеющего r_0 элементов с качеством P_0 , r_1 элементов с качеством P_1 и т. д., равна

$$\frac{n!}{r_0! r_1! \dots r_k!} p_0^{r_0} p_1^{r_1} \dots p_k^{r_k}. \quad (5.78)$$

Характеристическая функция, относящая переменную t_i к p_i ($i=1, \dots, k$), имеет вид

$$(p_0 + p_1 e^{it_1} + p_2 e^{it_2} + \dots + p_k e^{it_k})^n. \quad (5.79)$$

С помощью (5.79) можно обычным образом найти моменты и смешанные моменты. Мы ограничимся здесь тем, что выпишем некоторые семиинварианты до четвертого порядка для $k=3$. (Остальные семиинварианты тех же порядков получаются из соображений симметрии.)

$$\left. \begin{aligned} \kappa_{1000} &= n p_0, \quad \kappa_{2000} = n p_0 q_0, \quad \kappa_{1100} = -n p_0 p_1, \\ \kappa_{3000} &= n p_0 q_0 (q_0 - p_0), \quad \kappa_{2100} = -n p_0 p_1 (q_0 - p_0), \\ &\quad \kappa_{1110} = 2n p_0 p_1 p_2, \\ \kappa_{4000} &= n p_0 q_0 (1 - 6 p_0 q_0), \quad \kappa_{3100} = -n p_0 p_1 (1 - 6 p_0 q_0), \\ \kappa_{2200} &= -n p_0 p_1 \{ (q_0 - p_0) (q_1 - p_1) + 2 p_0 p_1 \}, \\ \kappa_{2110} &= 2n p_0 p_1 p_2 (q_0 - 2 p_0), \quad \kappa_{1111} = -6n p_0 p_1 p_2 p_3. \end{aligned} \right\} \quad (5.80)$$

Здесь для краткости положено, как обычно, $q_i = 1 - p_i$.

Двумерное биномиальное распределение

5.31 Элементы генеральной совокупности могут также классифицироваться в соответствии с двумя качествами, например: голубоглазый или не голубоглазый и мужчина или женщина. (Вообще можно было бы рассматривать случай, когда имеется несколько качеств с несколькими категориями в каждом из них, но такая ситуация, ни в чем существенном не отличаясь, привела бы лишь к большей громоздкости и сложности.)

Пусть наличие или отсутствие одного качества обозначается соответственно P , Q ($Q=1-P$), а другого P' , Q' ($Q'=1-P'$). Относительные количества элементов различных классов могут

быть представлены следующей таблицей:

	P'	Q'	Всего
P	p_{11}	p_{10}	p
Q	p_{01}	p_{00}	q
Всего	p'	q'	1

(5.81)

С некоторых точек зрения получающееся здесь распределение можно рассматривать как мультиномиальное, задаваемое выражением

$$(p_{00} + p_{01} + p_{10} + p_{11})^n. \quad (5.82)$$

Обычно, однако, нам более интересно распределение числа элементов с качествами P и P' , чем с обоими качествами P и P' и т. д. Х. ф. совместного распределения числа элементов с качеством P и с качеством P' равна

$$\varphi(t) = (p_{00} + p_{01}e^{\theta_1} + p_{10}e^{\theta_2} + p_{11}e^{\theta_1 + \theta_2})^n, \quad (5.83)$$

где, как обычно, $\theta_1 = it_1$, $\theta_2 = it_2$. Положим

$$p_{(11)} = p_{11} - pp'. \quad (5.84)$$

Разлагая $\log \varphi(t)$, находим семиинварианты двумерного распределения

$$\left. \begin{aligned} \kappa_{11} &= np_{(11)}, \quad \kappa_{21} = np_{(11)}(q - p), \quad \kappa_{31} = np_{11}(1 - 6pq), \\ \kappa_{22} &= np_{(11)}\{(q - p)(q' - p') - 2p_{(11)}\}. \end{aligned} \right\} \quad (5.85)$$

Другие семиинварианты, такие как κ_{12} , можно найти, пользуясь симметрией. Конечно, семиинварианты вида κ_{20} являются семиинвариантами соответствующих одномерных распределений, задаваемых биномиальными выражениями $(q + p)^n$ и $(q' + p')^n$.

5.32 Простое вычисление показывает, что если величина $p_{(11)}$, определяемая равенством (5.84), равна нулю, то х. ф. принимает вид

$$(q + pe^{\theta_1})^n (q' + p'e^{\theta_2})^n. \quad (5.86)$$

Качества при этом независимы. Если перенести начало отсчета в точку, координаты которой равны средним значениям вариантов, то (5.83) перейдет в

$$\begin{aligned} \log \varphi(t) &= n \log (p_{00} + p_{01}e^{\theta_1} + p_{10}e^{\theta_2} + p_{11}e^{\theta_1 + \theta_2}) - np\theta_1 - np'\theta_2 = \\ &= \frac{1}{2} n (pq\theta_1^2 + 2p_{(11)}\theta_1\theta_2 + p'q'\theta_2^2) + o(n). \end{aligned} \quad (5.87)$$

А если перейти к нормированным вариантам (дисперсии исходных вариантов равны соответственно npq и $np'q'$) и устремить n

к бесконечности, то в пределе получим

$$\log \varphi(t) = \frac{1}{2} (\theta_1^2 + 2\rho\theta_1\theta_2 + \theta_2^2), \quad (5.88)$$

где

$$\rho = \frac{p_{(11)}}{(pq p' q')^{1/2}}. \quad (5.89)$$

В примере 3.17 было показано, что (5.88) есть логарифм х. ф. распределения

$$dF = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} (x_1^2 - 2\rho x_1 x_2 + x_2^2) \right\} dx_1 dx_2, \\ -\infty \leq x_1, x_2 \leq \infty. \quad (5.90)$$

Таким образом, двумерное биномиальное распределение стремится к распределению (5.90), которое можно рассматривать как двумерный аналог нормального распределения.

Варианты x_1 и x_2 независимы тогда и только тогда, когда $\rho=0$, или, равносильно, когда $p_{(11)}=0$ или $p_{11}=pp'$. Это условие в свою очередь означает, что среди элементов с качеством P' доля тех элементов, которые обладают качеством P , та же, что и их доля среди элементов с качеством Q' , так как если такое условие выполнено, то $p_{10}=p-p_{11}=pq'$. Таблица (5.81) принимает при этом вид

	P'	Q'	Всего
P	pp'	pq'	p
Q	qp'	qq'	q
Всего	p'	q'	1

(5.91)

Подобного рода ситуации мы обсудим более подробно, когда будем заниматься выборочным изучением качественных признаков. Заметим еще здесь же, что «независимость», как мы определили ее выше, согласуется с обычным смыслом, который вкладывается в это понятие, когда речь идет о независимости качеств.

Уишарт (1949) дал общую формулу для семинивариантов многомерного мультиномиального распределения.

УПРАЖНЕНИЯ

5.1 Показать, что для биномиального распределения

$$\kappa_{r+1} = pq \frac{d\kappa_r}{dp}, \quad r > 1.$$

Пользуясь этим уравнением, доказать равенства ($c=pq$; $g=q-p$):

$$\kappa_2 = pc, \quad \kappa_3 = pcg, \quad \kappa_4 = n(c - 6c^2), \quad \kappa_5 = ng(c - 12c^2),$$

$$\kappa_6 = n(c - 30c^2 + 120c^3), \quad \kappa_7 = ng(c - 60c^2 + 360c^3),$$

$$\kappa_8 = n(c - 126c^2 + 1680c^3 - 5040c^4).$$

(Ср. Фриш (1926), Холдейн (1939). Холдейн приводит формулы до κ_{12} включительно.)

5.2 Показать, что для *неполных* моментов биномиального распределения (относительно среднего)

$$\mu_r = \sum_{i=p}^n (j-np)^r \binom{n}{j} p^j q^{n-j}$$

(суммирование ведется, начиная с некоторого p , представляющего собой одно из значений вариаты) выполняются уравнения (5.10), т. е.

$$\mu_{r+1} = pq \left(nr\mu_{r-1} + \frac{d\mu_r}{dp} \right). \quad (\text{Романовский, 1925.})$$

5.3 Полагая $T_j = \binom{n}{j} p^j q^{n-j}$, показать, что неполные моменты биномиального распределения даются формулами

$$\mu_0 = \sum_{j=p}^n T_j,$$

$$\mu_1 = pqT_p,$$

$$\mu_2 = pqT_p \{p - (n+1)p\} + npq\mu_0,$$

$$\mu_3 = pqT_p [\{p - (n+1)p\}^2 + pq(2n-1)] + npq(q-p)\mu_0,$$

и вообще

$$\mu_r = pqT_p (p-np)^{r-1} + npq \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_{j-p} \sum_{j=0}^{r-2} \binom{r-1}{j} \mu_{j+1}.$$

(Фриш (1926). Эти формулы обобщают соотношения (5.9) на неполные моменты.)

5.4 Показать, что среднее отклонение биномиального распределения равно

$$2(np + \xi) \binom{n}{np + \xi} p^{np + \xi} q^{nq - \xi + 1},$$

где $(np + \xi)$ — наименьшее целое число, превосходящее np .

(Джонсон, 1957.)

5.5 Из соотношения (5.56) вывести рекуррентную формулу

$$\{(1+E)^r - E^r\} (npq\mu_0 - p\mu_1) = \mu_{r+1}$$

для моментов биномиального распределения и аналогичную формулу

$$\{(1+E)^r - E^r\} \lambda\mu_0 = \mu_{r+1}$$

для моментов пуассоновского распределения.

(К. Пирсон, 1924a).

5.6 Показать, что если распределение типа (5.24) хорошо аппроксимируется биномом $(Q+P)^n$ (т. е. совокупностью элементов его разложения), то

$$nP = n\bar{p}, \quad nQP = n\bar{p}\bar{q} - n\sigma_p^2,$$

так что $P = \bar{p} + \sigma_p^2/\bar{p}$ и, следовательно, P и n положительны.

Если, однако, вместо распределения типа (5.24) взять распределение типа (5.28), то

$$P = \bar{p} - (n-1)\sigma_p^2/\bar{p},$$

так что P , а следовательно, и n могут быть отрицательными.

(«Стюдент», 1919.)

5.7 (Неймановское распределение типа А). Варианта r принимает значения $r = 0, 1, \dots$ с частотами

$$f_r = \frac{\lambda_2^r}{r!} e^{-\lambda_1} \sum_{j=0}^{\infty} \frac{j^r}{j!} (\lambda_1 e^{-\lambda_2})^j, \quad \lambda_1, \lambda_2 > 0.$$

Показать, что производящая функция факториальных семинвариантов равна $\lambda_1 (e^{\lambda_2 t} - 1)$ и что $\mu'_1 = \lambda_1 \lambda_2$, $\mu_2 = \lambda_1 \lambda_2 (1 + \lambda_2)$. Найти также х. ф. и обратить ее, используя упражнение 4.22.

(Ю. Нейман (J. Neyman), 1940, Ann. Math. Statist. 10, 35).

5.8 (Распределение Пойа—Эппли). Варианта r принимает значения $r = 0, 1, \dots$ с частотами

$$f_0 = e^{-\lambda_1},$$

$$f_r = e^{-\lambda_1} \tau^r \sum_{j=1}^r \binom{r-1}{j-1} \frac{1}{j!} \left\{ \frac{\lambda_1 (1-\tau)}{\tau} \right\}^j, \quad r \geq 1.$$

Показать, что производящая функция факториальных семинвариантов равна $\lambda_1 t / (1 - \tau - \tau t)$, и найти среднее и дисперсию. (Упражнения 5.7 и 5.8 содержат примеры распределений числа индивидуумов в колониях, размер которых тоже имеет некоторое распределение. В обоих случаях индивидуумы распределены по пуассоновскому закону; в 5.7 размер колоний тоже пуассоновский, а в 5.8 он имеет J-образное распределение $(1-\tau)\tau^{\rho-1}$, $\rho = 1, 2, \dots$).

5.9 Показать, что если частоты двух симметричных биномиальных распределений порядка n суперпозируются так, что r -й член одного распределения прибавляется к $(r+1)$ -му члену другого, то в результате получаются частоты симметричного биномиального распределения порядка $n+1$ (с точностью до множителя). Вывести отсюда, что если складываются два нормальных распределения, у которых дисперсии одинаковы, а средние отличаются только на небольшую часть стандартного отклонения, то получающееся распределение почти нормально.

5.10 Показать, что если

$$y = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2/\sigma^2\right),$$

то

$$\int_{-\infty}^{\infty} y^2 dx = \frac{1}{2\sigma \sqrt{\pi}}.$$

Следовательно, если нормальное распределение группируется в интервалах с общей частотой N_1 и если N_2 — сумма квадратов частот, то величина

$$\frac{N_1^2}{2N_2 \sqrt{\pi}} = 0,282095 \frac{N_1^2}{N_2}$$

будет оценкой для σ . Проверить, что для данных таблицы 1.7 эта оценка дает значение σ , равное 2,553 (ошибка около 1%).

(Юл, 1938.)

5.11 (Трехмерное пуассоновское распределение). Показать, что когда p_{01} , p_{10} , p_{11} в (5.82) малы, а $np_{11} = \lambda_3$, $np_{10} = \lambda_2 - \lambda_3$, $np_{01} = \lambda_1 - \lambda_3$ конечны, то распределение стремится к предельному, общий член которого имеет вид

$$\frac{\lambda_3^i (\lambda_1 - \lambda_3)^j (\lambda_2 - \lambda_3)^k}{i! j! k!} e^{-\lambda_1 - \lambda_2 + \lambda_3}.$$

5.12 Используя лапласовскую непрерывную дробь (5.73), показать, что

$$\frac{x}{x^2 + 1} < R(x) < \frac{x^2 + 2}{x^3 + 3x}.$$

5.13 Показать, что отношение Миллса $R(x)$ равно

$$R(x) = x \int_{-\infty}^{\infty} \frac{1}{u^2 + x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du.$$

Заменяя интеграл суммой $h \sum_{j=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2}j^2h^2}}{x^2 + j^2h^2}$ и пользуясь первыми

тринадцатью ее членами, проверить, что при $h = \frac{1}{4}$, $x = 1$ будет $R = 0,15865524$.

(Дэс, 1956. Истинное значение R есть 0,15865525.)

5.14 Показать, что если для разложения отношения Миллса в лапласовскую непрерывную дробь (5.73) s -ю подходящую дробь c_s обозначить a_s/b_s , то

$$a_0 = 0, \quad a_1 = 1, \quad b_0 = 1, \quad b_1 = x,$$

$$c_s - c_{s-1} = \frac{(-1)^{s+1} (s-1)!}{b_{s-1} b_s},$$

$$c_s - c_{s-2} = \frac{(-1)^s x (s-2)!}{b_{s-2} b_s}$$

и, следовательно, при $x > 0$

$$c_0 < c_2 < c_4 < \dots < R(x) < \dots < c_5 < c_3 < c_1.$$

(Шентон, 1954.)

5.15 Показать, что отношение Миллса равно

$$R(x) = \int_0^{\infty} \exp\left(-\frac{1}{2}t^2 - tx\right) dt$$

и что

$$R(x) b_s - a_s = \frac{(-1)^s \int_0^{\infty} (t-x)^s e^{-\frac{1}{2}t^2} dx / \sqrt{2\pi}}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}$$

(см. предыдущее упражнение). Пользуясь неравенством Шварца

$$\left| \frac{\int f_1^2 dx \int f_1 f_2 dx}{\int f_1 f_2 dx \int f_2^2 dx} \right| > 0$$

при

$$f_1 = (x-t)^{\frac{1}{2}s} e^{-\frac{1}{4}t^2}, \quad f_2 = (x-t)^{\frac{1}{2}s+1} e^{-\frac{1}{4}t^2},$$

доказать, что

$$R > \frac{\alpha_{2s} + (2s)! \sqrt{x^2 + 8s + 4}}{2\beta_{2s}},$$

$$R < \frac{2\gamma_{2s+1}}{\alpha_{2s+1} + (2s+1)! \sqrt{x^2 + 8s + 8}},$$

где

$$\alpha_s = c_s a_{s+2} + c_{s+2} a_s - 2c_{s+1} a_{s+1},$$

$$\beta_s = b_s b_{s+2} - b_{s+1}^2, \quad \gamma_s = a_s a_{s+2} - a_{s+1}^2.$$

Вывести отсюда неравенство Бирибаума

$$\frac{1}{2} \{-x + \sqrt{x^2 + 4}\} < R(x) < 4/\{3x + \sqrt{x^2 + 8}\}, \quad x > 0$$

(берется положительное значение корня).

(Шентон, 1954.)

5.16 Пусть $v(x) = 1/R(x)$ и $\lambda(x) = dv/dx = v(v-x)$. Показать, что $0 < \lambda < 1$. Полагая

$$\pi(x) = e^{-x^2} / \int_{-\infty}^x e^{-\frac{1}{2}u^2} du \int_x^{\infty} e^{-\frac{1}{2}u^2} du,$$

проверить, что

$$d\pi/dx = 2x\pi(x) \{\lambda(x') - 1\}, \quad \text{где } -|x| \leq x' \leq |x|.$$

Отсюда, очевидно, вытекает, что $\pi(x)$ монотонно убывает при положительных x .

(Сэмпфорд, 1953.)

5.17. Рассматривается интеграл

$$\frac{1}{2\pi} \int_{-a}^a \int_{-a}^a \exp\left\{-\frac{1}{2}(x^2 + y^2)\right\} dx dy.$$

Показать, что он меньше интеграла от той же функции, взятого по кругу площади $4a^2$ с центром в начале координат. Пользуясь этим фактом, вывести неравенство

$$\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}u^2} du < \frac{1}{2} \{1 - \exp(-2x^2/\pi)\}^{\frac{1}{2}}, \quad x > 0.$$

5.18 Убедиться, что функция

(Пойа, 1945.)

$$\bar{R}(x) = y = e^{\frac{1}{2}x^2} \int_0^x e^{-\frac{1}{2}u^2} du$$

удовлетворяет дифференциальному уравнению

$$\frac{dy}{dx} = xy + 1,$$

и доказать равенство

$$y = \frac{x}{1} + \frac{x^3}{1 \cdot 3} + \frac{x^5}{1 \cdot 3 \cdot 5} + \frac{x^7}{1 \cdot 3 \cdot 5 \cdot 7} + \dots$$

(Пойа, 1945.)

5.19 Показать, что для мультиномиального распределения (5.77)

$$\kappa(r_1, r_2, \dots, r_i + 1, \dots, r_k) = a_i \frac{\partial}{\partial a_i} \kappa(r_1, r_2, \dots, r_i, \dots, r_k),$$

где $a_i = p_i/p_0$ и семинвариант справа имеет порядок r_i по P_i и т. д. (Гулдберг, 1935). Пользуясь этим, вывести соотношения (5.80).

5.20 Проверить справедливость формул (5.85) и выражения (5.86).

5.21 В некотором эксперименте наблюдаются группы индивидуумов. Показать, что если число групп имеет пуассоновское распределение, а число индивидуумов в группе — логарифмическое, то общее число наблюдаемых индивидуумов имеет отрицательное биномиальное распределение.

(Кэнуй (Quenouille M. H.) (1949),
Biometrics 5, 162.)

5.22 Пусть z — сумма n независимых вариантов x_i , каждая из которых имеет распределение с х. ф. $\varphi_1(t)$, и пусть n — дискретно распределенная варианта с х. ф. $\varphi_2(t)$, принимающая значения $0, 1, 2, \dots$. Показать, что х. ф., отвечающая варианту z , равна $\varphi_2\left(\frac{\log \varphi_1(t)}{t}\right)$ и что с положительной вероятностью $z=0$, даже если x_i имеют непрерывное распределение.

В частности, если n имеет пуассоновское распределение с параметром λ , то х. ф. z равна $\exp\{\lambda[\varphi_1(t)-1]\}$. Вывести отсюда результаты упражнений 5.7 и 5.21.

5.23 Используя результат упражнения 4.16, показать (ср. 5.13), что если параметр пуассоновского распределения сам имеет распределение III типа (5.31), то в результате получится отрицательное биномиальное распределение (5.33).

5.24 Условное распределение x при заданных μ и λ нормально со средним μ и дисперсией λ . Показать, что если $\varphi(t, u) = M\{\exp(it\mu + iu\lambda)\}$ ($\varphi(t, u)$ — х. ф. совместного распределения λ и μ), то х. ф. безусловного распределения x равна $\varphi\left(t, \frac{1}{2}it^2\right)$.

5.25 Если в упражнении 5.24 μ и λ распределены независимо с х. ф., соответственно, $\varphi_1(t) = (1 + t^2\sigma^2)^{-1}$, $\varphi_2(t) = (1 - 2iut\sigma^2)^{-m}$, m — целое положительное (относительно этих х. ф. см. упражнения 4.3 и 4.18), то распределение x совпадает с распределением суммы $m+1$ независимых вариантов, каждая из которых распределена так же, как μ .

5.26 Показать, что для пуассоновского распределения (5.18) $\mu'_{[r]} = \lambda^r$, а для гипергеометрического распределения (5.48) $\mu'_{[r]} = n^{[r]}(Np)^{[r]}/N^{[r]}$.

СТАНДАРТНЫЕ РАСПРЕДЕЛЕНИЯ — (2)

6.1 В этой главе мы продолжим изучение стандартных распределений статистической теории, начатое в предыдущей главе. Естественно, конечно, желание описать различного рода частотные распределения (значительное число примеров содержится в главе 1) гибкой системой математических формул. Здесь мы рассмотрим три возможных прихода к описанию распределений. Первый из них, ведущий свое начало от Карла Пирсона, основан на отыскании семейства кривых, при помощи которого можно удовлетворительно представить встречающиеся на практике распределения. Второй подход, идущий от Брунса, Грама, Шарлье и Эджворта, основан на представлении данной плотности распределения в виде ряда от производных нормальной плотности. И, наконец, третий подход, данный Эджвортом и другими авторами, состоит в отыскании такой функции от исходных вариантов, распределение которой, по крайней мере приближенно, можно было бы выразить через известное распределение.

Распределения Пирсона

6.2 В 5.28 было отмечено, что в предельном случае гипергеометрическое распределение совпадает с распределением, плотность которого подчиняется уравнению

$$\frac{df}{dx} = \frac{(x-a)f}{b_0 + b_1x + b_2x^2}. \quad (6.1)$$

Это уравнение может быть рассмотрено с несколько иной точки зрения. Если иметь в виду унимодальные распределения, рассмотренные в главе 1, то представляется интересным изучить тот класс плотностей, которые: (а) имеют единственную моду, т. е. $df/dx=0$ в некоторой точке $x=a$; (б) имеют гладкое соприкосновение с осью x на концах интервала, где сосредоточено распределение, т. е. $df/dx=0$, когда $f=0$. Нетрудно видеть, что, вообще говоря, решения уравнения (6.1) удовлетворяют этим условиям. На самом деле далее станет ясно, что среди распре-

делений семейства (6.1) существуют и такие распределения, которые имеют J- и U-образную форму.

Семейство плотностей, определяемых формулой (6.1), известно под названием «семейства распределений Пирсона». Прежде чем переходить к нахождению точных решений, мы рассмотрим некоторые свойства, присущие этому семейству в целом. Имеем

$$(b_0 + b_1x + b_2x^2)df = (x - a)f dx$$

или

$$x^n (b_0 + b_1x + b_2x^2) \frac{df}{dx} = x^n (x - a)f.$$

Интегрируя левую часть по частям и предполагая, что полученные интегралы существуют, мы получаем

$$\begin{aligned} [x^n (b_0 + b_1x + b_2x^2)f]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \{nb_0x^{n-1} + (n+1)b_1x^n + \\ + (n+2)b_2x^{n+1}\} f dx = \int_{-\infty}^{\infty} x^{n+1}f dx - a \int_{-\infty}^{\infty} x^n f dx. \end{aligned} \quad (6.2)$$

Предположим, что выражение в квадратных скобках обращается в нуль на концах распределения или что $\lim_{x \rightarrow \pm\infty} x^{n+2}f \rightarrow 0$, если распределение имеет бесконечный размах. Тогда, пользуясь обозначениями для моментов из (6.2), получим

$$-nb_0\mu'_{n-1} - (n+1)b_1\mu'_n - (n+2)b_2\mu'_{n+1} = \mu'_{n+1} - a\mu'_n$$

или

$$nb_0\mu'_{n-1} + \{(n+1)b_1 - a\}\mu'_n + \{(n+2)b_2 + 1\}\mu'_{n+1} = 0. \quad (6.3)$$

Эти уравнения позволяют определить старшие моменты по младшим. В самом деле, все моменты могут быть выражены через коэффициенты a , b_0 , b_1 и b_2 и моменты $\mu_0 (=1)$ и μ'_1 . И обратно, эти четыре константы можно выразить через моменты μ'_1 , μ'_2 , μ'_3 и μ'_4 или через три момента, взятых относительно среднего: μ_2 , μ_3 , μ_4 . Полагая в (6.3) последовательно $n=0, 1, 2, 3$, мы находим уравнения для a , b_0 , b_1 , b_2 :

$$\left. \begin{aligned} a &= -\frac{\mu_3(\mu_4 + 3\mu_2^2)}{A} = -\frac{V\mu_2 V\beta_1(\beta_2 + 3)}{A'}, \\ b_0 &= -\frac{\mu_2(4\mu_2\mu_4 - 3\mu_3^2)}{A} = -\frac{\mu_2(4\beta_2 - 3\beta_1)}{A'}, \\ b_1 &= -\frac{\mu_3(\mu_4 + 3\mu_2^2)}{A} = -\frac{V\mu_2 V\beta_1(\beta_2 + 3)}{A'}, \\ b_2 &= -\frac{(2\mu_2\mu_4 - 3\mu_3^2 - 6\mu_2^3)}{A} = -\frac{(2\beta_2 - 3\beta_1 - 6)}{A'}, \end{aligned} \right\} \quad (6.4)$$

где

$$A = 10\mu_4\mu_2 - 18\mu_3^2 - 12\mu_2^3, \quad A' = 10\beta_2 - 18 - 12\beta_1. \quad (6.5)$$

В этих формулах нулевое значение принято в качестве среднего, так что распределения семейства (6.1) полностью определяются своими четырьмя первыми моментами.

6.3 Из уравнения (6.1) следует, что мода равна $x=a$. Из (6.4) для пирсоновской меры асимметрии *)

$$Sk = \frac{-a}{V\mu_2} = \frac{V\beta_1(\beta_2+3)}{10\beta_2 - 12\beta_1 - 18}, \quad (6.6)$$

что, впрочем, было отмечено ранее (см. формулу (3.87)).

Далее, если $a=0$, то

$$\frac{d^2f}{dx^2} = \frac{d}{dx} \frac{xf}{b_0 + b_1x + b_2x^2} = \frac{f}{(b_0 + b_1x + b_2x^2)^2} (b_0 - b_2x^2). \quad (6.7)$$

Поэтому точки перегиба графика плотности распределения определяются соотношением

$$x^2 = \frac{b_0}{b_2}. \quad (6.8)$$

Следовательно, у плотности из семейства Пирсона существует не более чем две точки перегиба, и если их действительно две, то они отстоят от моды на одинаковом расстоянии. Но может, конечно, случиться, что одна из точек перегиба находится вне области, где сосредоточено распределение.

6.4 Принимая моду за начало отсчета, уравнение (6.1) можно записать в виде

$$\left. \begin{aligned} \text{или} \quad \frac{d}{dx} (\log f) &= \frac{x-a}{B_0 + B_1(x-a) + B_2(x-a)^2} \\ \frac{d}{dX} (\log f) &= \frac{X}{B_0 + B_1X + B_2X^2} \end{aligned} \right\} \quad (6.9)$$

Таким образом, точное выражение для плотности f можно получить путем интегрирования правой части уравнения (6.9).

Здесь следует различать два основных типа распределений, соответствующих тем случаям, когда знаменатель в (6.9) имеет действительные или мнимые корни.

Тип I (бета-распределение)

6.5 Пусть

$$B_0 + B_1X + B_2X^2 = B_2(X + \alpha_1)(X - \alpha_2), \quad \alpha_1, \alpha_2 > 0,$$

*) Происхождение обозначения Sk объясняется тем, что по-английски асимметрия — skewness. (Прим. перев.)

тогда

$$\begin{aligned} \frac{d}{dX}(\log f) &= \frac{X}{B_2(X + a_1)(X - a_2)} = \\ &= \frac{a_1}{B_2(a_1 + a_2)} \frac{1}{(X + a_1)} + \frac{a_2}{B_2(a_1 + a_2)} \frac{1}{(X - a_2)}. \end{aligned}$$

Отсюда

$$f = k(X + a_1)^{\frac{a_1}{B_2(a_1 + a_2)}}(X - a_2)^{\frac{a_2}{B_2(a_1 + a_2)}}, \quad (6.10)$$

что можно записать в форме

$$f = k \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2}, \quad (6.11)$$

где

$$\frac{m_1}{a_1} = \frac{m_2}{a_2}.$$

Размах этого распределения сосредоточен в интервале $(-a_1, a_2)$, поэтому, интегрируя в этих пределах, найдем

$$1 = k \int_{-a_1}^{a_2} \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2} dx,$$

откуда, положив $x = (a_1 + a_2)y - a_1$, получим

$$\begin{aligned} 1 &= k \int_0^1 y^{m_1} (1 - y)^{m_2} \frac{(a_1 + a_2)^{m_1 + m_2 + 1}}{a_1^{m_1} a_2^{m_2}} dy = \\ &= \frac{k(a_1 + a_2)^{m_1 + m_2 + 1}}{a_1^{m_1} a_2^{m_2}} B(m_1 + 1, m_2 + 1). \end{aligned}$$

Этим равенством определяется константа k . Таким образом,

$$f = \frac{a_1^{m_1} a_2^{m_2}}{(a_1 + a_2)^{m_1 + m_2 + 1} B(m_1 + 1, m_2 + 1)} \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2}. \quad (6.12)$$

В качестве начала отсчета здесь выбрана мода. Если в качестве начальной точки взять начало распределения, то получим

$$f = \frac{a_2^{m_2}}{(a_1 + a_2)^{m_1 + m_2 + 1} B(m_1 + 1, m_2 + 1)} x^{m_1} \left(1 - \frac{x - a_1}{a_2}\right)^{m_2}$$

или, полагая $a_1 + a_2 = 1$,

$$f = \frac{1}{B(m_1 + 1, m_2 + 1)} x^{m_1} (1 - x)^{m_2}. \quad (6.13)$$

6.6 Обычно это выражение записывают в следующей форме, которая стала почти стандартной:

$$f = \frac{1}{\beta(p, q)} x^{p-1} (1-x)^{q-1}, \quad p, q > 0. \quad (6.14)$$

Соответствующее распределение сосредоточено на интервале $(0, 1)$. Отвечающая (6.14) функция распределения является неполной бета-функцией, а само распределение часто называется бета-распределением. Если p и q оба превышают единицу, то кривая плотности имеет единственную моду $(p-1)/(p+q-2)$ и обращается в нуль на концах интервала $(0, 1)$. Если либо p , либо q лежит между 0 и 1, то одна из крайних ординат бесконечна и распределение имеет J-образную форму.

Условие, что $B_0 + B_1X + B_2X^2$ имеет действительные корни противоположных знаков, которое было использовано при выводе формулы (6.14), эквивалентно тому, что B_0 и B_2 имеют противоположные знаки, что в свою очередь эквивалентно условию $B_1^2/(4B_0B_2) < 0$. В терминах β_1 и β_2 это условие в силу (6.4) можно выразить неравенством

$$\frac{\beta_1(\beta_2 + 3)^2}{4(2\beta_2 - 3\beta_1 - 6)(4\beta_2 - 3\beta_1)} < 0. \quad (6.15)$$

Величина в левой части обычно обозначается буквой κ и служит критерием при различении основных типов распределения Пирсона.

6.7 Если корни $B_0 + B_1X + B_2X^2$ действительны и одного знака, то аналогичным образом можно показать, что

$$f = kx^{-q_1}(x-a)^{q_2}, \quad q_1 > q_2 - 1, \quad (6.16)$$

где $a \leq x \leq \infty$, если $a > 0$, и $-\infty \leq x \leq a$, если $a < 0$. Это распределение называется VI типом Пирсона, но поскольку с помощью простого преобразования $y = a/x$ оно сводится к распределению I типа, то нет надобности изучать его отдельно. Величина κ в (6.15) для этого случая больше единицы. В форме

$$f = \frac{kx^{p-1}}{(1+x)^{p+q}}, \quad 0 \leq x \leq \infty, \quad (6.17)$$

это распределение известно как бета-распределение второго рода. Здесь $1/k = B(p, q)$, как и для распределения типа I. Распределения (6.14) и (6.17) мы будем называть соответственно бета-распределением первого и второго рода с параметрами p и q .

Тип IV

6.8 Если корни $B_0 + B_1X + B_2X^2$ мнимые, то (вводя очевидные обозначения)

$$\frac{d}{dX}(\log f) = \frac{X}{B_2 \left\{ \left(X + \frac{B_1}{2B_2} \right)^2 + \frac{B_0}{B_2} - \frac{B_1^2}{4B_2^2} \right\}} = \frac{X}{B_2 \{(x + \gamma)^2 + \delta^2\}},$$

откуда

$$\log f = \log k + \frac{1}{2B_2} \log \{(X + \gamma)^2 + \delta^2\} - \frac{\gamma}{B_2\delta} \operatorname{arctg} \frac{X + \gamma}{\delta},$$

$$f = k \{(X + \gamma)^2 + \delta^2\}^{\frac{1}{2B_2}} \exp \left\{ -\frac{\gamma}{B_2\delta} \operatorname{arctg} \frac{X + \gamma}{\delta} \right\}.$$

Распределения с такой плотностью принадлежат системе Пирсона IV типа. Их плотности обычно записываются в виде

$$f = k \left(1 + \frac{x^2}{a^2} \right)^{-m} \exp \left\{ -v \operatorname{arctg} \left(\frac{x}{a} \right) \right\}. \quad (6.18)$$

Эти распределения сосредоточены на всей прямой и являются унимодальными. Но практически с такими распределениями трудно обращаться, несмотря на существование специальных таблиц. Дело в том, что функцию распределения, являющуюся интегралом от (6.18), не удастся выразить через элементарные функции. Проще всего для ее отыскания воспользоваться численным интегрированием функции плотности. Отметим, что величина x из (6.15) в этом случае лежит между 0 и 1.

Тип III (гамма-распределение)

6.9 Пирсон различает еще девять других типов, ряд из которых либо тривиален, либо недостаточно интересен. Мы остановимся на одном из типов, находящих значительное теоретическое применение. Некоторые другие типы можно найти в упражнениях к этой главе. Если в (6.9) $B_2 = 0$, то распределение имеет вид

$$f = k \left(1 + \frac{x}{a} \right)^p e^{-px/a}, \quad -a \leq x \leq \infty. \quad (6.19)$$

Здесь в качестве начала выбрано модальное значение. Если за начало отсчета взять начальную точку распределения и подходящим образом выбрать масштаб, то мы придем к той форме

распределения, которая уже встречалась ранее (примеры 3.6, 3.12):

$$f = \frac{1}{\Gamma(\lambda)} x^{\lambda-1} e^{-x}, \quad \lambda > 0, 0 \leq x < \infty. \quad (6.20)$$

Для этого распределения величина λ бесконечна. Кривая распределения (график плотности) унимодальна, за исключением случаев, когда значение λ меньше или равно единице. В этом случае распределение J-образно. Это распределение известно под названием распределения III типа, или гамма-распределения (в силу того, что функция этого распределения является неполной Γ -функцией).

6.10 Было установлено, что пирсоновские распределения нередко хорошо соответствуют результатам наблюдений. Другое достоинство этих распределений (и, в частности, распределений I и III типов), как мы это увидим в дальнейших главах, состоит в том, что ими можно с хорошей точностью приближать теоретические распределения, зная их моменты. Систематическое изложение техники подбора аппроксимирующих распределений было дано Элдертоном (1938а). Мы ограничимся здесь указанием общих принципов и разберем один пример подгонки, который, по-видимому, является одним из наиболее трудных случаев.

6.11 Все распределения Пирсона определяются своими четырьмя первыми моментами $\mu_1, \mu'_2, \mu'_3, \mu'_4$, за исключением некоторых вырожденных распределений, задаваемых меньшим количеством моментов. Пирсоновский метод подгонки состоит в следующем:

1. Определяются первые четыре момента для эмпирического распределения, отвечающего результатам наблюдений.

2. Вычисляются значения β_1, β_2 и величина λ (см. (6.15)), и, следовательно, определяется тип распределений.

3. Эмпирические моменты приравниваются моментам подходящего распределения, которые выражены в терминах его параметров.

4. Полученные уравнения разрешаются относительно неизвестных параметров и, следовательно, находится искомое распределение.

Следующий пример иллюстрирует этот процесс.

Пример 6.1

В таблице 1.15 дано (в нижней строке) распределение 9440 бобов по их длине. (См. также таблицу 6.1 на стр. 223.) Требуется подогнать одну из кривых Пирсона к этому распре-

делению. С учетом поправок Шеппарда для моментов получим

$$\mu'_1(14,5) = -0,190783898$$

$$\mu_2 = 3,238424951$$

$$\mu_3 = -5,306566352$$

$$\mu_4 = 50,999624044$$

$$\beta_1 = 0,829135838, \quad \sqrt{\beta_1} = -0,910569, \quad \beta_2 = 4,862944362.$$

Для величины x , определенной в (6.15), находим

$$x = \frac{\beta_1(\beta_2 + 3)^2}{4(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)} = \frac{51,262}{84,040}.$$

Это число лежит между 0 и 1, и, следовательно, подходящее распределение должно принадлежать типу IV. Полагая в уравнении (6.18) $\theta = x/a$ и $2m - 2 = r$, мы найдем

$$\mu'_n = k \int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} a^{n+1} \cos^{r-n}\theta \sin^n\theta e^{-v\theta} d\theta.$$

Отсюда, интегрируя по частям ($\cos^{r-n}\theta \sin\theta$ берем в качестве одной из частей), получим

$$\mu'_n = \frac{a}{r-n+1} \{ (n-1) a \mu'_{n-2} - v \mu'_{n-1} \},$$

что является частным случаем (6.3). Следовательно, моменты относительно среднего выражаются формулами:

$$\mu'_1 = -\frac{av}{r},$$

$$\mu_2 = \frac{a^2}{r^2(r-1)}(r^2 + v^2),$$

$$\mu_3 = -\frac{4a^3v(r^2 + v^2)}{r^3(r-1)(r-2)},$$

$$\mu_4 = \frac{3a^4(r^2 + v^2)\{(r+6)(r^2 + v^2) - 8v^2\}}{r^4(r-1)(r-2)(r-3)},$$

откуда

$$r = \frac{6(\beta_2 - \beta_1 - 1)}{2\beta_2 - 3\beta_1 - 6},$$

$$v = \frac{r(r-2)\sqrt{\beta_1}}{\sqrt{\{16(r-1) - \beta_1(r-2)^2\}}},$$

$$a = \sqrt{\left[\frac{\mu_2}{16} \{16(r-1) - \beta_1(r-2)^2\} \right]}.$$

Воспользовавшись значениями β_1 , β_2 и μ_2 , найдем

$$\begin{aligned} r &= 14,69772, & m &= 8,34886, \\ v &= 18,38043, & a &= 4,15949. \end{aligned}$$

Здесь надо сделать замечание относительно знаков. Определение знаков у r и m не вызывает трудностей; a и v положительны, поскольку $\sqrt{\beta_1}$ надо считать отрицательным.

До сих пор все вычисления выполнялись непосредственно. Для вычисления константы k нам нужны таблицы интеграла от плотности. Такие таблицы были созданы Пирсоном *). Вычислив значения k , находим

$$f = 0,395121 \left(1 + \frac{x^2}{17,30134} \right)^{-8,34886} \exp \left(-18,38043 \operatorname{arctg} \frac{x}{4,15949} \right).$$

При некотором терпении можно вычислить и значения ординат плотности распределения и интеграл от плотности в заданных пределах. В таблице 6.1 (третий столбец) приведены эти значения для сравнения их с наблюдаемыми частотами.

6.12 При подгонке кривых Пирсона методом моментов, указанным выше, нужно иметь в виду следующее. Если наша цель (как, например, в страховом деле) состоит в том, чтобы получить математическое выражение для распределения, которое удовлетворительно бы описывало наблюдения, хорошо бы градуировалось и интерполировалось, то подгонка с помощью моментов обычно удовлетворительна. Но этот метод, однако, заслуживает критики, когда наблюдаемые данные рассматриваются как выборки из некоторой генеральной совокупности, и нам желательно найти математическое представление этой совокупности. В таких случаях моменты, подсчитанные по наблюдениям, являются только оценками моментов совокупности и, вообще говоря, они не являются наиболее эффективными оценками параметров, характеризующих генеральную совокупность. Мы отложим дальнейшее обсуждение этих вопросов до второго тома (глава 17).

6.13 Другие системы кривых можно получить, воспользовавшись разложением плотности в ряды. В математике и физике хорошо известно, что часто функции можно с пользой представить или в виде степенных рядов (ряды Тейлора), или в виде тригонометрических функций (ряды Фурье). Ни одно из этих представлений нельзя признать очень подходящим для разложения плотностей, и поэтому мы рассмотрим другое множество функций с гораздо более обещающими возможностями.

*) Они помещены в первых двух изданиях *Tables for Statisticians*, Part I, но затем в последующих изданиях были опущены.

Полиномы Чебышева — Эрмита

6.14 Пусть, как и в 5.21,

$$\alpha(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

и

$$D = \frac{d}{dx}.$$

Рассмотрим последовательные производные $\alpha(x)$ по x . Имеем

$$\begin{aligned} D\alpha(x) &= -x\alpha(x), \\ D^2\alpha(x) &= (x^2 - 1)\alpha(x), \\ D^3\alpha(x) &= (3x - x^3)\alpha(x) \end{aligned}$$

и т. д. Очевидно, результаты дифференцирования будут представлять собой полиномы по x , умноженные на $\alpha(x)$. Определим полиномы Чебышева — Эрмита $H_r(x)$ тождеством

$$(-D)^r \alpha(x) = H_r(x) \alpha(x). \quad (6.21)$$

Ясно, что $H_r(x)$ является полиномом степени r с единичным коэффициентом при x^r . Условимся также считать $H_0 = 1$. Мы имеем

$$\alpha(x-t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2 + tx - \frac{1}{2}t^2\right) = \alpha(x) \exp\left(tx - \frac{1}{2}t^2\right),$$

причем в силу теоремы Тейлора

$$\alpha(x-t) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} t^j D^j \alpha(x) = \sum_{j=0}^{\infty} \frac{t^j}{j!} H_j(x) \alpha(x).$$

Следовательно, $H_r(x)$ есть коэффициент при $t^r/r!$ в разложении $\exp\left(tx - \frac{1}{2}t^2\right)$. Таким образом,

$$H_r(x) = x^r - \frac{r^{[2]}}{2 \cdot 1!} x^{r-2} + \frac{r^{[4]}}{2^2 \cdot 2!} x^{r-4} - \frac{r^{[6]}}{2^3 \cdot 3!} x^{r-6} + \dots \quad (6.22)$$

Первые десять полиномов имеют следующий вид:

$$\left. \begin{aligned} H_0 &= 1, \\ H_1 &= x, \\ H_2 &= x^2 - 1, \\ H_3 &= x^3 - 3x, \\ H_4 &= x^4 - 6x^2 + 3, \end{aligned} \right\} \quad (6.23)$$

$$\left. \begin{aligned} H_5 &= x^5 - 10x^3 + 15x, \\ H_6 &= x^6 - 15x^4 + 45x^2 - 15, \\ H_7 &= x^7 - 21x^5 + 105x^3 - 105x, \\ H_8 &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105, \\ H_9 &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x, \\ H_{10} &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945. \end{aligned} \right\} \quad (6.23)$$

6.15 Полиномы Чебышева — Эрмита обладают рядом интересных свойств. Дифференцируя тождество

$$\exp\left(tx - \frac{1}{2}t^2\right) = \sum_{j=0}^{\infty} \frac{t^j H_j(x)}{j!}$$

по x и приравнявая соответствующие коэффициенты при t^r , получаем

$$\frac{d}{dx} H_r(x) = r H_{r-1}(x), \quad (6.24)$$

и вообще

$$D^j H_r(x) = r^{(j)} H_{r-j}(x). \quad (6.25)$$

Дифференцируя то же самое тождество по t и приравнявая коэффициенты при t^{r-1} , получим

$$H_r(x) - x H_{r-1}(x) + (r-1) H_{r-2}(x) = 0. \quad (6.26)$$

Из (6.24) и (6.26) находим

$$\frac{d^2 H_r(x)}{dx^2} - x \frac{dH_r(x)}{dx} + r H_r(x) = 0. \quad (6.27)$$

Известно также (см. Шарлье, 1931), что уравнение (относительно x) $H_r(x) = 0$ имеет r действительных корней, каждое из которых не больше по абсолютной величине, чем $\sqrt{\frac{1}{2}r(r-1)}$.

Таблицы значений первых шести полиномов для $x=0(0,01)4$ были даны Йоргенсенем (1916).

6.16 Полиномы Чебышева — Эрмита обладают важным свойством ортогональности:

$$\int_{-\infty}^{\infty} H_m(x) H_n(x) \alpha(x) dx = \begin{cases} 0, & m \neq n, \\ n!, & m = n. \end{cases} \quad (6.28)$$

В самом деле, интегрирование по частям при $m \leq n$ приводит к выражению

$$\begin{aligned} \int_{-\infty}^{\infty} H_m H_n \alpha dx &= (-1)^n \int_{-\infty}^{\infty} H_m D^n \alpha dx = \\ &= (-1)^n [H_m D^{n-1} \alpha]_{-\infty}^{\infty} + (-1)^{n-1} \int_{-\infty}^{\infty} \frac{dH_m}{dx} D^{n-1} \alpha dx. \end{aligned}$$

Член в квадратных скобках обращается в нуль, и в силу (6.24) интеграл становится равным

$$m(-1)^{n-1} \int_{-\infty}^{\infty} H_{m-1} D^{n-1} \alpha dx.$$

Продолжая этот процесс далее, мы получим нуль, если $m \neq n$, и $m!$, если $m = n$.

Ряды Грама — Шарлье типа А

6.17 Предположим, что плотность распределения может быть формально разложена в ряд по производным от $\alpha(x)$ (ниже мы рассмотрим условия, при которых подобное разложение законно). Тогда

$$f(x) = \sum_{j=0}^{\infty} c_j H_j(x) \alpha(x).$$

Умножая $f(x)$ на $H_r(x)$ и интегрируя в пределах от $-\infty$ до ∞ , мы получим в силу условия ортогональности (6.28)

$$c_r = \frac{1}{r!} \int_{-\infty}^{\infty} f(x) H_r(x) dx. \quad (6.29)$$

Читатель, знакомый с гармоническим анализом, легко усмотрит здесь аналогию между вычислением постоянных c_r и способом вычисления коэффициентов в рядах Фурье. Подставляя в (6.29) точное выражение для $H_r(x)$, даваемое (6.22), найдем

$$c_r = \frac{1}{r!} \left\{ \mu'_r - \frac{r^{[2]}}{2 \cdot 1!} \mu'_{r-2} + \frac{r^{[4]}}{2^2 \cdot 2!} \mu'_{r-4} - \dots \right\}. \quad (6.30)$$

В частности, для моментов относительно среднего значения имеем

$$\left. \begin{aligned} c_0 &= 1, & c_5 &= \frac{1}{120} (\mu_5 - 10\mu_3), \\ c_1 &= 0, & c_6 &= \frac{1}{720} (\mu_6 - 15\mu_4 + 45\mu_2 - 15), \\ c_2 &= \frac{1}{2} (\mu_2 - 1), & c_7 &= \frac{1}{5040} (\mu_7 - 21\mu_5 + 105\mu_3), \\ c_3 &= \frac{1}{6} \mu_3, & c_8 &= \frac{1}{40320} (\mu_8 - 28\mu_6 + 210\mu_4 - \\ & & & - 420\mu_2 + 105). \\ c_4 &= \frac{1}{24} (\mu_4 - 6\mu_2 + 3), \end{aligned} \right\} \quad (6.31)$$

Тем самым формальное разложение дает

$$f(x) = \alpha(x) \left\{ 1 + \frac{1}{2} (\mu_2 - 1) H_2 + \frac{1}{6} \mu_3 H_3 + \right. \\ \left. + \frac{1}{24} (\mu_4 - 6\mu_2 + 3) H_4 + \dots \right\}. \quad (6.32)$$

Если $f(x)$ — плотность нормированного распределения, то

$$f(x) = \alpha(x) \left\{ 1 + \frac{1}{6} \mu_3 H_3 + \frac{1}{24} (\mu_4 - 3) H_4 + \dots \right\}. \quad (6.33)$$

Это разложение носит название ряда Грама — Шарлье типа А.

Эджвортова форма рядов типа А

6.18 Рассмотрим преобразование Фурье функции $H_r(x) \alpha(x)$. Поскольку

$$\sqrt{2\pi} \alpha(t) = e^{-\frac{1}{2} t^2} = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2} dx,$$

то

$$\sqrt{2\pi} \frac{d^r}{dt^r} \alpha(t) = (-1)^r \sqrt{2\pi} H_r(t) \alpha(t) = \int_{-\infty}^{\infty} i^r x^r \frac{e^{itx}}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2} dx,$$

и поэтому преобразование Фурье произведения $x^r \alpha(x)$ есть $i^r \sqrt{2\pi} H_r(t) \alpha(t)$. Обратно

$$x^r \alpha(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} i^r \sqrt{2\pi} H_r(t) \alpha(t) dt.$$

Меняя x и t местами, найдем

$$\sqrt{2\pi} (-i)^r t^r \alpha(t) = \int_{-\infty}^{\infty} e^{-ixt} H_r(x) \alpha(x) dx,$$

и, следовательно, меняя знак у t , находим, что преобразованием произведения $H_r(x) \alpha(x)$ является $\sqrt{2\pi} i^r t^r \alpha(t)$.

Рассмотрим теперь выражение

$$\{\exp(\kappa_r D^r)\} \alpha(x). \quad (6.34)$$

Его характеристическая функция имеет вид

$$\begin{aligned} \int_{-\infty}^{\infty} e^{itx} \exp(\kappa_r D^r) \alpha(x) dx &= \int_{-\infty}^{\infty} e^{itx} \sum \left(\frac{\kappa_r^j D^{rj}}{j!} \right) \alpha(x) dx = \\ &= \sum \frac{\kappa_r^j}{j!} \int_{-\infty}^{\infty} e^{itx} D^{rj} \alpha(x) dx = \sum \frac{\kappa_r^j}{j!} \int_{-\infty}^{\infty} e^{itx} (-1)^{rj} H_{rj}(x) \alpha(x) dx = \\ &= \sum \frac{\kappa_r^j}{j!} \sqrt{2\pi} (-i)^{rj} t^{rj} \alpha(t) = \sqrt{2\pi} \alpha(t) \exp\{\kappa_r (-it)^r\}. \end{aligned} \quad (6.35)$$

Аналогично можно получить, что х. ф. выражения

$$\exp\left\{-\frac{\kappa_1 - a}{1!} D + \frac{\kappa_2 - b}{2!} D^2 - \frac{\kappa_3}{3!} D^3 + \frac{\kappa_4}{4!} D^4 + \dots\right\} \alpha(x) \quad (6.36)$$

равна

$$\begin{aligned} \sqrt{2\pi} \alpha(t) \exp\left\{\frac{\kappa_1 - a}{1!} it + \frac{\kappa_2 - b}{2!} (it)^2 + \right. \\ \left. + \frac{\kappa_3}{3!} (it)^3 + \frac{\kappa_4}{4!} (it)^4 + \dots\right\}. \end{aligned} \quad (6.37)$$

Вообще, если теперь $\alpha(x)$ заменить ненормированной функцией

$$\beta(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}} = \alpha\{(x-m)/\sigma\},$$

то характеристическая функция выражения

$$\exp\left\{-\frac{\kappa_1 - a}{1!} D + \frac{\kappa_2 - b}{2!} D^2 - \frac{\kappa_3}{3!} D^3 + \frac{\kappa_4}{4!} D^4 \dots\right\} \beta(x) \quad (6.38)$$

будет равна

$$\begin{aligned} \sqrt{2\pi} \alpha(t\sigma) e^{itmt} \exp\left\{\frac{\kappa_1 - a}{1!} it + \frac{\kappa_2 - b}{2!} (it)^2 + \right. \\ \left. + \frac{\kappa_3}{3!} (it)^3 + \frac{\kappa_4}{4!} (it)^4 \dots\right\}, \end{aligned} \quad (6.39)$$

в чем нетрудно убедиться с помощью аналогичных рассуждений.

Предположим, что (6.38) представляет собой плотность некоторого распределения. Соответствующая производящая функция семиринвариантов есть логарифм от выражения (6.39), она равна

$$\frac{(\kappa_1 - a + m) it}{1!} + \frac{\kappa_2 - b + \sigma^2}{2!} (it)^2 + \frac{\kappa_3}{3!} (it)^3 + \frac{\kappa_4}{4!} (it)^4 + \dots,$$

и поэтому семинварианты суть $\kappa_1 = a + m$, $\kappa_2 = b + \sigma^2$, $\kappa_3, \kappa_4, \dots, \kappa_r, \dots$. Положив $a = m$ и $b = \sigma^2$, мы получим распределение с семинвариантами $\kappa_1, \kappa_2, \kappa_3, \dots$. Если эти величины на самом деле являются семинвариантами некоторого распределения, то сумма ряда (6.38) должна быть плотностью этого распределения, лишь бы: (1) ряд сходил к некоторой функции плотности и (2) это распределение однозначно определялось бы своими моментами.

Если выражать плотность распределения в нормированной форме, то тогда $\kappa_1 = 0$, $\kappa_2 = 1$ и выражение (6.38) равно

$$\exp \left\{ -\kappa_3 \frac{D^3}{3!} + \kappa_4 \frac{D^4}{4!} \dots \right\} \alpha(x) = f(x), \quad (6.40)$$

где функция $\beta(x)$ заменена на $\alpha(x)$, поскольку $m = 0$ и $\sigma^2 = 1$.

Такой ряд был получен впервые Эджвортом (1904) совершенно другим способом, исходя из теории элементарных ошибок. Уравнение (6.40) формально идентично с (6.33), и читатель, знакомый с историей вопроса о представлении плотности в виде рядов, возможно озадачится тем, что Эджворт утверждал, будто бы его ряд отличен от рядов типа А и дает лучшее представление плотности. Объяснение состоит в том, что для практических целей приходится брать только конечное число членов в этих рядах и пренебрегать остальными. Если мы возьмем первые k членов в (6.33), то результат будет вообще отличен от того результата, который мы получим, взяв в экспоненте в (6.40) первые $(k - 1)$ членов. Дело в том (см. пример 6.3), что с точки зрения теории элементарных ошибок члены в (6.33) не стремятся регулярно к нулю, так что, вообще говоря, последующие члены нельзя считать малыми по сравнению с предыдущими.

6.19 Для нормированных распределений соотношения (6.31) в терминах семинвариантов записываются следующим образом:

$$\left. \begin{aligned} c_0 &= 1, & c_1 &= c_2 = 0, \\ c_3 &= \frac{\kappa_3}{6}, \\ c_4 &= \frac{\kappa_4}{24}, \\ c_5 &= \frac{\kappa_5}{120}, \\ c_6 &= \frac{1}{720} (\kappa_6 + 10\kappa_3^2), \\ c_7 &= \frac{1}{5040} (\kappa_7 + 35\kappa_4\kappa_3), \\ c_8 &= \frac{1}{40320} (\kappa_8 + 56\kappa_5\kappa_3 + 35\kappa_4^2). \end{aligned} \right\} \quad (6.41)$$

Подгонка некоторых функций плотности к распределению длины бобов (таблица 1.15)
(по Преториусу (Pretorius, 1930) и Джонсону (Johnson, 1949a))

Длина бобов (мм) (1)	Наблюдаемые частоты (2)	IV тип Пирсона (3)	Тип А Грама — Шарлье (три члена) (4)	Тип А Грама — Шарлье (четыре члена) (5)	Тип А Грама — Шарлье (пять членов) (6)	Логнормальное распределение (7)	Тип <i>SU</i> Джонсона (8)
> 17,25	—	1,4	163	—15,2	2,0	—	0,1
17,0	6	28,5	12,8	13,7	—35,3	10,1	2,0
16,5	55	299,3	25,6	116,6	22,3	280,5	32,2
16,0	275	1181,6	241,7	370,4	438,1	290,1	290,1
15,5	1129	2132	1012,7	926,2	1214,0	1255,2	1151,5
15,0	2082	2229,8	2155,4	1833,0	1866,9	2163,8	2130,3
14,5	2294	1788,4	2593,0	2506,4	2112,8	2179,6	2240,6
14,0	1787	968,9	713,4	2082,6	1916,7	1598,5	1642,5
13,5	929	503,6	280,7	921,3	1183,4	965,0	970,6
13,0	437	243,7	258,7	132,1	371,2	515,3	508,7
12,5	199	113,8	206,2	178,1	66,9	254,5	249,3
12,0	115	52,5	98,7	117,0	101,2	119,5	118,0
11,5	70	24,2	29,6	43,5	107,1	54,4	55,2
11,0	36	11,3	5,9	10,0	54,0	24,3	25,7
10,5	18	5,4	—	—	15,4	10,7	12,1
10,0	7	1,9	—	—	—	4,8	5,8
9,5	1	—	—	—	3,3	2,1	2,7
< 9,25	—	—	—	—	—	1,7	2,6
Всего	9 440	9 440	9 440	9 440	9 440	9 440	9 440

Примечание: Фигурные скобки означают, что указанные частоты представляют собой суммы малых частот, отвечающих строкам, охватываемым скобками.

6.20 Практически в представлении плотности в виде ряда типа А в расчет приходится принимать только несколько первых членов. Слагаемые в $H_r(x)$ имеют коэффициенты, зависящие от μ_r , и надо иметь в виду, что в силу выборочных флуктуаций при $r > 4$ они определяются достаточно плохо. Если выборочными эффектами можно пренебрегать, то в рассматриваемых рядах следует брать больше членов. Но обычно нет надобности привлекать члены старше H_6 . Мы будем поэтому изучать вопрос о том, насколько хорошо плотность распределения можно представить в виде

$$\alpha(x) \left(1 + \frac{\kappa_3}{6} H_3 + \frac{\kappa_4}{24} H_4 + \frac{\kappa_5}{120} H_5 + \frac{\kappa_6 + 10\kappa_3^2}{720} H_6 \right), \quad (6.42)$$

в надежде на то, что роль отброшенных членов невелика. Попутно заметим, что в противоположность системе Пирсона может быть легко получена функция распределения для плотностей, представленных в виде рассмотренных рядов. В самом деле, если

$$f(x) = \sum a_r H_r(x) \alpha(x),$$

то

$$\int_{-\infty}^x f(x) dx = \sum a_r \int_{-\infty}^x H_r(x) \alpha(x) dx = - \sum a_r H_{r-1}(x) \alpha(x). \quad (6.43)$$

Пример 6.2

Рассмотрим подгонку ряда типа А к распределению данных о бобах по их длине из примера 6.1. Мы уже нашли первые четыре момента. Для нормированного распределения имеем

$$\mu_3 = -0,910569, \quad \mu_4 = 4,862944,$$

а также

$$\mu_5 = -12,574125, \quad \mu_6 = 53,221083.$$

Следовательно, искомое приближение имеет следующий вид:

$$9440\alpha(x) \{ 1 - 0,151762H_3 + 0,0776227H_4 - \\ - 0,0289036H_5 + 0,0142735H_6 \}.$$

Столбцы (4), (5) и (6) в таблице 6.1 показывают частоты, полученные в том случае, когда брались первые 3, 4 и 5 членов этого ряда. При просмотре чисел, полученных с учетом четырех и пяти членов, видно, что они дают худшие приближения, чем ряд с учетом только трех членов. К тому же ряд с пятью членами дает отрицательные частоты на одном из концов и моду, рав-

ную 12 мм, что противоречит исходным данным. Таким образом, представление с помощью рядов типа A не вполне удовлетворительно и не лучше того представления, которое дают кривые Пирсона IV типа.

Тетрахорические функции

6.21. Численные значения таких функций, как $H_r(x)\alpha(x)$, можно находить:

а) непосредственно, путем вычисления значений полинома и умножения на $\alpha(x)$;

б) с помощью таблиц $H(x)$ и $\alpha(x)$;

в) из таблиц так называемых эрмитовых вероятностных функций с отрицательным индексом (British Association Mathematical Tables, vol. 1, 1946). Эти последние дают значения функ-

ций $Hh_n(x) \equiv (-D)^{n-1} e^{-\frac{1}{2}x^2}$ с десятью знаками для n от 0 до 7 и $x = -7(0,1)6,5$. Заметим, что здесь опущен множитель $1/\sqrt{2\pi}$. Аналогичная функция была табулирована К. Пирсоном в форме

$$\tau_r(x) = \frac{H_{r-1}(x)\alpha(x)}{(r!)^{1/2}}. \quad (6.44)$$

Она известна под названием тетрахорической функции *) по причинам, которые станут ясными из второго тома при рассмотрении оценки коэффициента корреляции в таблицах 2×2 .

6.22 До сих пор предполагалось, что плотность обладает сходящимся рядом типа A . Мы не будем здесь вдаваться в дискуссию о том, при каких условиях это в действительности так. Ограничимся только тем, что предостережем читателя от ошибок, которые здесь можно сделать, и приведем без доказательства некоторые теоремы.

(1) Крамёр (1926). Если $f(x)$ есть функция с непрерывной производной такой, что

$$\int_{-\infty}^{\infty} \left(\frac{df}{dx} \right)^2 e^{\frac{1}{2}x^2} dx < \infty,$$

и если $f(x)$ стремится к нулю при $|x| \rightarrow \infty$, тогда $f(x)$ можно представить в виде ряда

$$f(x) = \sum_{j=0}^{\infty} \frac{c_j}{j!} D^j \alpha(x), \quad (6.45)$$

*) Семизначные таблицы тетрахорических функций $\tau_r(x)$ для $r=0(1)21$ и $x=0(0,002)4$ опубликованы в сборнике «Таблицы нормального интеграла вероятностей, нормальной плотности и ее нормированных производных» (под ред. Н. В. Смирнова), М., Изд-во АН СССР, 1960. (Прим. ред.)

где

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

Этот ряд сходится абсолютно и равномерно для $-\infty \leq x \leq \infty$.

(2) Другая теорема Крамёра основана на одном результате Гэлбрапа. Если $f(x)$ имеет ограниченную вариацию на каждом конечном интервале и если

$$\int_{-\infty}^{\infty} |f(x)| e^{\frac{1}{4}x^2} dx$$

существует, тогда ряд (6.45) сходится всюду к сумме $\frac{1}{2}\{f(x+0) + f(x-0)\}$. Сходимость равномерна на каждом конечном интервале непрерывности $f(x)$.

Крамёр также показал, что эта последняя теорема не может быть существенно улучшена за счет уточнения поведения $f(x)$ на бесконечности. В самом деле, рассмотрим функцию $f(x) = e^{-\lambda x^2}$. В силу (6.26) и (6.24) имеем

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-\lambda x^2} H_r(x) dx &= \int_{-\infty}^{\infty} e^{-\lambda x^2} x H_{r-1} dx - (r-1) \int_{-\infty}^{\infty} e^{-\lambda x^2} H_{r-2} dx = \\ &= \left[\frac{e^{-\lambda x^2}}{-2\lambda} H_{r-1} \right]_{-\infty}^{\infty} + \frac{r-1}{2\lambda} \int_{-\infty}^{\infty} e^{-\lambda x^2} H_{r-2} dx - (r-1) \int_{-\infty}^{\infty} e^{-\lambda x^2} H_{r-2} dx = \\ &= (r-1) \left(\frac{1}{2\lambda} - 1 \right) \int_{-\infty}^{\infty} e^{-\lambda x^2} H_{r-2}(x) dx. \end{aligned}$$

Если r нечетно, то интеграл обращается в нуль, так как H_r — нечетная функция. Если r четно, скажем $2r$, то интеграл равен

$$(2r-1)(2r-3) \dots 1 \left(\frac{1}{2\lambda} - 1 \right)^r \sqrt{\frac{\pi}{\lambda}} = \sqrt{\frac{\pi}{\lambda}} \frac{(2r)!}{2^r r!} \left(\frac{1}{2\lambda} - 1 \right)^r.$$

Соответствующий коэффициент при H_{2r} в ряде типа А равен $\left(\frac{1}{2\lambda} - 1 \right)^r / (2^r r!)$. В точке $x=0$ $H_{2r} = (-1)^r (2r)! / (2^r r!)$ и ряд имеет вид

$$\sum_{r=0}^{\infty} \sqrt{\frac{\pi}{\lambda}} \frac{(2r)!}{2^{2r} (r!)^2} \left(1 - \frac{1}{2\lambda} \right)^r.$$

В силу формулы Стирлинга r -й член этого ряда (скажем u_r) эквивалентен

$$\left(1 - \frac{1}{2\lambda}\right)^r \frac{1}{V\lambda r},$$

так что

$$u_r^{1/r} \sim 1 - \frac{1}{2\lambda}.$$

Следовательно, при $\lambda < 1/4$ ряд расходится.

6.23 С точки зрения статистики, однако, более важен вопрос не о том, когда бесконечный ряд дает представление для функции плотности, а когда с помощью конечного числа членов можно получить удовлетворительную аппроксимацию. Ясно, что:

а) сумма конечного числа членов таких рядов может привести к отрицательным значениям частот, в особенности на «хвостах» распределения (как в примере 6.2);

(б) ряды Шарлье в форме (6.33) могут вести себя нерегулярно в том смысле, что сумма k членов может давать худшее приближение, чем сумма $(k - 1)$ членов.

Бартон и Деннис (Barton and Dennis) (1952, *Biom.* **39**, 425) показали, что если пренебрегать семиинвариантами выше четвертого порядка, то представления плотностей формулами (6.33) и, в особенности, (6.42) могут оказаться неунимодальными и даже отрицательными, если только коэффициент асимметрии β_1 недостаточно близок к нулю. Например, если $|\beta_1| \geq 0,25$, то (6.42) определяет заведомо неунимодальную функцию; при $|\beta_1| \geq 0,5$ эта функция наверняка принимает отрицательные значения; ряды (6.33) приемлемы в более широкой области, но и они дают заведомо неунимодальные функции при $|\beta_1| \geq 0,7$.

Использование конечных рядов приводит к успеху только в случае распределений с умеренной асимметрией. Для таких распределений, как распределение Пирсона, и некоторых других, которые мы рассмотрим позже, подобная аппроксимация может быть довольно хорошей. Во многих статистических задачах значительный интерес представляет поведение распределения на его хвостах. Надо иметь в виду, что аппроксимация хвостов конечными рядами может оказаться крайне неудовлетворительной.

Пример 6.3

В качестве иллюстрации нерегулярного поведения членов ряда типа А рассмотрим аппроксимацию гамма-распределения:

$$dF = \frac{1}{\Gamma(\lambda)} x^{\lambda-1} e^{-x} dx, \quad 0 \leq x < \infty, \quad \lambda > 0.$$

Его характеристическая функция (см. пример 3.6) выражается

формулой

$$\varphi(t) = \frac{1}{(1-it)^\lambda}$$

и поэтому $\kappa_r = \lambda(r-1)!$, а для распределения в нормированной форме

$$\kappa_r = \frac{(r-1)!}{\lambda^{\frac{1}{2}r-1}}.$$

Из способа образования членов в (6.41) видно, что коэффициент c_r есть сумма членов $\kappa_r, \kappa_{r-3}\kappa_3, \dots (\kappa_{q_1} \dots \kappa_{q_m})$, где $(q_1, \dots, q_m)$ образуют такое разбиение числа r , в котором нет q , меньших чем 3. Ясно, что поскольку κ_q — величина порядка $\lambda^{1-\frac{1}{2}q}$, то наибольший порядок относительно λ имеет тот член вида $\kappa_{q_1} \dots \kappa_{q_m}$, который содержит наибольшее число элементов. Например, если $r=9$, то наибольший порядок имеет член $(3^3) = (\kappa_3\kappa_3\kappa_3)$, если $r=8$, то член $(4^2) = (\kappa_4\kappa_4)$, и т. д.

Эти рассуждения помогают определить порядок членов ряда типа А относительно λ . Так,

член	...	c_0	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
порядок относительно λ		0	$-\frac{1}{2}$	-1	$-1\frac{1}{2}$	-1	$-1\frac{1}{2}$	-2	$-1\frac{1}{2}$	-2	$-2\frac{1}{2}$	-2

С ростом порядка малости относительно λ члены убывают, но нерегулярно, и поэтому здесь не приходится говорить о малости последующих членов по сравнению с предыдущими. Асимптотические качества таких рядов, очевидно, требуют более подробного изучения в каждом отдельном случае.

Ряды типа В

6.24 Шарлье предложил также ряд типа В, основанный на производных пуассоновской частоты $e^{-\lambda}\lambda^x/x!$ по λ , или, что то же самое, на первых разностях по x . Этот подход имеет некоторую математическую привлекательность и был использован для аппроксимации распределений, похожих на пуассоновское. Однако большого распространения он не получил. Другая форма (тип С), предложенная Шарлье с целью избежать отрицательных частот, также не выдержала испытаний практикой *).

*) Можно также получить разложения в терминах производных от гамма-и бета-распределений. Соответствующими полиномами являются полиномы Лагерра и Якоби, см. Г. Сеге, Ортогональные многочлены, М., Физматгиз,

Нормализация функций плотности

6.25 Некоторые важные теоретические распределения, возникающие в статистике, оказываются зависимыми от некоторой переменной n таким образом, что когда n стремится к бесконечности, эти распределения стремятся к нормальному. Для больших n нормальная аппроксимация часто оказывается вполне удовлетворительной, но для малых или умеренных n она может быть несостоятельной. Тем не менее и в этом случае мы можем постараться использовать нормальное приближение, рассматривая преобразованный вариант

$$\xi = b_0 + b_1x + b_2x^2 + b_3x^3 + \dots, \quad (6.46)$$

где b_n имеет порядок $n^{-1/2}$ или меньший. Выбирая подходящим образом b_n , мы можем добиться, чтобы распределение ξ было ближе к нормальному, чем распределение x . Более того, зная распределение ξ (оно предполагается нормальным) и используя формулу (6.46), можно найти само распределение x .

Рассмотрим, в самом деле, эджвортово разложение (6.38) типа A

$$\exp \left\{ -\frac{\kappa_1 - m}{1!} D + \frac{\kappa_2 - \sigma^2}{2!} D^2 - \frac{\kappa_3}{3!} D^3 + \dots \right\} \beta(x). \quad (6.47)$$

Мы сохранили коэффициенты при D и D^2 , поскольку аппроксимация может оказаться лучшей, если брать у ξ среднее m и дисперсию σ^2 , отличающиеся от среднего и дисперсии у x .

Предположим теперь, что семинварианты κ_r имеют порядок n^{1-r} (этот случай встречается весьма часто). Выберем m и σ^2 таким образом, чтобы выполнялись условия

$$\kappa_1 - m = l_1 \sigma, \quad l_1 = O(n^{-1/2}),$$

$$\kappa_2 - \sigma^2 = l_2 \sigma^2, \quad l_2 = O(n^{-1}).$$

Тогда σ^2 имеет порядок κ_2 , т. е. n^{-1} , и, следовательно,

$$\frac{\kappa_r}{\sigma^r} = O(n^{1-r/2}).$$

Тем самым (6.47) можно представить как

$$\exp \left\{ -l_1 \sigma D + \frac{1}{2} l_2 \sigma^2 D^2 - \frac{1}{6} l_3 \sigma^3 D^3 + \right. \\ \left. + \frac{1}{24} l_4 \sigma^4 D^4 - \frac{1}{120} l_5 \sigma^5 D^5 + \frac{1}{720} l_6 \sigma^6 D^6 - \dots \right\} \beta(x), \quad (6.48)$$

1962; Дурбин и Ватсон (Durbin and Watson) (1951, Biom. 38, 159, в особенности стр. 172). Уоллес (Wallace) (1958, Ann. Math. Statist. 29, 635) изучал вопрос о справедливости общих асимптотических разложений (в частности, по отличию от нормальных плотностей) и проблеме о нормализации, обсуждаемую в 6.25—6.26. [О разложениях в терминах производных от персоновских плотностей см. Большев (1963), Теория вероят. и ее примен. 8, 129. (Прим. ред.)]

где l_1 и l_3 порядка $O(n^{-1/2})$, l_5 порядка $O(n^{-3/2})$, l_2 и l_4 порядка $O(n^{-1})$, l_6 порядка $O(n^{-2})$ и т. д.

Разлагая оператор по степеням D и сохраняя только члены, содержащие l до порядка $O(n^{-2})$ включительно, мы получим оператор

$$\begin{aligned}
 & 1 - l_1 \sigma D + \frac{1}{2} l_2 \sigma^2 D^2 - \frac{1}{6} l_3 \sigma^3 D^3 + \frac{1}{24} l_4 \sigma^4 D^4 - \frac{1}{120} l_5 \sigma^5 D^5 + \\
 & + \frac{1}{720} l_6 \sigma^6 D^6 + \frac{1}{2} \left(l_1^2 \sigma^2 D^2 + \frac{1}{4} l_2^2 \sigma^4 D^4 + \frac{1}{36} l_3^2 \sigma^6 D^6 + \frac{1}{576} l_4^2 \sigma^8 D^8 - \right. \\
 & - l_1 l_2 \sigma^3 D^3 + \frac{1}{3} l_1 l_3 \sigma^4 D^4 - \frac{1}{12} l_1 l_4 \sigma^5 D^5 + \frac{1}{60} l_1 l_5 \sigma^6 D^6 - \frac{1}{6} l_2 l_3 \sigma^5 D^5 + \\
 & + \frac{1}{24} l_2 l_4 \sigma^6 D^6 - \frac{1}{72} l_3 l_4 \sigma^7 D^7 + \frac{1}{360} l_3 l_5 \sigma^8 D^8 \left. \right) + \\
 & + \frac{1}{6} \left(-l_1^3 \sigma^3 D^3 - \frac{1}{216} l_3^3 \sigma^9 D^9 + \frac{3}{2} l_1^2 l_2 \sigma^4 D^4 - \frac{1}{2} l_1^2 l_3 \sigma^5 D^5 - \frac{1}{8} l_1^2 l_4 \sigma^6 D^6 + \right. \\
 & + \frac{1}{288} l_3^2 l_4 \sigma^{10} D^{10} - \frac{1}{12} l_1 l_3^2 \sigma^7 D^7 + \frac{1}{24} l_2 l_3^2 \sigma^8 D^8 + \frac{1}{2} l_1 l_2 l_3 \sigma^6 D^6 + \\
 & + \frac{1}{24} l_1 l_3 l_4 \sigma^8 D^8 \left. \right) + \frac{1}{24} \left(l_1^4 \sigma^4 D^4 + \frac{1}{1296} l_3^4 \sigma^{12} D^{12} + \frac{2}{3} l_1^3 l_3 \sigma^6 D^6 + \right. \\
 & \left. + \frac{1}{6} l_1^2 l_3^2 \sigma^8 D^8 + \frac{1}{54} l_1 l_3^3 \sigma^{10} D^{10} \right). \quad (6.49)
 \end{aligned}$$

Результат действия этого оператора в применении к $\beta(x)$ будет тем же самым, как если в выражении (6.49) оператор $\sigma^r D^r$ заменить на функцию $(-1)^r H_r \left(\frac{x-m}{\sigma} \right)$ и результат умножить на $\beta(x)$.

Функция распределения получается путем интегрирования этого выражения; ее значение в точке $m + \sigma x$ равно (члены расположены в соответствии с возрастанием их порядка малости относительно n)

$$\begin{aligned}
 & \int_{-\infty}^x \alpha(x) dx + \alpha(x) \left[- \left(l_1 + \frac{1}{6} l_3 H_2 \right) - \left(\frac{1}{2} l_1^2 H_1 + \frac{1}{2} l_2 H_1 + \frac{1}{6} l_1 l_3 H_3 + \right. \right. \\
 & + \frac{1}{24} l_4 H_3 + \frac{1}{72} l_3^2 H_5 \left. \right) - \left(\frac{1}{6} l_1^3 H_2 + \frac{1}{2} l_1 l_2 H_2 + \frac{1}{12} l_1^2 l_3 H_4 + \right. \\
 & + \frac{1}{12} l_2 l_3 H_4 + \frac{1}{24} l_1 l_4 H_4 + \frac{1}{120} l_5 H_4 + \frac{1}{72} l_1 l_3^2 H_6 + \frac{1}{144} l_3 l_4 H_6 + \\
 & + \frac{1}{1296} l_3^3 H_8 \left. \right) - \left(\frac{1}{24} l_1^4 H_3 + \frac{1}{8} l_2^2 H_3 + \frac{1}{4} l_1^2 l_2 H_3 + \frac{1}{36} l_1^3 l_3 H_5 + \right. \\
 & + \frac{1}{12} l_1 l_2 l_3 H_5 + \frac{1}{48} l_1^2 l_4 H_5 + \frac{1}{48} l_2 l_4 H_5 + \frac{1}{120} l_1 l_5 H_5 + \frac{1}{720} l_6 H_5 +
 \end{aligned}$$

$$+ \frac{1}{144} l_1^2 l_3^2 H_7 + \frac{1}{144} l_2 l_3^2 H_7 + \frac{1}{1152} l_4^2 H_7 + \frac{1}{144} l_1 l_3 l_4 H_7 + \\ + \frac{1}{720} l_3 l_5 H_7 + \frac{1}{1296} l_1 l_3^3 H_9 + \frac{1}{1728} l_3^2 l_4 H_9 + \frac{1}{31 \cdot 104} l_3^4 H_{11} \Big]. \quad (6.50)$$

6.26 Пусть ξ — нормальная варианта. Определим ξ как функцию от x таким образом, что

$$G(\xi) = \int_{-\infty}^{\xi} \alpha(y) dy = F(x), \quad (6.51)$$

где функция распределения $F(x)$ задается рядом (6.50) типа А. По формуле Тейлора

$$G(\xi) = G\{x + (\xi - x)\} = G(x) + \sum_{r=1}^{\infty} \frac{(\xi - x)^r}{r!} \frac{d^r}{dx^r} G(x),$$

откуда, применяя (6.21), находим

$$G(\xi) = G(x) - \sum_{r=1}^{\infty} \frac{(x - \xi)^r}{r!} H_{r-1}(x) \alpha(x), \quad (6.52)$$

что в свою очередь равно (6.50).

Следующий шаг состоит в обращении этого ряда с целью получить разложение $(x - \xi)$ по степеням x . Пусть

$$x - \xi = a_0 + a_1 x + a_2 x^2 + \dots \quad (6.53)$$

(это есть не что иное, как выражение (6.46), слегка преобразованное в целях удобства). Мы видим, что если $x=0$, то $\xi = -a_0$ и $x - \xi$ имеет порядок $n^{-1/2}$; следовательно, с точностью до величины порядка n^{-2} из (6.52) и (6.23) при $x=0$ получаем

$$\xi - \frac{\xi^2}{2} (+0) + \frac{\xi^3}{6} (-1) = - \left(a_0 - \frac{a_0^3}{6} \right),$$

что совпадает с выражением в квадратных скобках в (6.50), когда $x=0$. Отсюда

$$a_0 = l_1 - \frac{1}{6} l_3 - \frac{1}{2} l_1 l_2 + \frac{1}{6} l_1^2 l_3 + \frac{1}{4} l_2 l_3 + \frac{1}{8} l_1 l_4 + \\ + \frac{1}{40} l_5 - \frac{7}{36} l_1 l_3^2 - \frac{15}{144} l_3 l_4 + \frac{52}{648} l_3^3.$$

Коэффициент a_1 в (6.50) мы можем найти, отождествляя коэффициенты при x и т. д. После некоторых алгебраических преобразований, записывая члены в убывающем порядке по n ,

получаем

$$\begin{aligned}
 x - \xi = & l_1 + \frac{1}{6} l_3 (x^2 - 1) + \frac{1}{2} l_2 x - \frac{1}{3} l_1 l_3 x + \frac{1}{24} l_4 (x^3 - 3x) - \\
 & - \frac{1}{36} l_3^2 (4x^3 - 7x) - \frac{1}{2} l_1 l_2 + \frac{1}{6} l_1^2 l_3 - \frac{1}{12} l_2 l_3 (5x^2 - 3) - \\
 & - \frac{1}{8} l_1 l_4 (x^2 - 1) + \frac{1}{120} l_5 (x^4 - 6x^2 + 3) + \frac{1}{36} l_1 l_3^2 (12x^2 - 7) - \\
 & - \frac{1}{144} l_3 l_4 (11x^4 - 42x^2 + 15) + \frac{1}{648} l_3^3 (69x^4 - 187x^2 + 52) - \\
 & - \frac{3}{8} l_2^2 x + \frac{5}{6} l_1 l_2 l_3 x + \frac{1}{8} l_1^2 l_4 x - \frac{1}{48} l_2 l_4 (7x^3 - 15x) - \frac{1}{30} l_1 l_5 (x^3 - 3x) + \\
 & + \frac{1}{720} l_6 (x^5 - 10x^3 + 15x) - \frac{1}{3} l_1^2 l_3^2 x + \frac{1}{72} l_2 l_3^2 (36x^2 - 49x) - \\
 & - \frac{1}{384} l_4^2 (5x^5 - 32x^3 + 35x) + \frac{1}{36} l_1 l_3 l_4 (11x^3 - 21x) - \\
 & - \frac{1}{360} l_3 l_5 (7x^5 - 48x^3 + 51x) - \frac{1}{324} l_1 l_3^3 (138x^3 - 187x) + \\
 & + \frac{1}{864} l_3^2 l_4 (111x^5 - 547x^3 + 456x) - \frac{1}{7776} l_3^4 (948x^5 - 3628x^3 + 2473x).
 \end{aligned} \tag{6.54}$$

Это и есть требуемое представление ξ в терминах x . По крайней мере с точностью до членов порядка n^{-2} варианта ξ будет распределена нормально.

Часто оказывается удобным выразить x в терминах ξ . Это нетрудно сделать, если заметить, что

$$\begin{aligned}
 x - \xi = g(x) &= g(\xi + x - \xi) = g(\xi) + (x - \xi) g'(\xi) + \dots = \\
 &= g(\xi) + g'(\xi) \{g(\xi) + (x - \xi) g'(\xi)\} + \dots
 \end{aligned}$$

Продолжая этот процесс, получим

$$\begin{aligned}
 x - \xi = & g(\xi) + g(\xi) g'(\xi) + g(\xi) g'^2(\xi) + \frac{1}{2} g^2(\xi) g''(\xi) + \\
 & + g(\xi) g'^3(\xi) + \frac{3}{2} g^2(\xi) g'(\xi) g''(\xi) + \frac{1}{6} g^3(\xi) g'''(\xi) + \dots
 \end{aligned} \tag{6.55}$$

Следовательно, используя значение ξ , задаваемое (6.54), после некоторых преобразований мы найдем

$$\begin{aligned}
 x - \xi = & l_1 + \frac{1}{6} l_3 (\xi^2 - 1) + \frac{1}{2} l_2 \xi + \frac{1}{24} l_4 (\xi^3 - 3\xi) - \frac{1}{36} l_3^2 (2\xi^3 - 5\xi) - \\
 & - \frac{1}{6} l_2 l_3 (\xi^2 - 1) + \frac{1}{120} l_5 (\xi^4 - 6\xi^2 + 3) - \frac{1}{24} l_3 l_4 (\xi^4 - 5\xi^2 + 2) +
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{324} l_3^2 (12\xi^4 - 53\xi^2 + 17) - \frac{1}{8} l_2^2 \xi - \frac{1}{16} l_2 l_4 (\xi^3 - 3\xi) + \\
 & + \frac{1}{720} l_6 (\xi^5 - 10\xi^3 + 15\xi) + \frac{1}{72} l_2 l_3^2 (10\xi^3 - 25\xi) - \\
 & - \frac{1}{384} l_4^2 (3\xi^5 - 24\xi^3 + 29\xi) - \frac{1}{180} l_3 l_5 (2\xi^5 - 17\xi^3 + 21\xi) + \\
 & + \frac{1}{288} l_3 l_4 (14\xi^5 - 103\xi^3 + 107\xi) - \frac{1}{7776} l_3^4 (252\xi^5 - 1688\xi^3 + 1511\xi).
 \end{aligned} \tag{6.56}$$

Описанный в этом пункте подход принадлежит Корнишу и Фишеру (1937). Эти же авторы (1960, *Technometrics* **2**, 209) нашли разложения более высоких порядков и создали таблицы для облегчения их использования.

Пример 6.4

Рассмотрим снова распределение из примера 6.3:

$$dF = \frac{1}{\Gamma(\lambda)} e^{-x} x^{\lambda-1} dx, \quad 0 \leq x < \infty.$$

Мы уже нашли, что после нормировки это распределение стремится к нормальному и что κ_r — величина порядка $\lambda^{-\frac{1}{2}r}$.

Положим в (6.48) l_1 и l_2 равными нулю, т. е. будем считать, что ξ и x имеют одинаковые средние и одинаковые дисперсии. Тогда

$$l_3 = 2\lambda^{-1/2}, \quad l_4 = 6\lambda^{-1}, \quad l_5 = 24\lambda^{-3/2}, \quad l_6 = 120\lambda^{-2}.$$

В этом случае формула (6.50) примет вид

$$\begin{aligned}
 G(x) - \alpha(x) \left\{ \frac{1}{3\lambda^{1/2}} H_2 + \frac{1}{4\lambda} H_3 + \frac{1}{18\lambda} H_5 + \frac{1}{5\lambda^{3/2}} H_4 + \frac{1}{12\lambda^{3/2}} H_6 + \right. \\
 \left. + \frac{1}{162\lambda^{3/2}} H_8 + \frac{1}{6\lambda^2} H_5 + \frac{1}{32\lambda^2} H_7 + \frac{1}{15\lambda^2} H_7 + \frac{1}{72\lambda^2} H_9 + \frac{1}{1944\lambda^2} H_{11} \right\}.
 \end{aligned}$$

Теперь в качестве простой иллюстрации найдем значение функции распределения для x при $\lambda=9$ и $x=12$. Среднее значение и дисперсия равны 9, так что отклонение, выраженное в единицах стандартного отклонения, есть $(12-9)/\sqrt{9}=1$.

Из (6.23) с дополнительным равенством для H_{11} находим, что

$$\begin{aligned}
 H_2 = 0, \quad H_3 = -2, \quad H_4 = -2, \quad H_5 = 6, \quad H_6 = 16, \quad H_7 = -20, \\
 H_8 = -132, \quad H_9 = 28, \quad H_{10} = 1216, \quad H_{11} = 936,
 \end{aligned}$$

Тогда искомое значение функции распределения приближенно равно сумме

$$\int_{-\infty}^1 \alpha(x) dx + \alpha(1)(0,0151635).$$

Значение нормальной функции берем из таблиц и находим $0,841345 + (0,2419707)(0,0151635) = 0,8450$

(все четыре десятичных знака совпадают с точными). Аппроксимация, очевидно, достаточно хорошая даже для такого небольшого значения, как $\lambda = 9$.

К этому результату можно прийти с помощью (6.54). Полагая $x=1$, мы найдем $\xi = 1,015386$. Значение функции нормального распределения в этой точке ξ равно 0,8450, что совпадает с предыдущим результатом.

Предположим теперь, что мы желаем найти то значение x , для которого $F(x) = 0,99$ при $\lambda = 15$. Нормальное отклонение ξ , соответствующее такому значению, может быть найдено из таблиц, и оно оказывается равным 2,326348. Согласно (6.56)

$$x - \xi = \frac{1}{3\lambda^{1/2}}(\xi^2 - 1) + \frac{1}{4\lambda}(\xi^3 - 3\xi) + \dots,$$

откуда в единицах стандартного отклонения

$$x = 2,69722.$$

Ненормированное значение x равно

$$15 + x \sqrt{15} = 25,45.$$

Этот результат является правильным с точностью до двух знаков.

Итак, примеры показывают, что в некоторых случаях, несмотря на то, что бесконечные ряды типа А могут и не сходиться, вполне удовлетворительную аппроксимацию уже можно получить с использованием лишь нескольких первых членов. Не приводя доказательств, отметим, что, применяя результаты Крамэра (1928), можно показать, что в рассмотренном примере асимптотическое разложение действительно существует.

Преобразования Джонсона

6.27 Идея преобразования одной варианты к другой, имеющей более простую плотность распределения, может быть развита дальше. Мы замечаем, что ничто вовсе не обязывает нас ограничиться только полиномиальными преобразованиями типа (6.46). Более того, подобного рода преобразования в действительности оказываются не наилучшими. Некоторые авторы изу-

чали более общий тип преобразований. В дальнейшем изложении мы будем следовать Джонсону (1940а).

Рассмотрим преобразование типа

$$\xi = \gamma + \delta g\left(\frac{x - \mu}{\lambda}\right), \quad (6.57)$$

где μ , λ , γ и δ — параметры, подлежащие подбору и g — некоторая подходящая функция. Для практических целей желательно, чтобы функция g сама не зависела от параметров и была бы монотонной; в противном случае это доставляло бы заботы относительно пределов изменения ξ и x . Без потери общности мы можем предположить, что g — монотонно возрастающая функция. Если мы запишем преобразование в форме

$$\frac{\xi - \gamma}{\delta} = g\left(\frac{x - \mu}{\lambda}\right), \quad (6.58)$$

то μ и γ будут играть роль параметров сдвига, а λ и δ — роль масштабных параметров. Последнюю пару мы можем считать положительной. Будем пытаться выбрать эти параметры и функцию g так, чтобы распределение ξ было или точно, или, по крайней мере, приближенно нормальным.

6.28 Без потери общности можно рассматривать только нормированные величины

$$y = \frac{x - \mu}{\lambda}. \quad (6.59)$$

Если ξ нормальна с плотностью вероятности $\alpha(\xi)$, то для распределения y получим уравнение

$$dF(y) = \alpha(\xi) \left| \frac{d\xi}{dy} \right| dy = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \{ \gamma + \delta g(y) \}^2 \right] \delta g'(y) dy, \quad (6.60)$$

где $g'(y) = dg(y)/dy$. Будем исследовать три типа систем:

- 1) система S_L или логарифмически нормальная (логнормальная): $g(y) = \log y$,
- 2) система S_B : $g(y) = \log \{y/(1 - y)\}$,
- 3) система S_U : $g(y) = \operatorname{argsh} y = \log \{y + \sqrt{y^2 + 1}\}$.

Возможны, конечно, и другие системы. Заметим, что многообразие форм кривых, получаемых только из этих трех типов, столь же обширно, как и пирсоновская система кривых.

Логнормальное распределение

6.29 Пусть

$$\xi = \gamma + \delta \log y. \quad (6.61)$$

Когда ξ изменяется от $-\infty$ до $+\infty$, y изменяется от 0 до ∞ . Плотность распределения варианты y имеет бесконечно высокий

порядок соприкосновения с осью абсцисс как в нуле, так и в ∞ ; согласно (6.60)

$$dF(y) = p(y) dy,$$

где

$$p(y) = \frac{\delta}{\sqrt{2\pi} y} \exp \left\{ -\frac{1}{2} (\gamma + \delta \log y)^2 \right\}, \quad y \geq 0.$$

Для каждого положительного r

$$\lim_{y \rightarrow \infty} y^r p(y) = \lim_{y \rightarrow \infty} \frac{\delta e^{-\frac{1}{2} \gamma^2}}{\sqrt{2\pi}} \frac{y^{r-1-\gamma\delta}}{\exp \left\{ \frac{1}{2} (\delta \log y)^2 \right\}} = 0$$

и

$$\lim_{y \rightarrow 0} y^r p(y) = \lim_{y \rightarrow 0} \frac{\delta e^{-\frac{1}{2} \gamma^2}}{\sqrt{2\pi}} \frac{y^{r-1-\gamma\delta}}{y^{1/2} \delta^2 \log y} = 0.$$

Это распределение унимодально, так как производная

$$\frac{dp(y)}{dy} = \frac{\delta}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} (\gamma + \delta \log y)^2 \right\} \left\{ -\frac{1}{y^2} - \frac{\delta}{y^2} (\gamma + \delta \log y) \right\}$$

внутри интервала $0 < y < \infty$ обращается в нуль только в одной точке, определяемой уравнением

$$1 + \delta(\gamma + \delta \log y) = 0.$$

6.30 Моменты величины y , взятые относительно 0, выражаются формулой

$$\begin{aligned} \mu'_r(y) &= \int_0^\infty y^r p(y) dy = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp \left\{ \frac{r(\xi - \gamma)}{\delta} \right\} \exp \left(-\frac{1}{2} \xi^2 \right) d\xi = \exp \left(\frac{1}{2} \frac{r^2}{\delta^2} - \frac{r\gamma}{\delta} \right). \end{aligned} \quad (6.62)$$

Поэтому, полагая $\omega = \exp(1/2\delta^2)$, $\rho = \exp(-\gamma/\delta)$, имеем

$$\left. \begin{aligned} \mu'_1 &= \omega \rho, \\ \mu'_2 &= \omega^2 \rho^2 (\omega^2 - 1), \\ \mu'_3 &= \omega^3 \rho^3 (\omega^2 - 1)^2 (\omega^2 + 2), \\ \mu'_4 &= \omega^4 \rho^4 (\omega^2 - 1)^2 (\omega^8 + 2\omega^6 + 3\omega^4 - 3) \end{aligned} \right\} \quad (6.63)$$

и

$$\beta_1 = (\omega^2 - 1)(\omega^2 + 2), \quad (6.64)$$

$$\beta_2 - 3 = (\omega^2 - 1)(\omega^6 + 3\omega^4 + 6\omega^2 + 6). \quad (6.65)$$

Асимметрия всегда положительна и распределение имеет эксцесс, превышающий нормальный. Функция плотности возрастает от нуля до точки максимума и затем более медленно убывает при $y \rightarrow \infty$.

6.31 Значения μ_2 и μ_3 определяют ω и ρ и, следовательно, δ и γ . Поэтому, отправляясь от параметров сдвига и масштаба, мы можем подгонять логнормальное распределение по второму и третьему моментам x -распределения. Логнормальное распределение не обладает общностью семейства (6.57), так как оно зависит фактически от трех, а не от четырех параметров. Выражение

$$\xi = \gamma + \delta \log \{(x - \mu)/\lambda\}$$

можно представить в виде

$$\xi = (\gamma - \delta \log \lambda) + \delta \log (x - \mu),$$

и, значит, слагаемое $\gamma - \delta \log \lambda$ можно рассматривать как единый параметр. Именно по этой причине при подгонке этого распределения используются только μ_2 и μ_3 и не используется μ_4 . Простейшая процедура подгонки, предложенная Уикселлом (1917), состоит в следующем:

Согласно (6.64), находим $t \equiv (\omega^2 - 1)^{1/2}$ как положительный корень уравнения

$$t^3 + 3t - \sqrt{\beta_1} = 0. \quad (6.66)$$

Тогда параметр сдвига μ определяется из (6.63) по формуле

$$\mu = \mu'_1(x) - \frac{\sqrt{\mu_2}}{t}. \quad (6.67)$$

Параметр ρ определяется значениями μ_2 и t , и, следовательно, γ и δ определяются значениями ω и β . Юэн (1933) дал таблицы для вычисления положительного корня уравнения (6.66).

За более исчерпывающими сведениями по истории, теории и применениям логнормального распределения мы отсылаем читателя к книге: J. Aitchison and J. A. C. Brown, *The Lognormal Distribution*, 1957, Cambridge University Press.

Пример 6.5

Для сравнения займемся подгонкой логнормальной кривой к распределению бобов по их длине, рассмотренному в предыдущем примере этой главы. Как мы уже видели, распределение бобов по их длине имеет «скос» влево, поэтому будем подгонять

логнормальную кривую с обращенной областью изменения переменной, т. е. таким образом, чтобы отрицательным значениям ξ отвечали бы наибольшие значения x . Это, конечно, не влияет на арифметический процесс подгонки, за исключением финального этапа.

Из примера 6.1 мы имеем $\sqrt{\beta_1} = -0,910569$. Поскольку мы рассматриваем теперь обращенную область изменения переменной, то $\sqrt{\beta_1}$ следует взять с обратным знаком; тогда уравнение (6.66) примет вид

$$t^3 + 3t + 0,910569 = 0.$$

Положительный корень $t = 0,294968$. Далее мы имеем (см. пример 6.1) $\mu_2 = 3,238425$. Тогда из второго уравнения (6.63) следует, что $\rho = 5,8516$. Заметим, между прочим, что из (6.65) для варианты ξ получим значение $\beta_2 = 4,510$, в то время как для x $\beta_2 = 4,863$ (см. пример 6.1). Отсюда видно, что наше преобразование, грубо говоря, не меняет значения β_2 . Итак, имеем

$$1/(2\delta^2) = \log_e \omega = \frac{1}{2} \log_e (1 + t^2) = 0,041718,$$

$$-\gamma/\delta = \log_e \rho = 1,7667,$$

то дает

$$\delta = 3,462, \quad \gamma = -6,116.$$

За единицу измерения, равную ширине интервала группировки, здесь принято $\frac{1}{2}$ мм, и, следовательно, согласно данным примера 6.1 варианта x имеет среднее значение $14,5 - \frac{1}{2} (0,19078) = 14,405$. Начало отсчета x в силу (6.67) равно

$$14,405 - \sqrt{\mu_2}/t = 17,455.$$

Тем самым кривая «начинается» в точке 17,455 мм.

Столбец (7) таблицы 6.1 на странице 223 дает соответствующие значения частот. Результат подгонки не очень хорош. Будучи более плохим, чем результат аппроксимации с помощью распределения типа IV, он все же лучше, чем соответствующее приближение с помощью ряда Грама — Шарлье.

6.32 Кривые системы S_B с $g(y) = \log \{y/(1-y)\}$ задают преобразование

$$\xi = \gamma + \delta \log \frac{x - \mu}{\lambda + \mu - x}. \quad (6.68)$$

Заметим, однако, что это преобразование можно получить как логнормальное $\xi = \gamma + \delta \log (x - \mu)$ с $y = 1 - \mu/x$. Мы можем

также записать

$$\xi = \gamma + \delta \log \{y/(1-y)\} = \gamma + 2\delta \operatorname{argth}(2y-1). \quad (6.69)$$

Эта формула дает возможность совершать численные подсчеты с помощью таблиц обратного гиперболического тангенса. Плотность распределения для y , скажем $p(y)$, задается формулой

$$p(y) = \frac{\delta}{\sqrt{2\pi}} \frac{1}{y(1-y)} \exp \left\{ -\frac{1}{2} \left(\gamma + \delta \log \frac{y}{1-y} \right)^2 \right\}, \quad (6.70)$$

$$0 \leq y \leq 1.$$

Без труда можно показать, что эта кривая имеет с осью абсцисс соприкосновение бесконечного порядка на обоих концах интервала изменения y .

6.33 Кривые плотностей вероятностей (6.70) могут быть, однако, бимодальны. Дифференцируя (6.70), мы найдем уравнение для модальных значений y :

$$2y - 1 = \delta \left(\gamma + \delta \log \frac{y}{1-y} \right).$$

Полагая $y = \frac{1}{2}(y' + 1)$, получим

$$y' - \gamma\delta = \delta^2 \log \frac{1+y'}{1-y'}.$$

В координатной плоскости (y', u) прямая линия $u = y' - \gamma\delta$ пересекает кривую $u = \delta^2 \log \left(\frac{1+y'}{1-y'} \right)$ в одной или трех точках. В первом случае кривая унимодальна, во втором — бимодальна. Можно показать (мы опускаем доказательство), что необходимые и достаточные условия для бимодальности состоят в том, что

$$\delta < 1/\sqrt{2}, \quad |\gamma| < \delta^{-1} \sqrt{1-2\delta^2} - 2\delta \operatorname{argth} \sqrt{1-2\delta^2}. \quad (6.71)$$

6.34 Моменты распределения обычно очень трудно определить в приемлемой для приложений форме. Поэтому для целей подгонки распределений чаще оказывается проще использовать некоторые другие величины, такие как, например, квантили, что и сделано в следующем примере.

(Джонсон, 1949а.)

Рассмотрим данные, приведенные в таблице 1.11 (облачность в Гринвиче). Условимся считать, что оценки состояния облачности, выраженные в баллах 0, 1, 2, ..., возникли в результате группировки по интервалам $[-0,5; 0,5)$, $[0,5; 1,5)$ и т. д. В таком

случае за начало отсчета можно принять $\mu = -0,5$, а в качестве масштабного параметра выбрать $\lambda = 11$. Оставшиеся два параметра γ и δ подберем так, чтобы получить точное совпадение для двух крайних класс-интервалов. Так, в первом классе $y = 1/11$ и $\log_e\{y/1 - y\} = -2,302$. Нормальное отклонение, отвечающее частоте $\frac{320}{1715} = 0,186589$, равно $-0,8965$, и, следовательно, согласно (6.69)

$$-0,8965 = \gamma - 2,3026 \delta.$$

Аналогично, для другого конца $y = 10/11$ и отклонение равно $0,2684$. Следовательно,

$$0,2684 = \gamma + 2,3026 \delta.$$

Таблица 6.2

Частотные кривые, соответствующие различным способам подгонки данных таблицы 1.11

Степень облачности	Наблюденные частоты	$S_B(1)$	$S_B(2)$	Тип I Пирсона
0	320	320,0	320,0	321,7
1	129	100,9	120,9	121,5
2	74	73,9	72,0	75,1
3	68	63,8	57,5	61,4
4	45	59,8	52,1	56,0
5	45	59,9	51,6	55,2
6	55	63,4	54,9	57,8
7	65	72,0	63,9	65,5
8	90	90,0	85,5	83,2
9	148	135,4	160,7	139,6
10	676	676,0	676,0	678,0
Всего	1715	1715,1	1714,9	1715,0

Полученные два уравнения дают

$$\gamma = 0,3110, \quad \delta = 0,25166.$$

Соответствующий результат подгонки помещен в таблице 6,2 в столбце $S_B(1)$. Кривая $S_B(2)$ — результат подгонки тем же самым путем в предположении, что группирование относится к интервалам $[0; 0,5)$, $[0,5, 1,5)$ и т. д., и что $\mu = 0$, а $\lambda = 10$. Представлен также результат подгонки с помощью кривой Пирсона I типа. Все три кривые имеют тот недостаток, что они не дают достаточно малых частот в окрестности центральной точки интервала изменения варианты,

6.35 Семейство кривых S_U определяется уравнением

$$\xi = \gamma + \delta \operatorname{argsh} y. \quad (6.72)$$

Соответствующие плотности выражаются формулой

$$p(y) = \frac{\delta}{\sqrt{2\pi}} \frac{1}{\sqrt{y^2 + 1}} \exp \left[-\frac{1}{2} [\gamma + \delta \log \{y + \sqrt{y^2 + 1}\}]^2 \right], \quad (6.73)$$

$$-\infty \leq y \leq \infty.$$

Графики их функций имеют высокий порядок соприкосновения с осью абсцисс на концах интервала изменения аргумента. Без труда можно показать, что эти кривые унимодальны и их моды лежат между медианой и нулем. Тем самым эти распределения имеют положительную или отрицательную асимметрию в зависимости от того, отрицательна или положительна величина γ . Моменты выражаются формулой

$$\mu'_r = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left(-\frac{1}{2} \xi^2 \right) \frac{1}{2r} \left\{ \exp \left(\frac{\xi - \gamma}{\delta} \right) - \exp \left(-\frac{\xi - \gamma}{\delta} \right) \right\}^r d\xi$$

и, в частности,

$$\left. \begin{aligned} \mu'_1 &= -\omega \operatorname{sh} \Omega, \\ \mu'_2 &= \frac{1}{2} (\omega^2 - 1) (\omega^2 \operatorname{ch} 2\Omega + 1), \\ \mu'_3 &= -\frac{1}{4} \omega^2 (\omega^2 - 1)^2 \{ \omega^2 (\omega^2 + 2) \operatorname{sh} 3\Omega + 3 \operatorname{sh} \Omega \}, \\ \mu'_4 &= \frac{1}{8} (\omega^2 - 1)^2 \{ \omega^4 (\omega^8 + 2\omega^6 + 3\omega^4 - 3) \operatorname{ch} 4\Omega + \\ &\quad + 4\omega^4 (\omega^2 + 2) \operatorname{ch} 2\Omega + 3(2\omega^2 + 1) \}, \end{aligned} \right\} \quad (6.74)$$

где $\omega = \exp(1/2\delta^2)$, $\Omega = \gamma/\delta$.

Соответствующий метод подгонки иллюстрируется следующим примером.

Пример 6.7

Рассмотрим еще раз распределение бобов по их длине, которое мы уже изучали в предшествующих примерах. Для распределения наблюдаемых величин, сгруппированных в класс-интервалы длины 0,5 мм:

$$\text{среднее} = 14,405 \text{ мм}, \quad \beta_1 = 0,829,$$

$$\mu_2 = (0,9036)^2, \quad \beta_2 = 4,863,$$

Мы можем подставить эти значения в уравнения (6.74) и разрешить их относительно ω и Ω . Этот процесс утомителен, и для его осуществления Джонсоном (1949а) была построена номограмма. По этой номограмме $\delta=2,64$, $\gamma=2,38$. Используя эти значения и соотношения

$$M_{\xi} = M \{(x - \mu)/\lambda\} = 1,1029, \quad D(x/\lambda) = (0,5948)^2,$$

находим μ и λ :

$$\lambda = 1,5192, \quad M_{\xi} = 14,399 + 1,1029\lambda = 16,0745.$$

Восьмой столбец в таблице 6.1 на стр. 223 показывает частоты, выровненные с помощью кривой S_U , соответствующей этим значениям констант. Подгонка выглядит столь же хорошей, как и подгонка с помощью кривых Пирсона IV типа.

6.36 Подгонка математических кривых к наблюдаемым данным имеет некоторый самостоятельный интерес, выходящий за рамки статистических приложений. Чтобы не перегружать данную главу, мы кратко сформулируем ряд других смежных с рассмотренными нами тем, которые обсуждались в статистической теории. Читателям, интересующимся подробностями, предоставляется возможность самостоятельно разобраться в рекомендуемой нами литературе.

Распределения Бёрра

6.37 Все три рассмотренные в этой главе системы предназначались для подгонки соответствующих кривых к *плотности* распределения. Мы видели, что в ряде случаев выполнение суммирования или интегрирования с целью нахождения функции распределения оказывается трудным делом. Естественно возникает вопрос, нельзя ли подобрать кривые сразу к функции распределения, а затем, в случае необходимости, получить плотность распределения относительно простой операцией дифференцирования.

Такой подход был применен Бёрром (1942) и Хатке (1949). Обобщая уравнение Пирсона (6.1), рассмотрим

$$dF = F(1 - F)g(x)dx, \quad (6.75)$$

где $g(x)$ — некоторая подходящим образом выбранная функция, которая должна быть неотрицательной при $0 \leq F \leq 1$ и в области изменения x . Решение уравнения (6.75) получаем сразу, так как

$$dF \left(\frac{1}{F} + \frac{1}{1-F} \right) = g(x)dx,$$

что дает

$$F(x) = [1 + \exp \{-G(x)\}]^{-1}, \quad (6.76)$$

где

$$G(x) = \int_{-\infty}^x g(t) dt.$$

Например, удобной является форма

$$F(x) = 1 - \frac{1}{(1 + x^\alpha)^\beta}, \quad 0 \leq x < \infty, \quad (6.77)$$

как, впрочем, и многие другие.

6.38 Метод подгонки, предложенный Бёрром, состоит в приравнении *кумулятивных моментов* (вычисленных по имеющимся данным)

$$M_j(a) = \int_a^\infty (x-a)^j \{1 - F(x)\} dx - \int_{-\infty}^a (x-a)^j F(x) dx \quad (6.78)$$

теоретическим значениям, выраженным в терминах параметров распределения. В действительности эта операция с кумулятивными моментами совпадает с той операцией, которую Пирсон проводил для обыкновенных моментов. Эта операция довольно утомительна, несмотря на то, что Бёрр и Хатке дали вспомогательные таблицы. Упрощенный метод подгонки, основанный (как в примере 6.6) на равенстве подогнанных и наблюдаемых частот в некоторых класс-интервалах, не нашел достаточно широкого применения.

Подгонка с помощью вероятностных моментов

6.39 Подгонка частотных распределений по методу моментов имеет некоторые недостатки особенно тогда, когда область значений варианты не ограничена и результаты наблюдений представляют собой лишь выборку из множества возможных наблюдений. Резко выделяющиеся элементы и «хвосты» частот оказывают сильное влияние на величину моментов. Возможен другой путь подгонки, состоящий в приравнении теоретических и эмпирических (т. е. вычисленных по выборке) значений так называемых *вероятностных моментов*

$$\omega_r = \int_{-\infty}^{\infty} [f(x)]^r dx. \quad (6.79)$$

Этот метод подробно был разобран Сичелом (1949) для одного из случаев, когда кривые Пирсона дают не очень хороший результат. Все, что говорилось в 6.12 об эффективности подгонки кривыми Пирсона здесь также остается в силе. С точки зрения, принятой в настоящей главе, соответствующий метод состоит в приравнивании эмпирических и теоретических значений величин (6.79). Вероятностные моменты требуют, однако, поправок на группирование, и метод не получил достаточно широкого распространения.

Моригути (Moriguti, 1952, Ann. Math. Statist. 23, 286) показал, что для каждого непрерывного распределения с дисперсией σ^2

$$\sigma^{r-1} \omega_r \geq \frac{2r}{3r-1} \left[\left(\frac{r-1}{3r-1} \right)^{1/2} / B \left(\frac{r}{r-1}, \frac{1}{2} \right) \right]^{r-1}.$$

Подгонка с помощью усеченных распределений

6.40 Мы кратко отметим тот факт, что некоторые наблюдаемые распределения аппроксимируются такими математическими выражениями, которые распространяются на ненаблюдаемые значения варианты. Возможен, конечно, и другой подход, при котором подгонка соответствующей кривой осуществляется только в том интервале, где расположены результаты наблюдений. Однако подгонка с помощью таких «усеченных» распределений очень трудна, к тому же моменты зависят от крайних точек, значения которых мало что говорят о характере истинного распределения. Систематическое рассмотрение этих вопросов начато совсем недавно и посвящено главным образом изучению отдельных частных случаев.

Двумерные распределения

6.41 Удовлетворительного метода для описания семейств двумерных плотностей, подобного тем методам, которые мы рассмотрели в одномерном случае, все еще не найдено. Некоторые усилия в этом направлении были сделаны Преториусом (1930) и Джонсоном (1949). В рассматриваемом случае можно ввести двумерные ряды Грама — Шарлье (см. упражнения 6.16 и 6.17). Если эта задача вызвала бы интерес, то возможно, что лучший путь ее решения состоял бы в том, чтобы искать такое преобразование переменных, которое бы делало переменные приблизительно независимыми, и тогда задача свелась бы к подгонке подходящих кривых к одномерным распределениям.

УПРАЖНЕНИЯ

6.1 Показать, что следующие плотности распределений принадлежат системе Пирсона, и начертить их графики для некоторых специальных значений параметров:

(а) (нормальное распределение)

$$f = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad -\infty \leq x \leq \infty;$$

(б) (распределение II типа)

$$f = \frac{1}{aB\left(\frac{1}{2}, m+1\right)} \left(1 - \frac{x^2}{a^2}\right)^m, \quad -a \leq x \leq a;$$

(в) (распределение VII типа)

$$f = \frac{1}{aB\left(\frac{1}{2}, m-1\right)} \left(1 + \frac{x^2}{a^2}\right)^{-m}, \quad -\infty \leq x \leq \infty.$$

6.2 Делая в случае необходимости простые преобразования вариант, показать, что каждое из следующих распределений принадлежит системе Пирсона:

$$dF = ke^{-\frac{1}{2}x^2} x^{v-1} dx, \quad dF = k(1-r^2)^{\frac{1}{2}(n-4)} dr,$$

$$dF = \frac{k dt}{(1+t^2/v)^{\frac{1}{2}(v+1)}}, \quad dF = k\eta^{p-2}(1-\eta^2)^{\frac{1}{2}(N-p-2)} d\eta.$$

(Все эти распределения играют важную роль в выборочной теории.)

6.3 Показать, что пирсоновские распределения

$$\frac{d \log f}{dx} = \frac{x}{B_0 + B_1 x + B_2 x^2}$$

сосредоточены на всей прямой, если $B_0 + B_1 x + B_2 x^2$ не имеет действительных корней; на полупрямой — если корни действительные и одного знака; на конечном интервале — если корни действительные и противоположных знаков.

6.4 Показать, что плотность распределения типа VI системы Пирсона может быть записана в виде

$$y = y_0 \left(1 - \frac{x^2}{a^2}\right)^{-m} \exp\left\{-v \arg \operatorname{th} \left(\frac{x}{a}\right)\right\}.$$

Сравнить с распределениями IV типа.

6.5 Показать, что для пирсоновских распределений III типа $2\beta_2 - 3\beta_1 - 6 = 0$ и, следовательно, величина κ бесконечна.

6.6 Если корни полинома $B_0 + B_1 X + B_2 X^2$ равны, то показать, что $\kappa = 1$ и что соответствующие этому случаю уравнения кривых Пирсона (V тип) могут быть записаны в следующем виде:

$$f = \frac{v^{p-1}}{\Gamma(p-1)} x^{-p} e^{-v/x}, \quad 0 \leq x \leq \infty.$$

6.7 Показать, что для пирсоновского распределения

$$\frac{df}{f} = \frac{(a+x) dx}{b_0 + b_1 x + b_2 x^2}$$

характеристическая функция подчиняется соотношению

$$b_2 \theta \frac{d^2 \varphi}{d\theta^2} + (1 + 2b_2 + b_1 \theta) \frac{d\varphi}{d\theta} + (a + b_1 + b_0 \theta) \varphi = 0,$$

где $\theta = it$. Вывести отсюда рекуррентные формулы для моментов.

Показать также, что логарифм производящей функции удовлетворяет соотношению

$$b_2 \theta \left\{ \frac{d^2 \psi}{d\theta^2} + \left(\frac{d\psi}{d\theta} \right)^2 \right\} + (1 + 2b_2 + b_1 \theta) \frac{d\psi}{d\theta} + (a + b_1 + b_0 \theta) = 0.$$

Пользуясь этим соотношением, получить рекуррентную формулу для семинвариантов:

$$\begin{aligned} \{1 + (r+2)b_2\} \kappa_{r+1} + r b_1 \kappa_r + r b_2 \left\{ \kappa_1 \kappa_r + \binom{r-1}{1} \kappa_2 \kappa_{r-1} + \right. \\ \left. + \binom{r-1}{2} \kappa_3 \kappa_{r-2} + \dots + \binom{r-1}{j} \kappa_{j+1} \kappa_{r-j} + \dots \right. \\ \left. \dots + \binom{r-1}{1} \kappa_{r-1} \kappa_2 + \kappa_r \kappa_1 \right\} = 0. \end{aligned}$$

6.8 (Ряд Шарлье типа В). Определим

$$\gamma(\lambda, x) = e^{-\lambda} \lambda^x / x!, \quad \nabla \gamma(\lambda, x-1) = \gamma(\lambda, x) - \gamma(\lambda, x-1),$$

$$G_r(\lambda, x) = \left\{ \frac{d^r}{d\lambda^r} \gamma(\lambda, x) \right\} / \gamma(\lambda, x);$$

показать, что

$$G_r(\lambda, x) = \{(-1)^r \nabla^r \gamma(\lambda, x)\} / \gamma(\lambda, x),$$

$$\sum_{x=0}^{\infty} (G_r G_s \gamma) = \begin{cases} 0, & r \neq s, \\ r! / \lambda^r, & r = s. \end{cases}$$

Если f может быть представлена в виде

$$f = \sum_j (b_j G_j \gamma),$$

то

$$b_j = \frac{\lambda^j}{j!} \sum_{x=0}^{\infty} (f G_j).$$

6.9 Показать, что не существует распределений, представимых в виде сходящегося ряда типа А, которые бы не определялись полностью своими моментами.

6.10 Пусть y есть функция от x , которую мы желаем приближенно представить в виде $y = \sum_{j=0}^r c_j H_j(x) \alpha(x)$. Показать, что тогда значения c_j падо

выбирать таким образом, чтобы они минимизировали интеграл

$$\int_{-\infty}^{\infty} \left\{ \frac{y - \sum_{j=0}^r c_j H_j(x) a(x)}{\sqrt{a(x)}} \right\}^2 dx.$$

6.11 Показать, что если функция плотности нормированного распределения разложена в ряд типа А, то коэффициенты при втором и третьем членах зависят соответственно от β_1 и β_2 и тем самым представляют собой меры асимметрии и эксцесса.

6.12 Пусть $f(x)$ и $g(x)$ — две дифференцируемые функции плотности с семинвариантами соответственно κ и κ' . Показать, что тогда формально

$$f(x) = \exp \left\{ \sum_{j=0}^{\infty} \frac{\kappa_j - \kappa'_j}{j!} \left(-\frac{d}{dx} \right)^j \right\} g(x).$$

6.13 Если y имеет распределение

$$dF = \frac{1}{\Gamma(\lambda)} y^{\lambda-1} e^{-y} dy, \quad 0 \leq y < \infty,$$

и $\xi = y + \delta \log y$, то семинварианты ξ определяются по формулам

$$\kappa_1(\xi) = \gamma + \delta \frac{d}{d\lambda} \log \Gamma(\lambda),$$

$$\kappa_r(\xi) = \delta^r \frac{d^r}{d\lambda^r} \log \Gamma(\lambda), \quad r > 1.$$

Следовательно,

$$\beta_1(\xi) = \frac{1}{\lambda} + O(\lambda^{-2}), \quad \beta_2(\xi) = 3 + \frac{2}{\lambda} + O(\lambda^{-2}).$$

Сравнить эти значения с соответствующими коэффициентами для y :

$$\beta_1(y) = \frac{4}{\lambda}, \quad \beta_2(y) = 3 + \frac{6}{\lambda}.$$

(Джонсон, 1949а.)

6.14 Если y имеет распределение

$$dF = \frac{\Gamma(\tau)}{\Gamma(\tau-\nu)\Gamma(\nu)} y^{\nu-1} (1+y)^{-\tau}, \quad 0 \leq y < \infty,$$

и $\xi = y + \delta \log y$, то для больших ν и $\tau - \nu$

$$\beta_1(\xi) = \frac{1}{\nu} - \frac{4}{\tau} + \frac{1}{\tau-\nu}, \quad \beta_2(\xi) = 3 + \frac{2}{\nu} - \frac{6}{\tau} + \frac{2}{\tau-\nu}.$$

Сравнить с

$$\beta_1(y) = \frac{4}{\nu} - \frac{4}{\tau} + \frac{16}{\tau-\nu}, \quad \beta_2(y) = 3 + \frac{6}{\nu} - \frac{6}{\tau} + \frac{30}{\tau-3}.$$

(Джонсон, 1949а.)

6.15 Если R распределена по закону

$$dF = \frac{\left\{ \Gamma\left(\frac{1}{2}n-1\right) \right\}^2}{\Gamma(n-2)} R^{\frac{1}{2}(n-4)} (1-R)^{\frac{1}{2}(n-4)}, \quad 0 \leq R \leq 1,$$

то тогда для

$$\xi = \frac{1}{2} \sqrt{n-3} \log \{R/(1-R)\}$$

имеем

$$\begin{aligned} \beta_1(\xi) &= 0, \\ \beta_2(\xi) &= 3 + \frac{\frac{1}{2} \left[\frac{d^2}{d\lambda^2} \log \Gamma(\lambda) \right]_{\lambda=\frac{1}{2}n-1}}{[\log \Gamma(\lambda)]^2_{\lambda=\frac{1}{2}n-1}}. \end{aligned}$$

(Джонсон, 1949a.)

6.16 Если x и y имеют совместное нормальное распределение

$$f(x, y) = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} (x^2 - 2\rho xy + y^2) \right\},$$

$$-\infty \leq x, y \leq \infty,$$

и $D_1 = d/dx$, $D_2 = d/dy$, то логарифм характеристической функции выражения

$$\exp \left\{ \kappa_{11} D_1 D_2 + \sum_{r+s \geq 3} (-1)^{r+s} \kappa_{rs} \frac{D_1^r D_2^s}{r! s!} \right\} \alpha(x) \alpha(y) \quad (A)$$

представляет собой выражение

$$-\frac{1}{2} (t_1^2 + 2\rho t_1 t_2) + \sum_{r+s \geq 3} \kappa_{rs} \frac{(it_1)^r (it_2)^s}{r! s!}.$$

Отсюда получить, что двумерная форма ряда Эджворта может быть представлена или в виде (A) или как

$$\exp \left\{ \sum_{r+s \geq 3} (-1)^{r+s} \kappa_{rs} \frac{D_1^r D_2^s}{r! s!} \right\} f(x, y).$$

(Кендалл, 1949b.)

6.17 Пользуясь выражением (A) из предыдущего упражнения, показать, что формально

$$\frac{\partial f}{\partial \kappa_{rs}} = (-1)^{r+s} \frac{D_1^r D_2^s}{r! s!} f, \quad r+s \geq 3 \quad \text{или} \quad r=s=1.$$

Получить эту формулу прямым подсчетом для двумерного нормального распределения с $r=s=1$, т. е. установить, что

$$\frac{\partial f}{\partial \rho} = \frac{\partial^2 f}{\partial x \partial y}.$$

(Кендалл, 1949b.)

6.18 Доказать, что кривые системы S_U из 6.35 унимодальные, и показать, что они имеют положительную или отрицательную асимметрию, в зависимости от отрицательности или положительности γ .

6.19 Показать, что следующие функции распределения являются допустимыми решениями уравнения (6.76) (параметры во всех случаях положительны):

$$F(x) = (1 + e^{-x})^{-\alpha}, \quad -\infty \leq x \leq \infty;$$

$$F(x) = (1 + ae^{-\operatorname{tg} x})^{-\alpha}, \quad -\frac{1}{2}\pi \leq x \leq \frac{1}{2}\pi;$$

$$F(x) = \left(\frac{2}{\pi} \operatorname{arctg} e^{-x}\right)^{\alpha}, \quad -\infty \leq x \leq \infty;$$

$$F(x) = \left(x - \frac{1}{2\pi} \sin 2\pi x\right)^{\alpha}, \quad 0 \leq x \leq 1.$$

(Бёрр, 1942.)

6.20 Некоторая варианта имеет нулевое среднее, единичную дисперсию и r -е семинварианты κ_r порядка $n^{1-\frac{1}{2}r}$, $r \geq 2$. С помощью разложения Эджворта показать, что с точностью до величин порядка n^{-1} мода равна $-\frac{1}{2}\kappa_3$ и медиана равна $-\frac{1}{6}\kappa_3$. Отсюда вывести приближенное соотношение

$$\frac{(\text{медиана} - \text{мода})}{(\text{среднее} - \text{мода})} = \frac{2}{3}.$$

(Холдейн, 1942.)

ИСЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ

7.1 В предыдущих шести главах рассматривалась теория статистических распределений с дескриптивной точки зрения. Мы видели, что распределения, встречающиеся в практике, обладают некоторыми свойствами регулярности, позволяющими представить их в математической форме, что они могут быть охарактеризованы некоторыми количествами, такими как моменты и семиинварианты, и что можно установить некоторые общие теоремы о функциях распределения и функциях плотности. Теперь мы начинаем изучение иного характера, именно исследование возможности делать значимые и объективные утверждения относительно генеральной совокупности или порождающего ее механизма, когда рассмотрению доступна лишь некоторая выборка возможных наблюдений. В широком смысле это и представляет собой предмет статистических выводов.

7.2 За исключением тривиальных случаев, невозможно делать выводы о генеральной совокупности по выборке с той определенностью, которая присуща дедуктивной логике. Наши утверждения или должны быть несколько неопределенно сформулированы, или им должны сопутствовать различного рода сомнения. В обычной речи используются разные слова и фразы для обозначения отношения к утверждениям, в истинности которых мы сомневаемся. Говорят, что нечто более или менее *вероятно*, более или менее *правдоподобно*, что имеется некоторая *уверенность* в справедливости утверждения или что имеются *шансы* против него; мы относимся к некоторому утверждению или как к *сомнительному*, или открытому для сомнений, или же как к требующему *подтверждения*. Многие из этих выражений приобрели техническое значение в статистике, которое будет указано в соответствующем месте. В сущности, они выражают различные аспекты одного и того же, а именно — неопределенности, с которой мы в силу обстоятельств относимся к отдельным утверждениям.

7.3 В следующей главе будут обсуждаться некоторые из этих идей в аспекте статистических выводов. Однако, прежде чем приступить к этому, мы должны дать краткое изложение ис-

числения, которое потребуется для того, чтобы сделать подход последовательным и систематическим, а именно исчисления вероятностей.

Некоторые авторы понимают под «теорией вероятностей» все, что касается неопределенных выводов. Другие включают в эту теорию только то, что связано с множеством повторяемых событий. При первом подходе «вероятность» рассматривают как неопределяемое понятие, подобно прямой линии в геометрии Евклида, и строят теорию, исходя из некоторых аксиом. При втором подходе обычно пытаются определить вероятность в терминах относительных частот событий и тем самым свести теорию к чисто математическому изучению абстрактных множеств или предельных свойств последовательностей. Наше собственное мнение (Кендалл, 1949а) состоит в том, что ни один из этих подходов сам по себе не достаточен для статистика, каково бы ни было его значение для математика. Представляется, однако, что каждый должен сделать для себя выбор и что тип предпочитаемой аксиоматики определяется психологическим складом, опытом и сферами интересов человека. В статистике слишком много спорить об основаниях теории вероятностей означает свидетельствовать свою незрелость.

7.4 В этой главе мы не рассматриваем теорию вероятностей в широком смысле. Она является ветвью научной методологии. Мы будем заниматься исчислением вероятностей, которое представляет собой ветвь математики. Английский язык в известной степени ответствен за смешение этих двух различных вещей. В средние века и в эпоху Возрождения существовало различие между понятием *probabilitas* (вероятность), которое относилось к величине степени уверенности, и понятием *доктрина шансов*, означающим учение о вычислении числа способов, по которым могут произойти некоторые классы событий. В семнадцатом и восемнадцатом веках доктрина шансов ассимилировалась с вероятностью так, что теперь одно и то же слово употребляется для обозначения и того и другого. Полезное различие утерялось. Мы можем принять, что «шансы» включаются в «вероятность», но мы будем различать *теорию вероятностей*, которая относится к использованию вероятностей, когда делаются утверждения относительно внешнего мира, и *исчисление вероятностей*, занимающееся математическим выведением следствий из некоторых аксиом и постулатов. К изложению этого последнего мы теперь и переходим.

7.5 Мы сформулируем основные правила исчисления, не делая попыток вывести их из более простых предложений. К счастью, каковы бы ни были разногласия относительно оснований, нет, или почти нет, разногласий относительно самих правил.

Во-первых, предполагается, что вероятность можно измерять по непрерывной шкале так, что всякая вероятность представляется действительным числом. Мы будем говорить просто, что вероятность какого-нибудь предложения *есть* действительное число x . Это допущение влечет среди прочего сравнимость всяких двух вероятностей, поскольку если вероятности измеряются числами x и y , то можно сказать, что первая из них больше, равна или меньше, чем вторая, в соответствии с одним из трех возможных соотношений $x > y$, $x = y$ или $x < y$.

7.6 Вероятность предложения q при данном p обозначается $P(q|p)$. *Правило 1* состоит в следующем:

$$\text{если } p \text{ влечет } q, \quad \text{то } P(q|p) = 1, \quad (7.1)$$

$$\text{если } p \text{ влечет не-}q, \text{ то } P(q|p) = 0. \quad (7.2)$$

Это правило определяет крайние точки шкалы вероятностей. Уверенность, что предложение неверно, представляется нулем, уверенность, что оно верно, — единицей. Всякая вероятность заключена между 0 и 1.

7.7 *Правило 2.* Если вероятности n взаимно исключающих предложений $q_1, \dots, q_n$ при данном p суть $P_1, \dots, P_n$, то вероятность того, что при данном p хотя бы одно из них верно, равна $P_1 + P_2 + \dots + P_n$.

Это правило обычно известно как «теорема сложения вероятностей». На языке учебников оно звучит так: вероятность того, что произойдет одно из взаимно исключающих событий, равна сумме вероятностей каждого из этих событий в отдельности.

7.8 Из правила 2 следует, что если предложения $q_1, \dots, q_n$ взаимно исключающие и одинаково вероятные при данном p и если Q — подмножество, содержащее m из них, причем одно из q обязательно верно, то вероятность того, что верное предложение принадлежит подмножеству Q , равна

$$P(Q|p) = m/n. \quad (7.3)$$

Это свойство является отправной точкой частотной теории вероятностей. Обычно оно формулируется в такой форме: если среди n взаимно исключающих и одинаково вероятных событий, одно из которых обязательно происходит, имеется m событий, наделенных некоторым характеристическим признаком A , то вероятность появления события с признаком A равна m/n . Использование указанного свойства в качестве *определения* вероятности некорректно, поскольку такое определение содержит в себе логический круг: в нем используется понятие «одинаково вероятные»,

7.9 Правило 3. Вероятность двух предложений q и r при данном p равна произведению вероятности q при данном p на вероятность r при данных q и p . Или символически

$$P(qr|p) = P(q|p)P(r|qp). \quad (7.4)$$

Так как q и r здесь симметричны, то имеем также

$$P(qr|p) = P(r|p)P(q|rp). \quad (7.5)$$

С частотной точки зрения это правило почти очевидно. Если в множестве из n элементов (a) обладают признаком A , (b) — признаком (B) и (ab) — обоими признаками, то правило утверждает, что

$$\frac{(ab)}{n} = \frac{(a)}{n} \frac{(ab)}{(a)} = \frac{(b)}{n} \frac{(ab)}{(b)},$$

т. е. сводится к простому арифметическому факту.

Применяя последовательно правило 3, получаем более общую формулу:

$$P(q_1q_2 \dots q_k|p) = P(q_1|p)P(q_2|q_1p)P(q_3|q_2q_1p) \dots \\ \dots P(q_k|q_{k-1} \dots q_1p). \quad (7.6)$$

Если, в частности,

$$P(qr|p) = P(q|p)P(r|p), \quad (7.7)$$

то в силу (7.4)

$$P(r|p) = P(r|qp). \quad (7.8)$$

В этом случае говорят, что q не связано с r при данном p . Знание q не влияет на вероятность r при данном p .

7.10 Приведенные три правила и различные их усовершенствования образуют базис исчисления вероятностей. Все дальнейшее представляет собой чистую математику. Мы предполагаем, что читатель имеет некоторое знакомство с элементами излагаемого предмета, и дадим лишь несколько примеров, для того чтобы подчеркнуть важность указанных правил.

Пример 7.1. Важность точного описания множества элементарных ходов

Подбрасываются три монеты. Какова вероятность того, что они падают все на герб или все на решетку?

Предположим, что вероятность выпадения герба на каждой монете равна $1/2$ и что результаты бросаний монет независимы. Тогда возможны восемь исходов: ГГГ, ГГР, ГРГ, ГРР, РГГ, РГР, РРГ, РРР, и все они равновероятны. В двух из них

все монеты падают одинаково — на герб или на решетку, так что искомая вероятность равна $\frac{1}{4}$.

Рассмотрим теперь такое рассуждение: имеются две возможности: либо результаты бросаний всех монет одинаковы, либо они одинаковы только для двух монет. Одна из этих двух возможностей доставляет нужное нам событие, и вероятность его равна поэтому $\frac{1}{2}$.

Рассмотрим еще такое рассуждение: имеются четыре возможности: выпадают три герба, два герба и одна решетка, две решетки и один герб, три решетки. Две из них доставляют нужное нам событие, и вероятность его равна поэтому $\frac{1}{2}$.

Рассмотрим, наконец, такое рассуждение: из трех монет две должны выпасть одинаково, а третья выпадает так же, как эти две, или иначе. Опять имеются две возможности и искомая вероятность равна $\frac{1}{2}$.

Все три последних рассуждения ошибочны. В них неравновероятным событиям приписывается равновероятность и затем применяется правило 2, которое здесь, конечно, незаконно. Например, в первом из этих рассуждений правильно говорится, что имеются две возможности, однако они неравновероятны при наших допущениях. Читатель может поинтересоваться и посмотреть, почему это так, а также проверить, что остальные два рассуждения неверны по той же самой причине.

Пример 7.2

Какова вероятность того, что при игре в бридж игрок имеет на руке после раздачи по меньшей мере два туза? Так поставленный вопрос не вполне точен, и в действительности имеются четыре вероятности, каждая из которых может представлять интерес:

(а) Вероятность того, что некоторый определенный игрок (скажем, Юг) имеет не меньше двух тузов, вне зависимости от расклада карт на других руках.

(б) Вероятность того, что какой-нибудь один игрок из четырех имеет не меньше двух тузов, вне зависимости от расклада карт на других руках.

(в) Вероятность того, что только Юг среди всех игроков имеет не меньше двух тузов.

(г) Вероятность того, что в точности один игрок из четырех имеет не меньше двух тузов.

Вычислим все эти вероятности по очереди.

(а) У Юга может быть 0, 1, 2, 3 или 4 туза. Колоду из 52 карт можно считать состоящей из двух частей — четырех тузов и остальных 48 карт. Вероятность того, что к Югу не приходит

ни одного туза, равна

$$p_0 = \frac{\binom{4}{0} \binom{48}{13}}{\binom{52}{13}}, \quad (7.9)$$

а вероятность прийти одному тузу —

$$p_1 = \frac{\binom{4}{1} \binom{48}{12}}{\binom{52}{13}}.$$

Отсюда находим искомую вероятность

$$p_a = 1 - (p_0 + p_1) = 1 - \frac{38 \cdot 37 \cdot 11}{49 \cdot 25 \cdot 17} = 0,257. \quad (7.10)$$

(б) Представляется соблазнительным думать, что вторая из рассматриваемых вероятностей равна четыре раза взятой первой, однако попадания по меньшей мере двух тузов различным игрокам не исключают друг друга. Что это неверно, ясно также из неравенства $4p_a > 1$. В действительности искомая вероятность p_b равна вероятности того, что тузы не распределяются каждому игроку по одному:

$$p_b = 1 - 4! \frac{\binom{48}{12} \binom{36}{12} \binom{24}{12} \binom{12}{12}}{\binom{52}{13} \binom{39}{13} \binom{26}{13} \binom{13}{13}};$$

множитель $4!$ появляется здесь из-за возможных распределений четырех тузов по игрокам. Производя вычисления, находим

$$p_b = 1 - \frac{13^3}{49 \cdot 25 \cdot 17} = 0,895. \quad (7.11)$$

(в) Чтобы получить третью вероятность, нужно из p_a вычесть вероятности того, что Юг и Север, Юг и Восток, Юг и Запад имеют каждый по два туза. Эти три события взаимно исключают друг друга, и вероятность каждого из них равна

$$p_{2,2} = \binom{4}{2} \frac{\binom{48}{11} \binom{37}{11} \binom{26}{13} \binom{13}{13}}{\binom{52}{13} \binom{39}{13} \binom{26}{13} \binom{13}{13}} = 0,02248;$$

множитель $\binom{4}{2}$ появляется здесь из-за возможных различных распределений четырех тузов в две пары. Искомая вероятность равна

$$p_c = p_a - 3p_{2,2} = 0,190. \quad (7.12)$$

(г) Четвертая вероятность есть сумма вероятностей четырех взаимно исключающих событий, каждая из которых равна p_c . Следовательно,

$$p_d = 4p_c = 0,759. \quad (7.13)$$

Пример 7.3. Индуктивные методы

В первых двух примерах числа интересующих нас исходов подсчитывались непосредственно. В широком классе случаев этот непосредственный подсчет затруднителен. В такой ситуации часто может быть применен индуктивный или полуиндуктивный метод решения. Мы проиллюстрируем один такой метод на довольно старой задаче.

Имеется n писем, каждому из которых соответствует определенный конверт. Письма распределены по конвертам случайным образом. Какова вероятность того, что ни одно из писем не вложено в соответствующий ему конверт?

Здесь под условием, что письма распределены «случайным образом», следует понимать равновероятность всех возможных распределений писем по конвертам. В силу правила 2 вопрос тогда сводится к следующей чисто алгебраической задаче: какую долю от общего числа случаев составляют те случаи, в которых ни одно письмо не лежит в соответствующем ему конверте?

Пусть u_n — число всех возможных случаев, в которых все письма распределены по конвертам неправильно. Рассмотрим какое-либо определенное письмо. Если это письмо лежит в конверте для другого письма, а это другое письмо лежит в конверте для исходного, что может случиться $n-1$ различными способами, то число возможных неверных распределений остальных $n-2$ писем равно u_{n-2} . А если выделенное письмо лежит не в своем конверте, что может произойти $n-1$ различными путями, и письмо от конверта, в котором лежит выделенное письмо, находится не в конверте от выделенного письма, то остальные письма могут неверно распределиться u_{n-1} способами. Следовательно, имеет место уравнение

$$u_n = (n-1)(u_{n-1} + u_{n-2}).$$

Переписывая его в виде

$$u_n - nu_{n-1} = -(u_{n-1} - (n-1)u_{n-2})$$

и полагая

$$v_n = u_n - nu_{n-1},$$

находим

$$v_n = -v_{n-1} = (-1)^{n-2} v_2,$$

т. е.

$$u_n - nu_{n-1} = (-1)^{n-2} (u_2 - 2u_1).$$

Но $u_1=0$ и $u_2=1$, поэтому

$$u_n - nu_{n-1} = (-1)^n$$

и, следовательно,

$$u_n = n! \left\{ \frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^n}{n!} \right\}.$$

Так как общее число различных возможных распределений писем по конвертам есть $n!$, то искомая вероятность равна

$$\frac{1}{2!} - \frac{1}{3!} + \dots + \frac{(-1)^n}{n!}, \quad (7.14)$$

т. е. равна сумме первых $n+1$ членов разложения e^{-1} (первые два члена сокращаются и в (7.14) не выписаны).

Вероятность на континууме

7.11 До сих пор мы рассматривали только вероятности событий конечного или дискретного характера. Наряду с этим возникает вопрос, можно ли придать какой-нибудь смысл вероятностям на континууме? Например, если в круг вписан квадрат и внутри круга случайно выбирается точка, то какова вероятность этой точке оказаться внутри квадрата? Если отрезок разбивается двумя точками на три отрезка, то какова вероятность того, что эти отрезки могут быть сторонами треугольника? Какова вероятность того, что $x < x_0$, где x — положительное действительное число, меньшее y_0 ? И тому подобное.

Все вероятности этого типа должны рассматриваться как пределы. Рассмотрим первый пример — круг с вписанным в него квадратом. Представим себе, что вся фигура разбита прямоугольной сеткой на маленькие клетки площади ε . Если предположить, что вероятность точке оказаться в клетке одинакова для всех клеток, то вероятность попадания точки в квадрат равна отношению числа клеток, лежащих в квадрате, к числу клеток, лежащих в круге; клетки, задевающие окружность или квадрат, при этом не учитываются, поскольку их значимость исчезает при $\varepsilon \rightarrow 0$. Искомую вероятность можно сделать сколько угодно близкой к отношению площади квадрата к площади окружности, если ε взять достаточно малым. Мы можем сказать, что искомая вероятность *есть* это отношение, т. е. равна иррациональному числу $2/\pi$.

Та же предельная вероятность получилась бы, если бы было взято иное разбиение, аппроксимирующее площади. Важно отметить, однако, что когда речь идет о вероятностях на континууме, необходимо точно указывать предельный процесс, с помощью которого они получаются. В противном случае вероятности на континууме не имеют смысла. Следующий пример иллюстрирует это обстоятельство.

Пример 7.4

Рассмотрим отрезок OA , и пусть B — точка, делящая его пополам. Какова вероятность того, что случайно выбранная из OA точка принадлежит OB ?

Предположим сначала, что отрезок OA разделен на n одинаковых отрезков длины OA/n каждый. Если интерпретировать случайный выбор точки как случайный выбор одного из этих отрезков, то искомая вероятность в пределе при $n \rightarrow \infty$ будет равна, очевидно, $1/2$, поскольку в отрезок OB попадает половина отрезков разбиения.

Восставим теперь перпендикуляр OP к отрезку OA и возьмем на нем точку P так, чтобы $OP=OA$. Представим себе пучок, образованный $n+1$ лучами, исходящими из точки P и пересекающими OA , включающий лучи, идущие через O и A ; предположим, что лучи эти делят угол $OPA (= \pi/4)$ на равные углы $(= \pi/4n)$. Лучи высекают на OA отрезки, и если равные углы считать равновероятными, то этим отрезкам тоже можно приписать равные вероятности, так как они противолежат равным углам. Если принять это соглашение, то при $n \rightarrow \infty$ вероятность точке принадлежать любому отрезку, лежащему в OA , пропорциональна, очевидно, соответствующему ему углу в точке P . Например, вероятность того, что точка принадлежит отрезку OP , будет равна $(\arctg \frac{1}{2}) / (\frac{1}{4} \pi)$.

Мы получили теперь иной ответ в сравнении с тем случаем, когда всем малым отрезкам из OA приписывалась одинаковая вероятность. В этом нет никакого противоречия: эти два ответа различны, поскольку для их получения использовались различные предельные процессы. Нетрудно видеть, что, двигая точку P по перпендикуляру к OA в точке O и рассматривая, как и выше, пучок лучей, вероятность точке находиться в отрезке OB можно сделать равной любому числу между $1/2$ и 1. Следовательно, совершенно ясно, что понятие вероятности на континууме существенно связано с предельным процессом, посредством которого эти вероятности определяются, исходя из конечного разбиения на равновероятные интервалы.

Понятие случайной величины

7.12 Пусть имеется дискретная конечная генеральная совокупность, элементы которой распределены в соответствии с вариантом x . Если из этой генеральной совокупности случайным образом выбирается элемент, то вероятность того, что соответствующее ему значение варианты есть x_0 , равна $f(x_0)$ — значению функции плотности в x_0 , поскольку $f(x_0)$ есть доля элементов, обладающих этим свойством. Далее, вероятность того, что значение варианты, соответствующей выбранному элементу, не превосходит x_0 , равна $F(x_0)$ — значению в x_0 функции распределения варианты x ; это непосредственно следует из правила 2 и определения функции распределения.

Указанный факт говорит о существенной связи между вероятностями и распределениями. Значения функции распределения дают вероятности того, что значение варианты на случайно выбранном элементе генеральной совокупности равно или меньше некоторого фиксированного числа. Мы должны еще, однако, рассмотреть вопрос о том, остается ли это утверждение верным для бесконечных и непрерывных генеральных совокупностей.

Предположим сначала, что генеральная совокупность бесконечна и дискретна. В этом случае мы не можем случайно выбрать элемент, однако можно представить себе выбор из конечной генеральной совокупности, приближающейся к рассматриваемой бесконечной. В этой конечной генеральной совокупности доля элементов, на которых варианта не превосходит x_0 , равна $F(x_0)$; учитывая должным образом природу предельного процесса, можно сказать, что и в случае бесконечной генеральной совокупности варианта не превосходит x_0 с вероятностью $F(x_0)$.

Аналогично для непрерывного распределения. В главе 1 непрерывное распределение рассматривалось как предельное для дискретного распределения, задаваемого формулой

$$\Delta F = f(x) \Delta x.$$

Если из соответствующей генеральной совокупности случайно выбирается элемент, то вероятность того, что он принадлежит интервалу Δx , равна $f(x) \Delta x$, поскольку полная частота равна единице и $f(x) \Delta x$ есть доля элементов, принадлежащих интервалу Δx . Переходя к пределу, мы можем сказать, что вероятность полученному значению не превзойти x_0 при случайном выборе элемента из непрерывной генеральной совокупности равна

$\int_{-\infty}^{x_0} dF = F(x_0)$. Следует, однако, помнить, что характер предельного перехода здесь должен быть точно установленным.

Всюду ниже под случайным выбором из непрерывной генеральной совокупности с распределением $dF = f(x)dx$ будет пониматься «предельный» случайный выбор при убывании Δx .

7.13 Выше мы говорили о выборе элементов «случайным» образом. В математической теории вероятностей случайность обычно определяется в терминах самой вероятности. Говорят, что элемент генеральной совокупности выбран случайно, если он выбран случайным методом; случайным же метод называется тогда, когда посредством него любой элемент генеральной совокупности выбирается равновероятно. Понятие случайности чрезвычайно важно в выборочной теории, и мы рассмотрим его довольно подробно в главе 9. В настоящий момент ограничимся замечанием, что, когда мы говорим о случайном выборе, мы в действительности имеем в виду метод выбора, наделяющий некоторые события одинаковыми вероятностями, и, следовательно, позволяющий применять теорию вероятностей *a priori*. Справедливость этого устанавливается, в конечном счете, эмпирически. На практике обнаружено, что существуют выборочные процессы, посредством которых элементы из генеральной совокупности извлекаются таким образом, что составляющие события могут рассматриваться как равновероятные; в выборочной теории часто встречаются выборки, порожденные такими процессами.

7.14 Величины, принимающие разные значения с различными вероятностями, привели к отдельному понятию — понятию случайной величины. В обычном анализе оно не появляется. Говоря о «переменной величине», имеют в виду числа, которые могут быть любыми из некоторого интервала; при этом не ставятся вопросы, следует ли одно из них рассматривать более часто, чем другие, и встречаются ли некоторые из них более часто в практике. Случайную величину, напротив, следует рассматривать как определяемую функцией распределения: случайная величина может принимать любое значение из заданного интервала, но значения эти выделяются функцией распределения.

7.15 Рассмотрим, что означает сложение случайных величин. В обычном анализе по заданным двум переменным ξ и η можно определить третью переменную

$$\zeta = \xi + \eta;$$

сложение означает просто, что если $\xi = x$ и $\eta = y$, то $\zeta = x + y$. Можно ли придать какой-либо разумный смысл ζ , когда ξ и η — случайные величины?

Если функция совместного распределения величин ξ и η есть $F_{1,2}$, то частота одновременного выполнения неравенств $\xi \leq x$, $\eta \leq y$ равна $F_{1,2}(x, y)$. Рассмотрим какое-нибудь значение z . Мы можем определить по $F_{1,2}$ частоту события $x + y \leq z$. Она равна интегралу

$$\iint dF_{1,2}(x, y).$$

взятому по области, в которой $x + y \leq z$.

Определяемая этим интегралом функция от z является функцией распределения, поскольку она не убывает, равна нулю при $z = -\infty$ и 1 при $z = +\infty$. Суммой ζ случайных величин ξ и η назовем случайную величину, задаваемую этой функцией распределения.

7.16 Более общим образом, пусть имеются случайные величины $\xi_1, \xi_2, \dots, \xi_n$ и пусть $x_1, x_2, \dots, x_n$ — их значения. С помощью уравнения

$$z = z(x_1, x_2, \dots, x_n) \quad (7.15)$$

можно определить случайную величину ζ , принимающую значения z : величина ζ задается функцией распределения, которая в (произвольной) точке z равна интегралу от $dF(x_1, x_2, \dots, x_n)$, взятому по множеству значений $x_1, x_2, \dots, x_n$, удовлетворяющему неравенству $z(x_1, x_2, \dots, x_n) \leq z$.

В частности, сумму случайных величин можно определить с помощью уравнения

$$z = x_1 + x_2 + \dots + x_n. \quad (7.16)$$

Мы можем даже написать

$$\zeta = \xi_1 + \xi_2 + \dots + \xi_n. \quad (7.17)$$

Запись (7.17) оправдывается тем фактом, что для таких выражений верны алгебраические законы коммутативности и ассоциативности; это нетрудно понять, убедившись, что

$$\xi_1 + \xi_2 = \xi_2 + \xi_1 \quad \text{и} \quad \xi_1 + (\xi_2 + \xi_3) = (\xi_1 + \xi_2) + \xi_3.$$

7.17 Обозначение (7.17) таит в себе известную опасность. Некоторые авторы предпочитают избегать использования слова «сумма» и говорить о свертке случайных величин. Это выражение, возможно, предпочтительней, однако оно не стало общепринятым среди английских авторов.

Мы не придали пока никакого смысла выражению типа $-\eta$. Наиболее естественное определение, которое мы и примем,

состоит в том, что в качестве $-\eta$ берется случайная величина, значения которой равны значениям величины η с обратным знаком. Принятое обозначение имеет свой недостаток: исходя из него, можно подумать, что если случайная величина ζ есть сумма двух независимых случайных величин ξ , η и из ζ вычитается ξ , то в результате получится η . Это неверно. В примере 7.5 мы увидим, что $\xi + \eta_1$ может иметь одинаковое распределение с $\xi + \eta_2$ (где $\eta_2 \neq \eta_1$), даже если случайные величины ξ , η_1 , η_2 независимы.

7.18 Отметим один весьма простой, но значительный результат о характеристических функциях независимых случайных величин. Если имеются n таких величин с распределениями соответственно $dF_1, \dots, dF_n$, то элемент частоты их суммы $\zeta = \xi_1 + \xi_2 + \dots + \xi_n$ равен интегралу от $dF_1 \dots dF_n$ по элементу объема, заключенному между z и $z + dz$. Следовательно, характеристическая функция суммы ζ , будучи интегралом от e^{itz} по области изменения z , равна

$$\begin{aligned} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{itz} dF_1 \dots dF_n = \\ = \int_{-\infty}^{\infty} e^{itx_1} dF_1 \int_{-\infty}^{\infty} e^{itx_2} dF_2 \dots \int_{-\infty}^{\infty} e^{itx_n} dF_n = \varphi_1 \varphi_2 \dots \varphi_n. \quad (7.18) \end{aligned}$$

Таким образом, х. ф. суммы независимых случайных величин равна произведению х. ф. и производящая функция семинвариантов суммы равна сумме производящих функций семинвариантов.

В нижеследующем примере 7.5 устанавливается, что две независимые случайные величины ξ и η могут иметь характеристические функции, совпадающие в некотором интервале, содержащем нуль. Если х. ф. ξ обращается в нуль вне этого интервала, то из (7.18) следует, что сумма двух независимых случайных величин, имеющих такое же распределение, как ξ , распределена одинаково с $\xi + \eta$.

Пример 7.5

Интегрируя обе части равенства

$$\int_0^{\infty} \frac{\sin ax}{x} dx = \frac{1}{2} \pi, \quad a > 0,$$

по α , имеем

$$\int_0^{\infty} \frac{\cos \beta x - \cos \alpha x}{x^2} dx = \frac{1}{2} \pi (\alpha - \beta), \quad \alpha > \beta.$$

Теперь легко показать, дважды используя тождества вида

$$\sin \theta \cos m\theta = \frac{1}{2} \{ \sin(m+1)\theta - \sin(m-1)\theta \},$$

что

$$\frac{2}{\pi} \int_0^{\infty} \frac{1 - \cos x}{x^2} \cos tx \, dx = \begin{cases} 1 - |t| & \text{при } |t| \leq 1, \\ 0 & \text{при } |t| > 1. \end{cases}$$

Следовательно распределение,

$$dF = \frac{1}{\pi} \frac{1 - \cos x}{x^2} dx, \quad -\infty \leq x \leq \infty,$$

имеет х. ф.

$$\varphi_1(t) = \begin{cases} 1 - |t|, & |t| \leq 1, \\ 0, & |t| > 1, \end{cases} \quad (7.19)$$

поскольку

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos x}{x^2} \sin tx \, dx = 0$$

в силу нечетности функции $\sin tx$. С другой стороны, пользуясь рядом Фурье

$$\cos \theta + \frac{1}{3^2} \cos 3\theta + \frac{1}{5^2} \cos 5\theta + \dots = \begin{cases} \frac{1}{8} (\pi^2 - 2\pi\theta), & 0 \leq \theta \leq \pi, \\ \frac{1}{8} (\pi^2 + 2\pi\theta), & -\pi \leq \theta \leq 0, \end{cases}$$

можно показать, что дискретное распределение

$$\left. \begin{aligned} f(0) &= \frac{1}{2}, \\ f(x) &= \frac{2}{x^2}, \quad x = (2m+1)\pi, \quad m=0, \pm 1, \pm 2, \dots, \end{aligned} \right\} \quad (7.20)$$

имеет периодическую с периодом 2π х. ф., равную

$$\varphi_2(t) = 1 - |t|, \quad |t| \leq 1. \quad (7.21)$$

Из (7.19) и (7.21) следует, что $\varphi_1(t)\varphi_2(t) = \varphi_1(t)\varphi_1(t)$ для всех t . Таким образом, сумма первой и второй случайных величин распределена одинаково с суммой двух случайных величин, одинаково распределенных с первой. Поэтому, вообще говоря, неверно, что если $\xi + \eta_1 = \xi + \eta_2$, то $\eta_1 = \eta_2$.

Пример 7.6

Утверждение, обратное к предложению 7.18, неверно. Если ξ и η независимы, то х. ф. их совместного распределения равна произведению соответствующих им х. ф. функций и *обратно*, однако из равенства х. ф. их суммы произведению х. ф. каждой из них в отдельности не следует их независимость.

Рассмотрим распределение

$$dF = \left\{ \frac{1}{2\pi} x^{-\frac{1}{2}} y^{-\frac{1}{2}} e^{-\frac{1}{2}(x+y)} + \right. \\ \left. + \varepsilon(x-y)(xy - x - y + 2)e^{-(x+y)} \right\} dx dy, \quad (7.22) \\ 0 \leq x, \quad y \leq \infty.$$

Очевидно, существует такое, возможно малое, ε , для которого заключенный в скобки член в правой части (7.22) неотрицателен при всех допустимых x и y . Поэтому (7.22) действительно может представлять функцию плотности распределения. Далее, если поменять местами x и y в члене, содержащем ε , то он изменит знак; следовательно, интеграл от него по области $x \geq 0, y \geq 0$ равен нулю. Полная частота равна поэтому интегралу от первого члена, т. е. $\left\{ 2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right) \right\}^2 / (2\pi) = 1$.

Х. ф. x и y получается интегрированием $\exp(it_1x + it_2y)$ по (7.22). Подсчет показывает, что

$$\varphi(t_1, t_2) = \frac{1}{(1 - 2it_1)^{\frac{1}{2}} (1 - 2it_2)^{\frac{1}{2}}} + \frac{2\varepsilon t_1 t_2 (t_1 - t_2)}{(1 - it_1)^3 (1 - it_2)^3}. \quad (7.23)$$

Для нахождения х. ф. x полагаем в (7.23) $t_2 = 0$. Получим

$$\varphi_1(t_1) = (1 - 2it_1)^{-\frac{1}{2}}.$$

Аналогично для х. ф. y , положив $t_1 = 0$, получим

$$\varphi_2(t) = (1 - 2it)^{-\frac{1}{2}}.$$

Полагая $t_1 = t_2$, находим, что х. ф. $x+y$ равна $(1 - 2it)^{-1}$, т. е. совпадает с произведением *) $\varphi_1(t)$ и $\varphi_2(t)$. Соотношение между

*) Можно указать и более простой пример. Пусть ξ подчиняется распределению Коши $dF = [\pi(1+x^2)]^{-1} (|x| < \infty)$. В таком случае х. ф. ξ есть $\exp(-|t|)$ и, значит, х. ф. $\xi + \xi = 2\xi$ есть $\exp(-2|t|)$. С другой стороны, если ξ и η независимы и подчиняются указанному выше распределению Коши, то х. ф. суммы $\xi + \eta$ также равна $\exp(-2|t|)$, (Прим. ред.)

характеристическими функциями оказывается таким же, как и в случае $\varepsilon=0$. Это происходит потому, что член в (7.23), содержащий ε , имеет множители t_1 , t_2 и $(t_1 - t_2)$.

7.19 Несмотря на некоторую опасность терминологического характера, в дальнейшем мы часто не будем делать различия между «случайной величиной» и «значениями случайной величины». Под «величиной» следует понимать обычную (неслучайную) величину, если только из контекста не следует, что имеется в виду случайная величина. В тех случаях, когда возможно недоразумение, мы будем указывать, что речь идет о случайной величине. Чтобы избежать различия между случайными величинами и их значениями (соответственно ξ и x), для обозначения случайных величин часто будут использоваться латинские буквы, а сама случайная величина там, где позволяет контекст, будет именоваться вариантой.

Выборочные распределения

7.20 Мы уже отмечали, что если из генеральной совокупности случайным образом выбирается элемент, то вероятность того, что соответствующее значение варианты не превосходит x , равна значению в точке x функции распределения $F(x)$. Подобным образом, если выбирается элемент из многомерной генеральной совокупности, то вероятность того, что соответствующее ему значение первой варианты не превосходит x_1 , второй — не превосходит x_2 , ..., n -й не превосходит x_n , равна значению в точке $(x_1, x_2, \dots, x_n)$ функции распределения $G(x_1, x_2, \dots, x_n)$. Далее, если варианты независимы и r -я варианта имеет функцию распределения $F_r(x_r)$, то указанная вероятность равна

$$F_1(x_1)F_2(x_2) \dots F_r(x_r).$$

Предположим теперь, что имеется выборочный процесс (ниже называемый нами выбором), который, будучи применен к (одномерной) генеральной совокупности, извлекает группу из n элементов. Если этот процесс повторяется, то он порождает многомерное распределение: каждый выбор доставляет n значений $x_1, x_2, \dots, x_n$. Природа этого многомерного распределения зависит от выборочного процесса и от генеральной совокупности. Его функция распределения G в точке $(x_1, \dots, x_n)$ равна вероятности того, что соответствующее случайному выбору значение первой величины не превосходит x_1 , второй — x_2 и т. д. Значения ξ можно поставить в соответствие n случайным величинам $\xi_1, \xi_2, \dots, \xi_n$.

Особенно важным типом выборочного процесса для статистической теории является тот, при котором функция распределе-

ния $G(x_1, x_2, \dots, x_n)$ распадается в произведение $F_1(x_1)F_2(x_2)\dots F_n(x_n)$. Выбор в этом случае называется простым. Распределения значений $x_1, x_2, \dots, x_n$ независимы друг от друга, и можно сказать, что выбор всякого элемента не зависит от выбора любого другого. Кроме того, если выбор случаен, то все функции $F_r(x)$ равны одной и той же функции $F(x)$ — функции распределения, соответствующей генеральной совокупности, из которой извлекается выборка. Таким образом, распределение значений, получаемых в выборках объема n применением простого случайного метода, задается выражением

$$dG(x_1, x_2, \dots, x_n) = dF(x_1)dF(x_2)\dots dF(x_n) = f(x_1)f(x_2)\dots f(x_n)dx_1dx_2\dots dx_n, \quad (7.24)$$

и вероятность того, что первое значение в такой выборке не превосходит x_1 и т. д., равна $F(x_1)F(x_2)\dots F(x_n)$. Заметим еще, что поскольку иксы входят в (7.24) симметрично, их порядок несуществен.

7.21 Пусть имеется выборка объема n и $x_1, x_2, \dots, x_n$ — соответствующие варианты. Как в (7.15), можно рассмотреть распределение некоторой функции от них $z=z(x_1, x_2, \dots, x_n)$, например их среднее или дисперсию. Распределение соответствующей случайной величины ξ называют выборочным распределением ξ (или z). Вероятность того, что значение ξ , получаемое при случайном выборе, не превосходит z , равна значению функции распределения для $z=z(x_1, x_2, \dots, x_n)$ в z . Для простого случайного выбора она получается интегрированием (7.24) по множеству тех иксов, для которых $z(x_1, x_2, \dots, x_n) \leq z$.

Пример 7.7

Пусть осуществляется простой случайный выбор объема два из нормальной генеральной совокупности

$$dF = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}, \quad -\infty \leq x \leq \infty.$$

В силу (7.24) распределение выбираемых значений, скажем x_1 и x_2 , задается формулой

$$dF = \frac{1}{2\pi\sigma^2} \exp\{-(x_1^2 + x_2^2)/2\sigma^2\} dx_1 dx_2. \quad (7.25)$$

Найдем интеграл от (7.25) по множеству $\frac{1}{2}(x_1 + x_2) \leq z$. Этот интеграл, как функция от z , есть функция распределения среднего выборки.

Положим

$$u = \frac{1}{2}(x_1 + x_2), \quad v = \frac{1}{2}(x_1 - x_2).$$

Распределение переходит при этом в

$$dF = \frac{1}{\pi\sigma^2} \exp \{-(u^2 + v^2)/\sigma^2\} du dv, \quad (7.26)$$

так что u и v независимы. Проинтегрировав по v , находим распределение для u :

$$dF = \frac{1}{\sigma\sqrt{2\pi}} \exp(-u^2/\sigma^2) du. \quad (7.27)$$

Таким образом, вероятность того, что $\frac{1}{2}(x_1 + x_2)$ не превосходит z , равна

$$\frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^z e^{-u^2/\sigma^2} du. \quad (7.28)$$

Иначе говоря, u распределено нормально с дисперсией $\frac{1}{2}\sigma^2$.

7.22 Оставшаяся часть этой главы посвящается некоторым важным для статистики теоремам о пределах сумм n независимых случайных величин при возрастании n . Установим прежде всего простой, но эффективный результат, принадлежащий Чебышеву, о частном случае которого говорилось в 3.35.

Пусть x — случайная величина и $g(x)$ — неотрицательная функция от x . Тогда для любого $k > 0$

$$\mathbf{P}\{g(x) \geq k\} \leq \mathbf{M}\{g(x)\}/k.$$

В самом деле, обозначив буквой R множество, на котором $g(x) \geq k$, имеем

$$\mathbf{M}\{g(x)\} = \int_{-\infty}^{\infty} g(x) dF \geq \int_R g(x) dF \geq k \int_R dF = k\mathbf{P}\{g(x) \geq k\}. \quad (7.29)$$

В частности, при $g(x) = \{x - \mathbf{M}(x)\}^2$ (7.29) сводится к неравенству Бьенэме — Чебышева (3.95).

Закон больших чисел

7.23 Пусть $\xi_1, \xi_2, \dots$ — последовательность независимых одинаково распределенных случайных величин с математическим ожиданием μ . Обозначим $\bar{\xi}_n$ среднее первых n из них:

$$\bar{\xi}_n = \frac{1}{n} \sum_{i=1}^n \xi_i.$$

При всех n математическое ожидание $\bar{\xi}_n$ равно μ . Закон больших чисел утверждает, грубо говоря, что при возрастании n величины $\bar{\xi}_n$ располагаются все ближе и ближе к μ . Более точно, для любого заданного $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P \{ |\bar{\xi}_n - \mu| > \varepsilon \} = 0. \quad (7.30)$$

Если предположить существование у ξ_1 дисперсии, равной, скажем, σ^2 , то доказать (7.30) очень просто. В этом случае дисперсия $\bar{\xi}_n$ существует и равна σ^2/n , так что в силу неравенства Бьенэме — Чебышева

$$P \{ |\bar{\xi}_n - \mu| > \varepsilon \} \leq \sigma^2/n\varepsilon^2.$$

Для заданного ε стоящая слева вероятность будет, следовательно, сколь угодно малой, если n достаточно велико.

7.24 Теорема остается верной и без условия существования σ^2 . Приведенное выше доказательство легко перенести на тот случай, когда существует $M(|\xi|^{\alpha})$ при $\alpha > 1$; если известно только существование $M(|\xi|)$, то требуется иное доказательство (см. упражнение 7.20). Имеются также обобщения на случай разнораспределенных слагаемых.

Усиленный закон больших чисел

7.25 Закон больших чисел утверждает наличие некоторого предельного свойства у сумм случайных величин. Так называемый «усиленный» закон

больших чисел относится к поведению последовательности $\Sigma_n \equiv \sum_{i=1}^n \xi_i$

при всех значениях n ; он утверждает как бы нечто о свойствах Σ_n по пути к пределу. В самом деле, он состоит в том, что для любых положительных ε и δ существует N такое, что, каково бы ни было $M > 0$,

$$P \left\{ \max_{N \leq n \leq N+M} |\bar{\xi}_n - \mu| \geq \varepsilon \right\} \leq \delta. \quad (7.31)$$

Закон больших чисел утверждает, что $|\bar{\xi}_n - \mu|$, в конце концов, становится малым, но не утверждает малость каждого значения. Возможно, что при некоторых n значения велики, однако такие случаи не могут встречаться часто. Согласно усиленному закону больших чисел вероятность такого события чрезвычайно мала. В случае одинаково распределенных величин усиленный закон имеет место при единственном неограничительном условии существования μ ; для неодинаково распределенных величин нужны еще другие дополнительные условия.

Центральная предельная теорема

7.26 В том случае, когда существуют дисперсии, можно доказать намного более точную теорему. Именно, оказывается, что распределения средних $\bar{\xi}_n$ становятся почти нормальными с математическим ожиданием μ и дисперсией σ^2/n . Более точно,

для всех t_1, t_2

$$\lim P \left\{ t_1 \leq \frac{\bar{\xi}_n - \mu}{\sigma/\sqrt{n}} \leq t_2 \right\} = \int_{t_1}^{t_2} \alpha(t) dt, \quad (7.32)$$

где $\alpha(t)$ — функция плотности нормального распределения. Теорема эта верна и для разнораспределенных ξ ; μ в этом случае заменяется на среднее из математических ожиданий слагаемых, а $\sigma/\sqrt{n}$ — на стандартное отклонение $\bar{\xi}_n$.

Предположим, что все ξ имеют конечный абсолютный момент третьего порядка и что, кроме того, $\left(\sum_1^n v_{3r} \right) / \sigma_n^3 \rightarrow 0$ при $n \rightarrow \infty$,

где $\sigma_n^2 = \sum_1^n \mu_{2r}$ (индекс r при моменте означает, что имеется в виду соответствующий момент случайной величины ξ_r). Для х. ф. φ_r величины $\xi_r - \mu'_{1r}$ имеет место разложение (см. 3.15)

$$\varphi_r \left(\frac{t}{\sigma_n} \right) = 1 - \frac{\mu_{2r}}{\sigma_n^2} \frac{t}{2} + \rho \frac{v_{3r}}{\sigma_n^3} \frac{|t|^3}{\sigma} \quad (7.33)$$

(здесь и ниже ρ — число, не превосходящее по модулю 1). Так как согласно 3.6 $(\mu_{2r})^{1/2} \leq (v_{3r})^{1/3}$, то из сделанного относительно моментов предположения вытекает, что для любого t правая часть (7.33) при возрастании n стремится к 1 равномерно по $r \leq n$. Отсюда при достаточно больших n

$$\log \varphi_r \left(\frac{t}{\sigma_n} \right) = - \frac{\mu_{2r}}{\sigma_n^2} \frac{t^2}{2} (1 + o(1)) + 2\rho \frac{v_{3r}}{\sigma_n^3} \frac{|t|^3}{\sigma},$$

и, следовательно, обозначая $\psi(t)$ логарифм х. ф. нормированной суммы $\left(\sum_1^n \xi_r - \mu'_{1r} \right) / \sigma_n$, имеем

$$\psi(t) = \sum_1^n \log \varphi_r \left(\frac{t}{\sigma_n} \right) = - \frac{t^2}{2} (1 + o(1)) + \frac{\rho}{3} |t|^3 \frac{\sum_1^n v_{3r}}{\sigma_n^3} \rightarrow - \frac{t^2}{2},$$

так что распределение этой нормированной суммы стремится к нормированному нормальному распределению. Полученный результат принадлежит Ляпунову (1901).

7.27 Центральную предельную теорему можно доказать при условии, не включающем в себя требование существования третьих моментов. Имеет место следующий результат, принадлежащий Линдбергу и Феллеру: для сходимости распределений нормированных сумм к нормированному нормальному

распределению и $\max_{1 \leq r \leq n} \mu_{2r}/\sigma_n^2 \rightarrow 0$ необходимо и достаточно, чтобы

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma_n^2} \sum_{r=1}^n \int_{|x - \mu'_{1r}| > \varepsilon \sigma_n} (x - \mu'_{1r})^2 dF_r = 0, \quad (7.34)$$

где $\sigma_n^2 = \sum \mu_{2r}$. Из стремления $\max_{1 \leq r \leq n} \mu_{2r}/\sigma_n^2$ к нулю вытекает, что ни одно из слагаемых не доминирует над остальными. Сходности к нормальному распределению может и не быть, если случайные величины не обладают вторыми моментами. Например, в главе 11 мы увидим, что среднее из n величин, каждая из которых имеет распределение Коши

$$dF = \frac{1}{\pi} \frac{dx}{1+x^2}, \quad -\infty \leq x \leq \infty, \quad (7.35)$$

тоже имеет распределение Коши. Этот результат легко получить, используя х. ф. (см. пример 4.2).

7.28 Центральная предельная теорема играет важную роль в статистической теории. Мы уже встречались с примерами ее применимости в случаях стремления биномиального и пуассоновского распределений к нормальному. Много других примеров встретится нам в дальнейшем. Центральную предельную теорему можно получить также для последовательностей зависимых случайных величин. С общим изложением этих вопросов можно познакомиться по книге Феллера (1950).

7.29 Вышеизложенные методы могут быть распространены на многомерные случайные величины. Пусть, например, ξ и η имеют совместное распределение. Множество из n элементов, извлеченных случайным образом из этого распределения, представляет собой n пар значений случайных величин: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Можно задаться вопросом, каковы распределения таких выборочных характеристик, как, например, смешанный момент $z = \sum_{i=1}^n x_i y_i / n$. Распределение смешанного момента получается интегрированием

$$dF(x_1, y_1) dF(x_2, y_2) \dots dF(x_n, y_n)$$

по множеству $\sum x_i y_i / n \leq z$. Обобщения на случай большего числа переменных очевидны.

7.30 Важно также отметить, что центральная предельная теорема может быть доказана и для многомерных случайных величин. В частности, аналогично теореме из 7.26 распределение нормированных сумм иксов и игреков с увеличением n стремится к двумерному нормальному распределению при схожих весьма общих условиях.

УПРАЖНЕНИЯ

7.1 Пусть $E_1, E_2, \dots, E_n$ представляют собой n совместимых событий (n событий совместимы, если любое число их от 0 до n может произойти одновременно). Пусть p_i — вероятность осуществления E_i (независимо от осуществления или неосуществления других событий), p_{ij} — вероятность совместного осуществления E_i и E_j (независимо от остальных), и т. д. Показать, что вероятность того, что ни одно из событий не произойдет, равна

$$1 - \sum p_i + \sum p_{ij} - \sum p_{ijk} + \dots + (-1)^n p_{12\dots n},$$

где суммирование ведется по всем возможным различным значениям индексов.

7.2 Имеется n совместимых событий. Доказать, что вероятность осуществления ровно r из них равна

$$\sum p_{(r)} - \binom{r+1}{1} \sum p_{(r+1)} + \binom{r+2}{2} \sum p_{(r+2)} + \dots + (-1)^{n-r} \binom{n}{n-r} p_{(n)},$$

где $p_{(j)}$ означает p с j индексами.

7.3 Показать, что по меньшей мере r событий из n совместимых событий осуществляются с вероятностью

$$\sum p_{(r)} - \binom{r}{1} \sum p_{(r+1)} + \binom{r+1}{2} \sum p_{(r+2)} + \dots + (-1)^{n-r+1} \binom{n-1}{n-r} p_{(n)},$$

где $p_{(j)}$ означает то же, что и в предыдущем упражнении.

7.4 Решить задачу о конвертах из примера 7.3 с помощью результата упражнения 7.2. Показать также, что в точности r писем попадают в соответствующие им конверты с вероятностью

$$\frac{1}{n!} \left\{ \binom{n}{r} (n-r)! - \binom{r+1}{1} \binom{n}{r+1} (n-r-1)! + \dots \right\} = \frac{1}{r!} \left\{ 1 - 1 + \frac{1}{2!} - \dots + \frac{(-1)^{n-r}}{(n-r)!} \right\}.$$

7.5 Имеется n совместимых независимых событий $E_1, \dots, E_n$. Вероятность того, что осуществляется только E_i , равна α_i . Показать, что вероятность p_i осуществления E_i (вне зависимости от осуществления или неосуществления остальных событий) дается формулой

$$p_i = \alpha_i / (\alpha_i + t),$$

где t — корень уравнения

$$\prod_{i=1}^n (\alpha_i + t) = t^{n-1}.$$

Показать также, что в общем случае имеются два значения p_i , удовлетворяющих требуемым условиям.

7.6 Из кучки, состоящей из неизвестного количества N фишек, игрок выбирает случайным образом горсть, содержащую n фишек. Проанализировать следующее рассуждение. Число N одинаково вероятно может быть четным и нечетным. Если оно нечетно, то вероятность n быть нечетным больше $1/2$,

а если оно четно, то эта вероятность равна $1/2$. Таким образом, вероятность того, что n нечетно, больше чем $1/2$, и игроку следует ставить на вынимание нечетного числа.

7.7 С течением времени t происходят некоторые события. Количество событий, происходящих в интервале времени (t_1, t_2) , зависит лишь от длины этого интервала; количества событий, происходящих в непересекающиеся промежутки времени, независимы. Показать, что вероятность того, что за время t не произойдет ни одного события, равна $e^{-\lambda t}$, где $0 < \lambda \leq \infty$.

7.8 Показать, что если a равномерно распределена в интервале $0 \leq a \leq 6$, а b — в интервале $0 \leq b \leq 9$, то с вероятностью $1/3$ уравнение $x^2 - ax + b = 0$ имеет два действительных корня.

7.9 На плоскости случайным образом выбирается прямая

$$x \cos \theta + y \sin \theta = p$$

так, что элементы $d\theta dp$ равновероятны. Показать, что если эта прямая пересекает замкнутую выпуклую кривую длины l_1 , то вероятность пересечения ею другой замкнутой выпуклой кривой длины l_2 , лежащей внутри области, ограниченной первой кривой, равна l_2/l_1 *).

Вывести отсюда, что если прямая пересекает окружность, то всякий фиксированный диаметр этой окружности она пересекает с вероятностью $1/\pi$.

7.10 Пусть выполнены условия предыдущего упражнения. Показать, что если заменить координаты, осуществив сдвиг начала и поворот вокруг него, то $dp'd\theta = dpd\theta$, где штрихи относятся к величинам в новых координатах. Иначе говоря, показать, что «случайность» прямой, выбираемой указанным образом, не зависит от фиксации системы координат.

7.11 На окружности случайно выбираются три точки. Показать, что на одной и той же полуокружности они оказываются с вероятностью $3/4$. (Считать, что в пределе элементарные отрезки дуги равновероятны.)

Объяснить ошибку в следующем рассуждении: пара точек *всегда* лежит на одной полуокружности, кончающейся в одной из них; вероятность того, что третья точка лежит на этой же полуокружности, равна $1/2$, следовательно, вероятность тем трем точкам оказаться на одной полуокружности равна $1/2$.

Рассмотреть также следующее рассуждение. Независимо, где выбрана первая точка. Представим себе, что окружность разрезана в этой точке и развернута в горизонтальный отрезок. Вероятность того, что вторая и третья

точки попадут в левую половину отрезка, равна $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$; аналогично для правой половины. Следовательно, три точки оказываются в одной полуокружности с вероятностью $1/2$.

7.12 Доказать, что если x , y — независимые случайные величины с характеристическими функциями ϕ_1 , ϕ_2 соответственно, то х. ф. величины $x - y$ равна $\phi_1(t)\phi_2(-t)$. Вывести отсюда, что если $x + y$ и $x - y$ распределены одинаково, то распределение y симметрично относительно нуля **).

*) p и 0 считаются здесь изменяющимися в конечных пределах, причем верхняя граница изменения $|p|$ должна быть достаточно большой. (Прим. перев.)

**) Это утверждение неточно: оно неверно без дополнительных предположений (например, $\phi_1 \neq 0$). В самом деле, если взять в качестве x величину с х. ф., равной нулю вне некоторого интервала, содержащего нуль (например, (7.19)), а в качестве y — величину с действительной х. ф., действительной в этом интервале, то $x + y$ и $x - y$ распределены одинаково, но распределение y несимметрично. Приведем здесь принадлежащий И. В. Островскому пример действительной х. ф., действительной на отрезке $-A \leq t \leq A$, $A > 0$. Пусть $\theta(t) \neq 0$ — вещественная, нечетная, дважды непрерывно дифференцируе-

7.13 Величины x и y имеют двумерное нормальное распределение вида (5.90). Найти х. ф. суммы $x+y$ и показать, что эта сумма распределена нормально с дисперсией $2(1+\rho)$. Вывести, что $|\rho| \leq 1$.

7.14 Показать, что сумма независимых случайных величин, каждая из которых имеет пуассоновское распределение (параметры могут быть различными), тоже распределена по Пуассону.

7.15 Пусть x — нормированная нормально распределенная случайная величина и $y=x^2$. Показать, что х. ф. совместного распределения x и y равна

$$\varphi_{x,y}(t_1, t_2) = (1 - 2it_2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \frac{t_1^2}{1 - 2it_2} \right\},$$

и получить из этой формулы х. ф. распределений x и y в отдельности. Показать также, что ковариация x и y равна нулю.

7.16 Из генеральной совокупности, которой соответствует неотрицательная варианта со средним μ , производится случайная выборка объема n . Пусть $x_1, x_2, \dots, x_n$ — наблюдаемые значения варианты. Показать, что ве-

роятность величине $\sum_{i=1}^n x_i$ превзойти некоторое число λ не превышает $n\mu/\lambda$.

7.17 Показать, что распределение среднего (скажем, u) двух независимых случайных величин, каждая из которых имеет распределение

$$dF = \frac{3}{2} x^2 dx, \quad -1 \leq x \leq 1,$$

есть

$$dF = \frac{3}{5} (3 - 15u + 20u^2 - 8u^5).$$

мая функция, суммируемая на всей оси $-\infty < t < \infty$ вместе со своими двумя производными и равная нулю при $-A \leq t \leq A$. Положим

$$\psi(x) = \frac{i}{2\pi} \int_{-\infty}^{\infty} \theta(t) e^{-itx} dt = \frac{1}{\pi} \int_0^{\infty} \theta(t) \sin tx dt.$$

Функция $\psi(x)$ вещественна, непрерывна и, кроме того, $\lim_{x \rightarrow \pm\infty} \psi(x) x^2 = 0$.

Поэтому $\varepsilon > 0$ можно взять столь малым, чтобы для всех x , $-\infty < x < \infty$, было справедливо неравенство

$$\frac{1}{1+x^2} + \varepsilon \psi(x) > 0.$$

Положим теперь

$$\psi(t) = k \int_{-\infty}^{\infty} \left\{ \frac{1}{1+x^2} + \varepsilon \psi(x) \right\} e^{itx} dx,$$

где $k > 0$ — постоянная, выбранная так, чтобы $\varphi(0) = 1$. Очевидно, $\varphi(t)$ — х. ф. Так как

$$\varphi(t) = \pi k e^{-|t|} + i k \varepsilon \theta(t),$$

то $\varphi(t)$ не вещественна, но при $-A \leq t \leq A$

$$\varphi(t) = \pi k e^{-|t|}.$$

(Прим. перев.)

7.18 Найти распределение суммы k независимых случайных величин, каждая из которых имеет распределение

$$dF = e^{-x} dx, \quad 0 \leq x < \infty.$$

Нарисовать графики функций плотности сумм для нескольких k и пронаблюдать приближение их к нормальной кривой.

7.19 Рассмотрим симметричную функцию $f(x)$, равную при $x \geq 0$

$$f(x) = \begin{cases} \frac{A}{x_p} (x + 1 - x_p), & x_p - 1 \leq x \leq x_p, \quad p = 1, 2, \dots; \\ \frac{A}{x_p} (-x + 1 + x_p), & x_p \leq x \leq x_p + 1, \quad p = 1, 2, \dots; \\ 0 & \text{в остальных случаях,} \end{cases}$$

где $x_p = (p+1)^2$. Показать, что при подходящем выборе A эта функция является функцией плотности. Показать также, что математическое ожидание $|x|$

не существует, но предел $\lim_{n \rightarrow \infty} \int_{-n}^n x f(x) dx$ конечен (т.е. математическое ожи-

дание x существует в смысле главного значения). Убедиться в том, что если изменить начало отсчета, то не будет существовать математического ожидания и в смысле главного значения. Вывести отсюда (беря в качестве одной из случайных величин постоянную), что если математическое ожидание принимать в смысле главного значения, то математическое ожидание суммы двух случайных величин не обязательно равно сумме их математических ожиданий.

(Фреше, 1937.)

7.20 Имеется последовательность ξ_j независимых одинаково распределенных случайных величин с х. ф. ф. Пользуясь разложением

$$\varphi(t) = 1 + \mu it + o(t),$$

показать, что

$$\lim_{n \rightarrow \infty} \psi_n(t) = \mu it,$$

где ψ_n — п. ф. с. среднего (первых) n случайных величин, и что, следовательно, эти средние имеют в пределе единичное распределение (см. пример 4.3). Иначе говоря, доказать, что для последовательности ξ_j справедлив закон больших чисел.

ВЕРОЯТНОСТЬ И СТАТИСТИЧЕСКИЕ ВЫВОДЫ

8.1 Исчисление вероятностей в форме, изложенной в предыдущей главе, дает возможность, отпавляясь от вероятностей простых событий, вычислять вероятность более сложных событий. Но на практике приходится обычно сталкиваться с обратной задачей: по результатам наблюдений мы желаем получить сведения или о той генеральной совокупности, из которой они получены, или о том механизме, который порождает наблюдаемые значения. Во втором томе мы систематически изучаем различные методы и технику статистических выводов в математической статистике. В настоящей главе мы излагаем только основы теории статистических выводов и приводим ряд ее результатов в той форме, в которой они нам понадобятся позже в этом томе.

Теорема Байеса

8.2 Пусть $q_1, \dots, q_n$ — непересекающиеся события (предложения) и H — имеющаяся информация. И пусть p — некоторое другое событие (предложение). Тогда согласно правилу умножения вероятностей (см. правило 3 в главе 7),

$$P(q_r, p|H) = P(p|H)P(q_r|p, H) = P(q_r|H)P(p|q_r, H),$$

откуда

$$P(q_r|p, H) = \frac{P(q_r|H)P(p|q_r, H)}{P(p|H)}. \quad (8.1)$$

Если мы предположим, что события q_i образуют полную систему событий, так что одно из них должно произойти, то из (8.1), суммируя по всем q , получим

$$1 = \sum_r \frac{P(q_r|H)P(p|q_r, H)}{P(p|H)}. \quad (8.2)$$

Подставляя выражение для $P(p|H)$ из (8.2) в (8.1), мы найдем

$$P(q_r|p, H) = \frac{P(q_r|H) P(p|q_r, H)}{\sum_r \{P(q_r|H) P(p|q_r, H)\}} = \frac{P(q_r, p|H)}{\sum_r P(q_r, p|H)}. \quad (8.3)$$

Это равенство известно под названием теоремы Байеса, устанавливающей, что вероятность события q_r при данных p и H пропорциональна вероятности события q_r при заданной информации H , умноженной на вероятность события p при условии q_r и H .

8.3 Предположим теперь, что некоторое событие является одним из событий, зависящих от взаимно исключающих друг друга событий $q_1, \dots, q_n$. Пусть $P(q_r|H)$ — вероятность события q_r . Каждое из этих событий может привести к некоторому событию p с вероятностями $P(p|q_r, H)$. Теорема Байеса дает нам обратные вероятности, т. е. вероятности событий q_r при условии, что произошло событие p . Вероятности $P(q_r|H)$ называются априорными, вероятности $P(q_r|p, H)$ — апостериорными, вероятность $P(p|q_r, H)$ может быть названа *правдоподобием*. Тогда теорему Байеса можно переформулировать следующим образом: апостериорная вероятность равна априорной, умноженной на правдоподобие.

Байесовский постулат

8.4 В приведенной выше форме теорема Байеса является простым логическим следствием теоремы умножения вероятностей и не вызывает сомнений. Критика же этой теоремы возникает не в связи с ее выводом, а в связи с ее применением. При практическом применении этой теоремы, когда желают выяснить, какое же из событий q в действительности имеет место, обычно подразумевают, что таким событием надо считать то, для которого апостериорная вероятность максимальна. Это эквивалентно принятию той гипотезы относительно событий q , для которой максимальна совместная вероятность q и p , как это видно из равенства (8.3). Возникающая здесь трудность вызвана тем, что вычисление апостериорной вероятности связано с необходимостью знания априорной вероятности. Но обычно эти априорные вероятности нам не известны. В этом случае, очевидно, нельзя найти и апостериорную вероятность и, следовательно, решить вопрос о том, какая из гипотез имеет место. Чтобы обойти эту трудность, Байес предлагает считать все априорные вероятности событий, относительно которых нам ничего не известно, одинаковыми. Это допущение, известное под

названием постулата Байеса, принципа равновероятности, является одним из наиболее спорных вопросов в статистической теории *). Прежде чем переходить к обсуждению этого постулата, уместно рассмотреть несколько примеров.

Пример 8.1

Урна содержит четыре шара, относительно которых известно, что они или (а) все белые, или (б) два из них белые и два черные. Вынимаем из урны случайным образом один шар и обнаруживаем, что он белый. Спрашивается, какова вероятность того, что все шары белые?

Итак, у нас имеются две гипотезы, q_1 и q_2 . При гипотезе q_1 вероятность получить белый шар равна 1, при q_2 — равна $1/2$. Из (8.1)

$$P(q_1 | p, H) = \frac{P(q_1 | H)}{P(q_1 | H) + \frac{1}{2} P(q_2 | H)},$$

$$P(q_2 | p, H) = \frac{\frac{1}{2} P(q_2 | H)}{P(q_1 | H) + \frac{1}{2} P(q_2 | H)}.$$

Согласно байесовскому постулату мы предполагаем, что

$$P(q_1 | H) = P(q_2 | H) = 1/2,$$

откуда

$$P(q_1 | p, H) = 2/3, \quad P(q_2 | p, H) = 1/3.$$

Поэтому если мы можем выбирать только между двумя возможностями (а) и (б), то мы должны выбрать ту, апостериорная вероятность которой максимальна, т. е. принять гипотезу (а).

Предположим теперь, что шар возвращен в урну и снова производится случайное извлечение шара. Если вновь вынутый шар оказался черным, то гипотеза (а) решительно отвергается. Но если же извлеченный шар белого цвета, то мы можем вычислить новую апостериорную вероятность, приняв в качестве априорной апостериорную вероятность, полученную после первого вынимания. Мы имеем тогда $P(q_1 | H) = 2/3$, $P(q_2 | H) = 1/3$, где H включает p , откуда, применяя (8.1), получим апостериорную

*) В современных формулировках этого постулата иногда допускается и не равновероятный характер априорного распределения (так, например, поступает Джеффрейс). Мы не будем здесь подробно вдаваться в объяснения, отложив их до второго тома, поскольку это слишком бы увело нас в сторону.

вероятность для q_1 и q_2 при новом событии, скажем p' ,

$$P(q_1 | p', H) = \frac{2}{3} / \left(\frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} \right) = \frac{4}{5},$$

$$P(q_2 | p', H) = \frac{1}{2} \cdot \frac{1}{3} / \left(\frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} \right) = \frac{1}{5}.$$

Ясно, что если мы повторим наш процесс заново и вновь получим белый шар, то апостериорная вероятность события (а) станет еще большей. Это, конечно, согласуется с интуитивным представлением о том, что если в процессе извлечения шаров (с возвращением) из урны мы не получаем черных шаров, то весьма правдоподобно, что этих шаров вовсе нет.

Пример 8.2

Рассмотрим предыдущий пример с несколько иной точки зрения. Предположим, что в единицу времени извлекается по одному шару (с последующим возвращением) и после n извлечений получили последовательность из n белых шаров. Вероятность этого события при гипотезе (а) равна 1, а при гипотезе (б) равна $1/2^n$. Из формулы (8.1) мы тогда получим (событие p означает, что все первые n шаров белого цвета)

$$P(q_1 | p, H) = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2} \cdot 1/2^n} = \frac{2^n}{2^n + 1}, \quad P(q_2 | p, H) = 1/(2^n + 1).$$

С ростом n вероятность $P(q_1 | p, H)$ стремится к единице, а вероятность $P(q_2 | p, H)$ — к нулю. Этот результат остается верным, какова бы ни была начальная априорная вероятность $0 < t \leq 1$ для гипотезы (а). В самом деле,

$$P(q_1 | p, H) = \frac{2^{nt}}{2^{nt} + (1-t)},$$

что стремится к единице. Этот факт также согласуется с общепринятыми представлениями: каковы бы ни были начальные вероятности, новые с каждым шагом все меньше и меньше зависят от них.

Пример 8.3

Из урны, наполненной шарами неизвестного цвета, извлекается случайно шар (с последующим возвращением). Этот процесс продолжается t раз, и каждый раз извлекаются шары черного цвета. Спрашивается, какова вероятность того, что при следующем извлечении шар будет также черным?

Так поставленный вопрос не допускает определенного ответа, поскольку возможно, что в урне находятся шары всевозможных цветов и оттенков, и поэтому нам неизвестно, какие же гипотезы надо сравнивать.

Предположим в связи с этим, что шары могут быть только двух цветов — черные и белые. Рассмотрим гипотезы: (1) все шары черные, (2) все черные за исключением одного, (3) все черные за исключением двух и т. д. Но даже и сейчас задача все еще остается неопределенной, поскольку необходимо знать общее число шаров. Будем предполагать, что их всего N . Позднее мы рассмотрим случай $N \rightarrow \infty$.

Пусть q_R — гипотеза о том, что в урне R шаров черного цвета и $N - R$ белого. Вероятность при каждом извлечении получить черный шар равна R/N и вероятность того, что подряд встретятся m черных шаров, равна $(R/N)^m$. Если все события q_R а priori равновероятны, то из (8.1) следует, что

$$P(q_R | p, H) = \frac{(R/N)^m}{\sum_{R=0}^N (R/N)^m}.$$

Вероятность получения черного шара при гипотезе q_R равна R/N . Поскольку события q_R взаимно исключают друг друга, то вероятность получить следующий шар черного цвета равна

$$\sum_{R=0}^N \frac{R}{N} P(q_R | p, H) = \frac{\sum (R/N)^{m+1}}{\sum (R/N)^m}. \quad (8.4)$$

Отсюда можно получить ответ на поставленный вопрос и в предельном случае. При $N \rightarrow \infty$ правая часть выражения (8.4) стремится к отношению определенных интегралов

$$\int_0^1 x^{m+1} dx / \int_0^1 x^m dx = \frac{m+1}{m+2}. \quad (8.5)$$

Этот результат представляет собой частный случай так называемого лапласовского правила следования. Некоторые энтузиасты этого правила иногда применяют его слишком неразборчиво. Так, они склонны считать, что из этого правила следует справедливость утверждения, что если некоторое событие наблюдалось m раз подряд, то шанс, что оно появится и в $(m+1)$ -й раз, в $(m+1)$ раз больше шанса, что оно не появится. Должно быть ясно, что это заключение незаконно.

8.5 Если встать на частотную точку зрения, то появляются принципиальные трудности, возникающие в связи с байесов-

ским постулатом. Если мы считаем, что вероятность есть мера наших сомнений, то тогда естественно полагать априорные вероятности одинаковыми, когда относительно них нам ничего неизвестно, поскольку все соответствующие предложения одинаково сомнительны. Частотная теория, однако, при применении постулата Байеса требует, чтобы события, отвечающие различным q , были равновероятны в генеральной совокупности. Многим статистикам, хотя и не всем, такое требование к Природе представляется слишком сильным. Постулат Байеса является одним из важнейших моментов в теории вероятностей. Многие приверженцы субъективной школы теории вероятностей принимают его. В то же время немало сторонников частотной школы, которые полностью его отвергают.

Все еще имеется так много разногласий по этому вопросу, что никакие точки зрения нельзя пока считать общепринятыми. Одно, однако, ясно: если кто-нибудь отвергает байесовский постулат, то он должен заменить его чем-либо другим. Проблема, которую Байес пытался решить, чрезвычайно важна для научных выводов, и едва ли представляется возможным, чтобы какая-либо научная мысль не давала некоторого решения проблемы, хотя бы интуитивного и эмпирического. Нам постоянно приходится оценивать степень верности гипотез по имеющимся данным, и борьба за существование, по словам Тиле, вынуждает нас обращаться к оракулам.

Максимальное правдоподобие

8.6 Трудности, связанные с обоснованием байесовского постулата, привели к поиску новых принципов и методов решения возникающих задач. Некоторые из этих принципов дают существенное продвижение в решении целого ряда частных задач. В частности, это относится к принципу наименьших квадратов и принципу, основанному на критерии χ^2 -квадрат, которые мы рассмотрим позже. Сейчас же мы остановимся на одном общем принципе — принципе максимального правдоподобия.

Возвращаясь к (8.3), мы можем записать теорему Байеса в такой форме:

$$P(q_r | p, H) \propto P(q_r | H) L(p | q_r, H), \quad (8.6)$$

где $L(p | q_r, H)$ — правдоподобие. Принцип максимального правдоподобия состоит в том, что при выборе одной из гипотез q_r следует выбирать ту из них, которая обращает в максимум L . Другими словами, мы принимаем ту гипотезу, которая дает наибольшую вероятность для наблюдаемого события. В то вре-

мя как байесовский подход основан на максимизации совместного распределения событий q_r и p , принцип максимального правдоподобия оперирует с обращением в максимум условной вероятности p при заданном q_r .

8.7 Нужно заметить, что принцип максимального правдоподобия не эквивалентен выбору гипотезы с наибольшей вероятностью. Некоторые защитники этого принципа полностью отрицают возможность вложить смысл в такие выражения, как «вероятность гипотез». Позже мы увидим, что практически различие между результатами, полученными с помощью принципа максимального правдоподобия и байесовского постулата, не столь существенно, как это можно было бы ожидать. Однако между этими принципами имеется принципиальное отличие.

При рассмотрении принципа максимального правдоподобия надо иметь в виду одну тонкость, нашедшую свое отражение в записи L вместо P . Обычная вероятностная функция дает вероятность события p при заданных q_r и H , при этом p варьируется, а q_r и H фиксированы. С точки зрения принципа максимального правдоподобия варьируется q_r , тогда как p и H считаются заданными. Эту функцию от q мы и называли правдоподобием, чтобы отличать ее от обычной вероятностной функции.

8.8 Предположим (как это часто случается в статистической практике), что наши гипотезы касаются значений некоторого численного параметра θ . К примеру, в случае двух альтернатив такими гипотезами могут быть следующие:

$$q_1 \equiv 0 < 0, \quad q_2 \equiv \theta \geq 0 \quad \text{или} \quad q_1 \equiv 0 = 1, \quad q_2 \equiv 0 = 2$$

и т. д.

Если θ пробегает только счетное число значений, то мы можем поставить вопрос о том, как каждому наблюдаемому событию p поставить в соответствие некоторое «наилучшее» значение θ . Байесовский метод основан на том, что «наилучшее» значение является наиболее вероятным значением. Из (8.3) мы видим, что таким значением должно быть то, которое максимизирует вероятность $P(q_r | p, H)$. Если мы ничего не знаем об априорных вероятностях $P(q_r | H)$, то в соответствии с байесовским принципом мы должны их считать одинаковыми. Тогда, как легко видеть, это эквивалентно отысканию такого q_r , для которого $L(p | q_r, H)$ максимально. Иначе говоря, в этом случае принцип максимального правдоподобия и байесовский постулат оказываются эквивалентными.

8.9 Это обстоятельство, однако, не выполнено, если параметр θ может принимать непрерывный ряд значений. В этом

случае выражения вида

$$P(q_r|H)$$

надо заменить на

$$P\left(\theta_0 - \frac{1}{2} d\theta_0 \leq 0 \leq \theta_0 + \frac{1}{2} d\theta_0 | H\right),$$

и тогда вместо (8.6) мы получим

$$\begin{aligned} P\left(\theta_0 - \frac{1}{2} d\theta_0 \leq 0 \leq \theta_0 + \frac{1}{2} d\theta_0 | p, H\right) &\propto \\ &\propto P\left(\theta_0 - \frac{1}{2} d\theta_0 \leq 0 \leq \theta_0 + \frac{1}{2} d\theta_0 | H\right) L\left(p | \theta_0 - \frac{1}{2} d\theta_0 \leq \right. \\ &\quad \left. \leq \theta \leq \theta_0 + \frac{1}{2} d\theta_0, H\right). \end{aligned} \quad (8.7)$$

Если требуется найти «наилучшую» оценку для θ , то мы должны в соответствии с байесовским постулатом взять равномерное распределение для θ и снова максимизировать L по различным значениям θ .

Предположим, что вместо различения гипотез относительно параметра θ нам надо различать, например, гипотезы о значениях некоторой функции $\varphi = \varphi(\theta)$. Скажем, вместо гипотез относительно значений дисперсии надо различать гипотезы о значениях стандартного отклонения. В этом случае мы должны рассмотреть уравнение (8.7), где вместо θ подставлено φ . При этом в соответствии с принципом Байеса мы должны считать, что априорное распределение равномерно. И тогда «наилучшая» оценка φ определяется путем максимизации L .

Но состоятелен ли такой подход? Если мы предположили, что θ имеет априори равномерное распределение, то мы не вправе предполагать то же самое и относительно φ , т. е. использование байесовского постулата представляется противоречивым. Принцип максимального правдоподобия свободен от этой неопределенности, поскольку, если $L(\theta)$ достигает максимума при некотором θ , то $L(\varphi)$ обращается в максимум при том значении φ , которое отвечает тому же самому θ . Действительно,

$$\frac{\partial L}{\partial \theta} = \frac{\partial L}{\partial \varphi} \frac{\partial \varphi}{\partial \theta}, \quad (8.8)$$

и обе части этого уравнения обращаются в нуль одновременно (если $\frac{\partial \varphi}{\partial \theta} \neq 0$).

8.10 Сделанное замечание является одним из доводов сторонников частотной школы теории вероятностей, отвергающих байесовский принцип и предпочитающих пользоваться принципом максимального правдоподобия. Но с нашей точки зрения

никакого противоречия в байесовском постулате нет, ибо оно вызвано неправильным пониманием существа дела. Будет показано, что байесовский постулат и принцип максимального правдоподобия дают тот же самый ответ в непрерывном случае, что и в дискретном, если правильно понимать предельный переход. Мы видели в 7.11, что значение вероятности в непрерывном случае существенно зависит от природы предельного перехода. Если под случайной величиной θ понимать ту случайную величину, которая получается в результате предельного процесса $\Delta\theta \rightarrow 0$, то тогда байесовский постулат будет давать другой ответ при условии, что θ рассматривается как случайная величина, получающаяся в результате аналогичного предельного процесса для φ . Несовпадение ответов объясняется точно так же, как и несовпадение ответов в примере 7.4. Поэтому кажущаяся несостоятельность байесовского постулата в действительности иллюзорна и возникает в силу неправильного понимания предельного перехода в случае непрерывных совокупностей.

Остается, конечно, верным, что во многих практических случаях мы не знаем механизма образования значений θ . Если нам требуется построить теорию различения, которая бы не зависела от нашего незнания априорного распределения для θ , то возражения против применения байесовского принципа остаются в силе и не относятся к принципу максимального правдоподобия. С другой стороны, все еще не существует убедительных доводов, почему мы должны принимать принцип максимального правдоподобия как научный принцип различения гипотез. Во втором томе мы увидим, что существует немало причин, оправдывающих применение этого принципа.

Пример 8.4

Рассмотрим пример 8.2 с позиций принципа максимального правдоподобия. Вероятность получить n белых шаров при гипотезе (а) равна единице, при гипотезе (б) равна 2^{-n} . Правдоподобие для (а) всегда больше, чем для (б), и, следовательно, мы должны выбирать гипотезу (а). Это верно при любом n . В частности, когда n возрастает, отношение правдоподобия стремится к отношению апостериорных вероятностей.

Пример 8.5

Рассмотрим n независимых результатов наблюдений, каждое из которых распределено нормально

$$dF = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} dx.$$

Если результаты наблюдений есть $x_1, x_2, \dots, x_n$, то функция правдоподобия равна

$$L = \frac{1}{\sigma^n (2\pi)^{\frac{1}{2}n}} \exp \left\{ -\frac{1}{2} \sum_{j=1}^n \left(\frac{x_j - \mu}{\sigma} \right)^2 \right\}.$$

Рассмотрим всевозможные значения μ , которые могут привести к этим наблюдениям $x_1, \dots, x_n$. Значение μ , максимизирующее L , удовлетворяет уравнению

$$\frac{\partial L}{\partial \mu} = 0, \quad \frac{\partial^2 L}{\partial \mu^2} < 0.$$

Поскольку функция L положительна, то задача нахождения «наилучшей оценки» для μ эквивалентна отысканию того значения, которое максимизирует $\log L$. Этот переход оказывается часто, как, например, здесь, более удобным. Имеем

$$\frac{\partial \log L}{\partial \mu} = \sum \left(\frac{x_j - \mu}{\sigma} \right) = 0, \quad (8.9)$$

и «наилучшая» оценка $\hat{\mu}$ для μ равна, следовательно,

$$\hat{\mu} = \bar{x} = \frac{1}{n} \sum x. \quad (8.10)$$

Легко видеть, что функция L имеет единственный максимум, поэтому оценка $\hat{\mu} = \bar{x}$ является единственной.

Если мы желаем оценивать одновременно μ и σ , то наряду с (8.9) надо рассмотреть уравнение

$$\frac{\partial \log L}{\partial \sigma} = -\frac{n}{\sigma} + \sum \frac{(x_j - \mu)^2}{\sigma^3} = 0, \quad (8.11)$$

которое дает

$$\hat{\sigma}^2 = \frac{1}{n} \sum (x - \mu)^2. \quad (8.12)$$

Заметим, что $\hat{\mu}$ не зависит от σ , тогда как $\hat{\sigma}$ зависит от μ . Выберем оценки $\hat{\mu}$ и $\hat{\sigma}$, которые одновременно удовлетворяют и (8.9) и (8.11). Тогда мы получим

$$\hat{\sigma}^2 = \frac{1}{n} \sum (x - \bar{x})^2. \quad (8.13)$$

Оценки (8.10) и (8.13) одновременно максимизируют правдоподобие.

УПРАЖНЕНИЯ

8.1 Некоторое событие постоянной вероятности $\bar{\omega}$ наблюдается r раз в n независимых испытаниях. Показать, что байесовский постулат и метод максимального правдоподобия дают для $\bar{\omega}$ оценку r/n . (При применении

теоремы Байеса можно предположить, что $\bar{\omega}$ равномерно распределена от 0 до 1.)

8.2 В условиях предыдущего упражнения испытания проводятся до тех пор, пока событие, осуществляющееся с вероятностью $\bar{\omega}$, не встретится r раз. Пусть при этом совершено n наблюдений. Показать, что и теорема Байеса, и метод максимального правдоподобия дают r/n в качестве подходящей оценки для $\bar{\omega}$.

8.3 Показать, что в упражнении 8.1 усредненное по всем выборкам значение r/n равно $\bar{\omega}$ и не равно $\bar{\omega}$ в упражнении 8.2.

8.4 В условиях предшествующих упражнений показать, что если проведено n испытаний и некоторое событие встретилось r раз, то вероятность того, что это событие встретится при следующем испытании, равна $(r+1)/(n+2)$. (Использовать более общую форму последовательного правила Лапласа.) Как это согласуется с результатом упражнения 8.1?

8.5 Некоторая урна содержит шары различных цветов, но число цветов M неизвестно. В результате извлечения n шаров с возвращением обнаружено $m (\leq n)$ различных цветов. Показать, что в качестве оценки для M метод максимального правдоподобия дает m .

8.6 В предшествующем упражнении, используя теорему Байеса, оценить M , если априорная вероятность для M пропорциональна 2^M ($M=1, 2, \dots, n$).

СЛУЧАЙНЫЙ ВЫБОР

Проблема выбора

9.1 В предыдущей главе мы несколько раз упоминали проблему выбора. Сформулирована она может быть очень просто: дана выборка из некоторой генеральной совокупности; определить по ней некоторые или все свойства этой генеральной совокупности. Уже отмечалось, что только в исключительных случаях можно делать утверждения о генеральной совокупности с полной определенностью и что, следовательно, необходимо ограничиться утверждениями менее категорического характера, выражаемыми в терминах вероятности.

9.2 Для того чтобы было возможным применять теорию вероятностей к этой проблеме, выбор должен быть случайным. Под случайной выборкой мы понимаем выборку, обладающую вычислимой для нее вероятностью. По поводу этого определения следует сделать два замечания.

Во-первых, оно содержит слово «вычислимую», а не «вычисленную», поскольку в практике часто нет необходимости в том, чтобы вероятности выбора были полностью вычислены для каждой выборки. Спецификация и контроль выборочной процедуры — это все, что требуется для применения теории вероятностей.

Во-вторых, в определении не требуется, чтобы каждая возможная выборка имела одну и ту же вероятность выбора. Если же это в самом деле так и последовательные извлечения независимы, то в соответствии с 7.20 выбор называется *простым случайным выбором*.

Для многих теоретических целей достаточно рассматривать простой случайный выбор, однако в третьем томе этой книги мы увидим, что поиск эффективности планирования выборочных исследований вселит в обдуманному отказу от простого случайного выбора. Дело здесь в том, что этот отказ не мешает применению теории вероятностей, если только (различные) ве-

роятности выбора однозначно определяются процессом выбора и, следовательно, вычислимы.

9.3. На практике мы часто встречаем неслучайные выборки, полученные некоторым целенаправленным способом. В таком случае невозможно делать точные вероятностные утверждения относительно генеральной совокупности и, когда нужно принять решение, приходится основываться на субъективных суждениях неудовлетворительного характера. Именно поэтому случайный выбор первостепенно важен при изучении по выборкам генеральной совокупности.

В дальнейшем будут рассматриваться только случайные выборки. Во избежание постоянных повторений, мы будем считать, если не оговорено противное, что все встречающиеся «выборки» и «выборочные распределения» относятся к простому случайному выбору.

9.4 Изучение случайного выбора полезно начать с рассмотрения типов генеральных совокупностей, из которых могут извлекаться выборки.

(а) Генеральная совокупность может быть конечной и реально существующей. Например: генеральная совокупность всех людей в Европе в фиксированный момент времени или генеральная совокупность яблок на данном дереве. Процесс выбора, при котором извлекается по одному элементу, если он будет длиться достаточно долго, в конце концов, исчерпает, очевидно, запас элементов. Такой выбор хотя и случаен, не является простым, поскольку последовательные извлечения не независимы.

Можно, однако, свести этот процесс к простому, возвращая выбранные элементы. Генеральная совокупность остается тогда одной и той же при каждом выборе. Первый из этих выборов иногда называют «выбором без возвращения», второй — «выбором с возвращением».

Кроме того, во многих случаях выбор можно считать приблизительно простым, даже когда не делается возвращений. Если генеральная совокупность велика по сравнению с объемом выборки, то извлечение относительно небольшого числа элементов, по существу, не меняет генеральной совокупности, которую, стало быть, можно считать приблизительно одной и той же при последовательных выборах.

(б) Выбор с возвращением из конечной генеральной совокупности в действительности может рассматриваться как выбор из бесконечной генеральной совокупности, поскольку он никогда не исчерпывает запаса элементов. Может, однако, встретиться необходимость рассмотрения бесконечной генеральной совокупности в ином смысле, именно предельной генеральной

совокупности. Мы можем, например, пожелать рассмотреть вероятность выборки из множества положительных целых чисел или из множества действительных чисел, лежащих между нулем и единицей. Последний случай появляется при выборе из непрерывного частотного распределения, когда соответствующую генеральную совокупность необходимо следует считать бесконечной.

Таким образом, если мы заменяем наблюдаемое распределение некоторым непрерывным математическим распределением, то одновременно мы заменяем конечную генеральную совокупность бесконечной. Извлечение случайных выборок из такой генеральной совокупности производится описанным в 7.11 способом — здесь должен использоваться предельный переход.

(в) В-третьих, генеральная совокупность может быть чисто гипотетической. Рассмотрим, например, бросания кости. Можно представлять себе повторные бросания как процесс выбора существующих элементов из несуществующей генеральной совокупности. В подобных случаях мы мысленно конструируем для выборки воображаемую генеральную совокупность.

Появление понятия гипотетической генеральной совокупности обуславливается концепцией частоты в теории вероятностей. Оно не нужно (и действительно полностью отвергнуто Джефрейсом) в том подходе, который рассматривает вероятность как неопределимую меру сомнения. Но если на вероятность смотреть как на относительную частоту, то, чтобы говорить о вероятности выборки типа той, которая дается бросанием кости или ростом пшеницы на участке земли, мы должны рассматривать выборку на основе некоторой генеральной совокупности.

Имеются очевидные логические трудности, когда такая выборка рассматривается как отбор — это отбор без выбора; еще большие трудности имеются относительно предположения случайности этого отбора. В самом деле, мы должны здесь попытаться вообразить, что все остальные элементы генеральной совокупности, которые лишь воображаемы, имеют одинаковую вероятность накинута мантию реальности и что, в некотором смысле, действительное событие избрано, чтобы стать реальным. Это довольно трудная концепция. В то же время следует допустить, что некоторые события, подобные бросаниям кости, происходят так, как если бы случайно выбирались элементы из некоторой реальной генеральной совокупности; соответственно представляется возможным эмпирическое обоснование понятия гипотетической генеральной совокупности,

Случайность в практике выбора

9.5 Слово «случайный», в его обычном разговорном смысле, применяется ко всякому методу выбора, не имеющему определенной цели. Мы говорим о случайном выборе из шляпы бумажных листков с написанными на них именами, о случайном выборе растения на поле пшеницы, о случайном выборе семейного бюджета из соответствующей генеральной совокупности, подразумевая при этом, что выбор происходит полностью наудачу.

Практически обнаружено, что выбор, производимый человеком, не является случайным в строгом смысле: он не доставляет одинаково часто те события, которые имеются основание считать равновероятными. Несколько примеров поясняют это обстоятельство.

Пример 9.1

При проведении одной из работ на Ротэмстэдской экспериментальной станции делались выборки растений пшеницы, по восемь растений в каждой. Шесть из них выбирались методами, которые можно считать действительно случайными и о которых речь будет идти ниже. Остальные два выбирались наудачу на глаз. Если бы в каждой из выборок все восемь растений расположили бы по величине, те два, которые выбраны на глаз, могли бы иметь любой номер от единицы до восьми; если бы эти два, как и остальные шесть, были выбраны случайно, они занимали бы эти места с примерно одинаковой частотой при большом числе выборок. Таблица 9.1 показывает действительно наблюдавшиеся частоты в двух различных случаях: а) 31 мая до образования колосьев и б) 28 июня, после того как колосья образовались.

Таблица 9.1

Распределение случайно выбранных растений по восьми категориям (Иэйтс, 1935)

Дата	Наблюдение	Числа, соответствующие определенным категориям								Полная частота
		1	2	3	4	5	6	7	8	
31 мая	Высота во время цветения	9	7	11	8	11	18	21	31	116
28 июня	Высота после образования колосьев	9	19	27	23	15	10	5	4	112

Отклонение действительных результатов от ожидаемых просто поразительно. 31 мая, до того как образовались колосья, наблюдатель предпочтительно выбирал более высокие растения, в то время как в июне он был сильно пристрастен к средним растениям, избегая выбирать низкие и высокие.

Таким образом, видно, что даже тренированный наблюдатель может осуществлять отбор с пристрастием и что это пристрастие не обязательно имеет одну и ту же направленность при различных обстоятельствах.

Пример 9.2

Таблица 9.2 показывает частоты последних цифр в большем числе измерений, произведенных четырьмя различными наблюдателями.

Сомнительно, что была какая-то действительная причина, обуславливающая более частое появление одних цифр по сравнению с другими, и можно с уверенностью предполагать, что отклонения от приблизительно одинаковой частоты появления цифр указывает на пристрастие наблюдателей.

Наблюдатель *A* оказывал предпочтение цифрам 0, 2, 8 и 9, избегая центра шкалы. Наблюдатель *B* довольно хорош — от-

Таблица 9.2

Пристрастность при регистрации результатов измерений: распределение последних цифр в измерениях, произведенных четырьмя наблюдателями (Юл, 1927)

Последняя цифра	Частота последней цифры в 1000 наблюдений			
	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
0	158	122	251	358
1	97	98	37	49
2	125	98	80	90
3	73	90	72	63
4	76	100	55	37
5	71	112	222	211
6	90	98	71	62
7	56	99	75	70
8	126	101	72	44
9	129	81	65	16
Полная частота	1 001	999	1 000	1 000

клонения его наблюдений от ожидаемых значений малы, но и он оказывал некоторое предпочтение цифре 0. Наблюдатель *C* плох — он округлял последнюю цифру примерно в одном из каждых двух наблюдений до 0 или 5. Наблюдатель *D*, очевидно, совсем плох — приблизительно в 57% его измерений последняя цифра округлена до 0 или 5.

Во всех наблюдениях производилось считывание шкалы: *A* делал измерения на чертежах с точностью до десятой миллиметра, *B*, *C* и *D* производили измерения голов живых существ с точностью до одного миллиметра. Отсюда можно заключить, что различные наблюдатели могут обладать различной степенью пристрастности даже при сравнимых обстоятельствах и что даже те из них, которые знают о возможности пристрастия и необходимости большой осторожности (как наблюдатель *A*), тем не менее не всегда могут избежать его.

Пример 9.3

Перед прибором (датчиком), состоящим из круглого диска, разделенного на десять равных секторов, занумерованных цифрами от 0 до 9, поставили наблюдателя. Диск вращался с большой скоростью, и время от времени перед ним вспыхивала электрическая лампочка на столь короткий момент, что диск казался неподвижным. Наблюдатель должен был смотреть на диск и записывать номер того сектора, который в момент вспышки находился против фиксированного указателя.

Этот прибор предназначался для получения случайных чисел и действительно их доставлял при работе с другим наблюдателем. Но наблюдатель, о котором говорилось выше, снимал показания с явным пристрастием. Частоты цифр по 10 000 произведенных им наблюдений представлены в таблице 9.3.

Таблица 9.3

Распределение цифр, снятых наблюдателем с датчика случайных чисел
(Кендалл и Бэбингтон Смит, 1939)

Цифра	0	1	2	3	4	5	6	7	8	9	Полная частота
Частота	1 083	865	1 053	884	1 057	1 007	1 081	997	1 025	918	10 000

Если бы наблюдатель был беспристрастен, то цифры появлялись бы приблизительно с одинаковой частотой; из таблицы видно, однако, что он имел пристрастие к четным цифрам и предубеждение против нечетных цифр 1, 3 и 9. Причина этого пристрастия неясна, поскольку наблюдатель не должен был

производить оценку (как в предыдущем примере), а должен был просто записать то, что он видел или думал, что видит. Объяснение, по-видимому, состоит в том, что он отдавал сильное предпочтение некоторым цифрам, т. е. в действительности видел не те цифры, или что его мозг контролировал и цензурировал зрительные восприятия. Здесь мы имеем дело с одной из чрезвычайно сильных форм психологического пристрастия.

Пример 9.4

В Англии и Уэльсе множество специальных служащих каждый год оценивают ожидаемый урожай некоторых культур. Предсказания делаются в различные периоды года, а после уборки урожая производится окончательная оценка. В таблице 9.4 представлены средние оценки урожая картофеля в различные периоды для ряда лет с 1929 по 1936.

Таблица 9.4

Отклонения при предсказании урожая: предсказания урожая картофеля в Англии и Уэльсе (тонны на акр *))
(из официальной сельскохозяйственной статистики)

Год	1 сентября		1 октября		1 ноября		Окончательная оценка (в %)
	Урожай	Отклонение от окончательной оценки (в %)	Урожай	Отклонение от окончательной оценки (в %)	Урожай	Отклонение от окончательной оценки (в %)	
1929	5,7	—17,4	6,2	—10,1	6,5	—5,8	6,9
1930	6,0	—7,7	6,1	—6,2	6,1	—6,2	6,5
1931	5,5	0,0	5,3	—3,6	5,3	—3,6	5,5
1932	6,4	—3,0	6,2	—6,1	6,3	—4,5	6,6
1933	6,4	—4,5	6,2	—7,5	6,4	—4,5	6,7
1934	6,0	—15,5	6,3	—11,3	6,7	—5,6	7,1
1935	5,6	—9,7	5,7	—8,1	6,0	—3,2	6,2
1936	6,0	—3,2	5,9	—4,8	5,8	—6,5	6,2

*) 1 акр $\cong$ 0,4 га. (Прим. перев.)

Эта таблица очень хорошо иллюстрирует постоянный пессимизм при предсказании урожая, свойственный всем соответствующим служащим в Англии (а также и в других странах). В таблице во всех случаях, кроме одного, предсказания ниже окончательной оценки. Эти служащие, по-видимому, неспособны понять на опыте, что они систематически недооценивают уро-

жай. Ничто в таблице не указывает на уменьшение разности между предсказанием и окончательной оценкой в течение рассматриваемого периода.

Следует также отметить, что приведенные в таблице оценки являются взвешенными средними большого числа независимых наблюдений. Часто встречающееся здесь неверное представление основано на предположении, что, хотя отдельные индивидуумы могут делать ошибки, их ошибки в совокупности взаимно уничтожаются. Наш пример показывает, что это, вообще говоря, неверно. Может появиться систематическое пристрастие, свойственное всем производящим оценку индивидуумам.

9.6 Предыдущие примеры достаточно хорошо показывают общий характер пристрастия наблюдателя. Тренированные наблюдатели могут проявлять пристрастие, даже сознавая свой недостаток, различные наблюдатели могут быть по-разному пристрастны при сходных обстоятельствах, один и тот же наблюдатель может быть по-разному пристрастен при различных обстоятельствах. Совершенно ясно, что следует искать истинную случайность в чем-то ином, а не в простом отсутствии целенаправленности наблюдателя. Могут встретиться люди, у которых психологические процессы столь хорошо сбалансированы, что они могут преднамеренно извлекать случайные выборки, однако небольшое число статистиков, которые экспериментировали в этом интересном направлении, обычно рассматривают себя как особо одаренных.

9.7 Как мы видели в главе 7, основная роль случайности в теории вероятностей состоит в обеспечении одинаковых вероятностей для некоторых элементарных событий. Можно сказать, что метод выбора является случайным для генеральной совокупности U , если при применении его к этой генеральной совокупности каждый элемент имеет одинаковую вероятность быть выбранным или, в терминах частоты, если при непрерывном применении этого метода к U все элементы выбираются приблизительно одинаково часто.

Но этого недостаточно. Предположим, что имеется генеральная совокупность, состоящая из двух элементов A и B , из которой производится выбор с возвращением. Метод, по которому A и B выбираются попеременно (в результате чего получается последовательность $ABAB...$), извлекает каждый элемент приблизительно одинаково часто, однако такого рода метод мы обычно не называем случайным. От случайного метода мы требуем, чтобы в подобной ситуации в результате его применения получалась последовательность, не содержащая никакой систематической упорядоченности. Не только одна-единственная

характеристика, но также и все возможные группы характеристик должны появляться одинаково часто.

9.8 Следует отметить еще одно обстоятельство. При извлечении выборки мы можем интересоваться лишь одной какой-нибудь вариантой, определяемой элементами генеральной совокупности, и может оказаться, что метод доставляет удовлетворительную случайную выборку в смысле этой варианты, но не в смысле других вариантов. Пусть, например, мы желаем сделать случайную выборку из населения некоторой определенной области. Если интересоваться такой вариантой, как цвет глаз, то, возможно, окажется достаточным выбрать некоторые дома, скажем каждый десятый, и из выбранных домов взять по одному жильцу, считая жильцов элементами выборки. Такой метод выбора не предоставляет каждому жителю одинакового шанса быть выбранным; однако если обратиться ко времени заселения домов, можно предполагать, что цвет глаз не влиял на географическое распределение людей, и, таким образом, если рассматривать распределение жителей по домам независимым от цвета глаз жителей, то выборку при исследовании цвета глаз можно считать случайной. Иначе было бы, если бы нас интересовал не цвет глаз, а доход, поскольку естественно ожидать, что дома, где живут более бедные люди, заселены более плотно, и выборка по одному человеку из каждого отобранного дома недостаточно представляла бы бедность. Такая выборка не была бы, следовательно, случайной по отношению к доходу.

Таким образом, метод, который обоснованно может считаться случайным для одной варианты, может не быть случайным для другой варианты.

Техника случайного выбора

9.9 Пусть имеется генеральная совокупность и фиксирована некоторая варианта. Как получить случайную выборку, т. е. как найти метод выбора, случайный для этой генеральной совокупности и для этой варианты? Ответ содержится частью в теории, частью в практике.

(а) Во-первых, следует требовать отсутствия явной связи между методом выбора и рассматриваемыми свойствами элементов генеральной совокупности. Этот метод и эти свойства должны быть независимы, насколько можно судить по нашим первоначальным знаниям. Например, если производить выборку на пшеничном поле, интересуясь высотой растений, то не следует употреблять метод, на который может влиять их высота. Например, нельзя осуществлять отбор, бросая обруч низко над пшеницей и затем выбирая растения, оказавшиеся внутри об-

руча после его падения (обруч может зацепиться за более высокие растения). При составлении выборки жителей некоторого города посредством выбора фамилий из телефонной книги, несомненно, будет иметься тенденция к выбору людей более состоятельных, и, следовательно, если в качестве варианты рассматривать богатство или какую-либо связанную с ним характеристику, такую как число детей, политические взгляды, уровень образования и т. д., то выборка тоже не будет случайной. Если бы мы интересовались такими характеристиками, как рост, цвет волос, группа крови, то выборка, возможно, была бы случайной, хотя во многих подобных случаях нетрудно представить причины, по которым варианта могла бы быть связана с уровнем жизни.

Если рассматривать вопрос с точки зрения субъективной теории вероятностей, то отсутствие знаний о соотношении между методом выбора и интересующей характеристикой может быть достаточным, чтобы обеспечить случайность, поскольку вероятности элементарных предложений являются тогда одинаковыми: вероятности определяются априорным субъективным отношением*). Если же принять частотную точку зрения, то подобное отсутствие знаний станет недостаточным, так как могут иметься неизвестные для наблюдателя соотношения, в силу которых элементарные предложения могут не быть приблизительно одинаково частыми. Тогда предполагается, что если осуществлены максимальные возможные усилия для нахождения существования каких-либо соотношений и если таковые при этом не обнаруживаются, то их вообще нет и, следовательно, метод можно считать случайным с большей или меньшей уверенностью. При этом подходе допущение случайности, в конечном счете, является частью общей неопределенности заключений о генеральной совокупности по выборке.

(б) Во-вторых, для оправдания использования какого-нибудь метода случайного выбора в новых обстоятельствах можно основываться на прошлом опыте его применения. Очевидно, это является экстраполяцией, и хотя большинство людей считает ее разумной, в этом нужно отдавать себе отчет. Субъективная теория вероятностей может включить эту экстраполяцию в свою сферу, поскольку определяемые этим методом вероятности можно оценить по первоначальному знанию; частотная же теория должна принимать экстраполяцию как дополнительное допущение.

*) По крайней мере, такова наша интерпретация этой ситуации. Специалисты по субъективной теории не обсуждали в печати проблему практического выбора сколько-нибудь подробно, и не исключено, что мы даем не вполне верное истолкование их взглядов, которое они, возможно, не примут.

9.10 Один из наиболее надежных методов получения случайных выборок состоит в конструировании модели генеральной совокупности и производстве выбора из этой модели. Можно, например, записать характеристики каждого элемента на карточку и производить выбор извлечением карточек из всего их набора, отвечающего всей генеральной совокупности. Этот метод применяется при лотереях, а сам процесс известен как лотерея, или выбор билетов. Он довольно эффективен, однако страдает на практике от двух недостатков: необходимости составлять генеральную совокупность карточек и опасности пристрастия при их выборе. Чтобы быть разумно удовлетворенным случайностью перемешивания, необходимо много усилий и труда, и та же цель может быть достигнута значительно более простым использованием случайных чисел, к рассмотрению которых мы теперь переходим.

Случайные числа

9.11 Наиболее легкий путь конструирования компактной генеральной совокупности состоит в сопоставлении каждому элементу целого положительного числа (проще всего это достигается нумерацией элементов числами 1, 2 и т. д.). Полученное множество целых чисел представляет собой компактную генеральную совокупность, и задача получения случайной выборки сводится к нахождению последовательности случайных чисел. Преимущества этого метода очевидны: не нужно конструировать никакую физическую модель генеральной совокупности; нумерация элементов может быть произведена любым удобным образом; одна и та же последовательность случайных чисел может применяться к любой перечислимой генеральной совокупности, так что любая последовательность случайных чисел имеет широкую область применения.

Одно обстоятельство следует здесь пояснить. Если нумерация элементов генеральной совокупности производится независимо от некоторых их характеристик, то любое множество чисел годится для получения выборки, случайной по отношению к этим характеристикам. Случайность в таком случае лежит, так сказать, в распределении целых чисел по элементам генеральной совокупности, а не в решении, какие числа взять для выборки. Практически, однако, подобного рода процедура значения не имеет, поскольку она заменяет лишь одну трудность другой — «случайной» нумерацией элементов генеральной совокупности. Обычно элементы генеральной совокупности нумеруют любым удобным образом, связанным с характеристиками эле-

ментов или нет, а затем выбирают случайное множество из множества всех чисел, соответствующих элементам генеральной совокупности.

9.12 Один из наиболее очевидных способов получения случайной выборки из пронумерованной генеральной совокупности состоит в использовании случайных чисел, взятых из какого-нибудь постороннего источника. Пусть, например, мы хотим произвести выборку из множества видимых на небе звезд. Будем игнорировать некоторые осложнения, возникающие из-за существования двойных звезд и неясных объектов. Так как положение звезды на небесной сфере определяется широтой и долготой, то нам нужна тогда последовательность случайных пар широты и долготы. На первый взгляд кажется разумным взять обычный атлас и выбрать числа, соответствующие расположенным в алфавитном порядке названиям мест в указателе, поскольку вряд ли можно ожидать какую-либо связь между распределением звезд на небе и распределением мест на земной поверхности. Небольшое размышление показывает, однако, что этот метод нехорош. На Земле существуют большие пространства суши и моря, на которых нет названных мест, — полюсы, пустыни, океаны; следовательно, в выборке может не оказаться чисел, соответствующих этим районам, и на небесной сфере будут исключаться области.

9.13 Можно было бы использовать иной метод: взять книгу, содержащую много цифр, например телефонный справочник или статистические или математические таблицы, открыть наудачу и выбрать те цифры, которые сразу бросаются в глаза или которые находятся сверху страницы, и т. п. Этот способ лучше предыдущего, но и он открыт некоторым возражениям.

(а) *Телефонные справочники.* Таблица 1.4 на стр. 22 показывает распределение 10 000 цифр, взятых из лондонского телефонного справочника. Цифры выбирались со страниц, на которых справочник открывался случайно; числа, состоящие меньше чем из четырех цифр, и числа, напечатанные жирным шрифтом, не использовались; из всех остальных четырехзначных чисел, напечатанных на странице, брались две последние цифры. Если бы числа были случайны, то следовало бы ожидать, что каждая цифра должна была бы встречаться около 1000 раз среди всех 10 000 цифр. В действительности же имеются довольно значительные отклонения от ожидаемого числа, и в одной из последующих глав мы увидим, что эти отклонения нельзя полностью отнести за счет выборочных флуктуаций. Заметно меньше 1000 имеется пятерок и девяток; это объясняется различными причинами: тенденцией избегать этих цифр из-за

сходного звучания *), резервирования чисел, оканчивающихся на 99, для испытательных целей телефонными служащими, и т. д. Очевидно, что с помощью подобных телефонных справочников нельзя построить таблицы случайных чисел.

(б) *Математические таблицы.* Очевидно, следует быть осторожным при использовании математических таблиц для построения случайных последовательностей. Возьмем, к примеру, таблицы логарифмов. Ясно, что имеется связь между последовательными логарифмами, объясняемая тем фактом, что если интервал невелик, то разности между табличными значениями приблизительно постоянны. Более того, справедлива любопытная теорема о цифрах в некоторых классах таблиц, дающая основание для теоретических сомнений относительно достоинств этого метода. Рассмотрим логарифмы по основанию 10 от натуральных чисел 1, 2 и т. д. Построим последовательность цифр, взяв k -ую цифру каждого логарифма. Оказывается, что относительная частота любой цифры в этой последовательности не стремится к пределу **) при возрастании длины последовательности, каково бы ни было k . Что именно происходит, по-видимому, неизвестно, однако представляется правдоподобным, что некоторые систематические эффекты начинают проявлять себя, и это, очевидно, ставит под сомнение случайность последовательности.

в) *Статистические таблицы.* Если имеется некоторое множество статистических данных таких, как численность населения городов и сельских районов, и если эти числа достаточно велики — скажем, четырех или более значные, — то есть основания предполагать, что последние их цифры будут случайными. И здесь, однако, использование таких таблиц требует осторожности: они могли быть составлены наблюдателем, пристрастным к некоторым цифрам; возможно, также имело место некоторое округление.

9.14 Необходимость каждому занимающемуся статистикой строить самому случайные последовательности была устранена публикацией различных таблиц случайных чисел. Имеется целый ряд их.

(а) Таблицы Типпетта, содержащие 41 600 цифр, взятых из отчетов по переписи населения и сгруппированных в 10 400 четырехзначных чисел (*Tracts for Computers*, No. 15).

(б) Таблицы Кендалла и Бэбингтона Смита, содержащие 100 000 цифр, сгруппированных по две и по четыре (*Tracts for*

*) По-английски 5 — five «фанв», 9 — nine «наин». (Прим. перев.)

**) Ср. И. Франел (J. Franel), *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* (1917) 62, 286.

Computers, No. 24). Эти числа были получены с помощью специально построенного прибора, кратко описанного в примере 9.3.

(в) Таблицы Фишера и Изйтса содержат 15 000 цифр, сгруппированных по две (*Statistical Tables for Biological, Agricultural and Medical Research*). Они были получены с помощью 15—19-значных таблиц логарифмов А. Томпсона и после того, как обнаружилось, что в них слишком много шестерок, улучшены.

(г) Рэнд Корпорейшен опубликовал миллион случайных цифр (*A Million Random Digits* (1955)), расположенных группами по пять. Они были получены с помощью специального рулеточного колеса; интересно отметить, что даже после нескольких проверок на наличие систематичности эти числа не являлись вполне случайными и их следовало еще улучшить.

Вышеупомянутые таблицы состоят из рандомизированных цифр. Волд (*Tracts for Computers*, No. 25) составил таблицу 25 000 нормальных отклонений, используя для этого таблицы Кендалла и Бэбингтона Смита и таблицы нормального интеграла. Филлер и другие (*Tracts for Computers*, No. 26) опубликовали 27 000 пар нормальных отклонений, коррелированных в заданных пределах. Книга, выпущенная Рэнд Корпорейшен, также содержит 100 000 нормальных отклонений *).

9.15 Прежде чем подробно рассматривать эти таблицы, будет, пожалуй, полезно привести несколько примеров их использования. Вот первые 100 цифр из таблиц Кендалла и Бэбингтона Смита.

Таблица 9.5

Случайные числа (*Tracts for Computers*, No. 24)

2 315	7 548	5 901	8 372	5 993	7 624	9 708	8 695	2 303	6 744
0 554	5 550	4 310	5 374	3 508	9 061	1 837	4 410	9 622	1 343
1 487	1 603	5 032	4 043	6 223	5 005	1 003	2 211	5 438	0 834
3 897	6 749	5 194	0 517	5 853	7 880	5 901	9 432	4 287	1 695
9 731	2 617	1 899	7 553	0 870	9 425	1 258	4 154	8 821	0 513

Пример 9.5

Произвести выборку объема 10 из генеральной совокупности 8585 мужчин, приведенной в таблице 1.7.

*) Кэнуй (Quepouille, 1959, *Biometrika* 46, 178) приводит 1000 случайных отклонений для каждого из девяти распределений: нормального, логнормального, экспоненциального, двойного экспоненциального, прямоугольного и трех разложений Эджворта. [В сборнике *Таблицы математической статистики* даны 12 500 случайных цифр и 2500 нормальных отклонений с тремя десятичными знаками. (Прим. ред.)]

Первое, что нужно сделать, — пронумеровать элементы генеральной совокупности. Здесь, как и во многих подобных случаях, одна нумерация уже произведена частотным распределением. Припишем мужчинам, принадлежащим группе 57— дюймов, номера 1 и 2, мужчинам, принадлежащим группе 58—, номера от 3 до 6 и т. д., те, которые оказались в группе 77—, будут иметь номера 8584 и 8585.

Возьмем теперь 10 четырехзначных чисел из таблиц; например, читая таблицу 9.5 по строке, имеем 2315, 7548, 5901, 8372, 5993, 7624, [9708], [8695], 2303, 6744, 0554, 5550.

Два числа, заключенных в квадратные скобки, больше 8585, и мы их поэтому не учитываем. Выберем теперь индивидуумы, соответствующие остальным десяти числам. Они оказываются принадлежащими интервалам 65—, 70—, 68—, 72—, 68—, 70—, 65—, 69—, 63—, 68— дюймов соответственно.

Среднее этих значений, если их считать равными серединам интервалов, которым они принадлежат, будет 68,24, в то время как среднее значение по всей генеральной совокупности равно 67,46.

Пример 9.6

Произвести выборку объема 12 из генеральной совокупности, о которой идет речь в нижеприводимой таблице с двумя входами, показывающей соотношение между инокуляцией и заболеванием холерой.

	Не заболевшие	Заболевшие	Всего
Инокулированные . . .	276 (0001—3312)	3 (3313—3348)	279
Не инокулированные .	473 (3349—9024)	66 (9025—9816)	539
Всего . . .	749	69	818

Всего в генеральной совокупности 818 элементов. Мы могли бы, конечно, взять трехзначные цифры из таблиц, например из таблицы 9.5:

231, 575, 485 и т. д.

Однако это не очень удобно, поскольку цифры не сгруппированы тройками. Удобнее взять, как и выше, четырехзначные

числа и каждому элементу генеральной совокупности сопоставить 12 чисел: например, первому поставить в соответствие числа от 0001 до 0012, второму — от 0013 до 0024, и т. д. Получающиеся при этом числа в приведенной выше таблице указаны в скобках. Числа, большие чем 9816, во внимание не принимаются.

Два опущенных в предыдущем примере числа можно теперь использовать, и все 12 выписанных там чисел приводят к следующим результатам:

	Не заболевшие	Заболевшие	Всего
Инокулированные . . .	3	0	3
Не инокулированные .	8	1	9
Всего . . .	11	1	12

Здесь, например, элемент, соответствующий числу 2315, принадлежит группе не заболевших инокулированных и т. д.

В этом примере так случилось, что ни один элемент из самой малой группы, состоящей из инокулированных и заболевших, не оказался выбранным. Предположим, что у нас получилась бы последовательность, содержащая числа

3314, 3323, 3333, 3341.

Все они соответствуют элементам, принадлежащим самой малой группе, которая содержит только три элемента. Если бы мы столкнулись с такой ситуацией, мы должны были бы решить, нужно ли производить выбор с возвращением или без него. В случае выбора без возвращения следовало бы предположить, что первые три числа из последовательности 3313—3348 исчерпали бы соответствующую группу элементов генеральной совокупности, и не принимать во внимание любые числа, соответствующие этой группе, встречающиеся впоследствии.

Отметим, что использование случайных чисел предполагает обычно выбор с возвращением. Если же мы хотим производить выбор без возвращения, то нам не следует использовать одно и то же число дважды. Это предполагает регистрирование тех чисел, которые уже использовались, — скучную и утомительную процедуру, исключая случаи, когда выборка или генеральная совокупность мала.

Пример 9.7

Построить последовательность случайных перестановок всех цифр от 1 до 5.

Поскольку цифры 0, 6, 7, 8 и 9 в данной задаче нас не интересуют, мы не будем принимать их во внимание в таблице случайных чисел. Читаем таблицу и отмечаем цифры в том порядке, в каком они записаны, например, в таблице 9.5 имеем 2315 7548 и т. д. Цифру 7 игнорируем так же, как и вторую пятерку, поскольку одна пятерка уже появилась. Получаем перестановку 23154. Далее продолжаем аналогично. Поскольку следующие цифры суть 8 5901 8372 5993 7624 ..., то второй перестановкой будет 51324 и т. д. *).

Пример 9.8

Произвести случайную выборку объема 10 из нормальной генеральной совокупности

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx.$$

Этот случай особенно интересен, поскольку здесь нам нужно сделать выборку из бесконечной генеральной совокупности. Как мы уже видели, такой процесс может рассматриваться только в предельном смысле.

Предположим, что числовая прямая, которую здесь можно идентифицировать с генеральной совокупностью, разбита на интервалы длины 0,1. Из таблицы значений нормального интеграла (см., например, Приложение, таблица 2), находим значения $F(x)$, приведенные на стр. 303.

Чтобы выбрать значение случайной переменной x , распределенной в соответствии с $F(x)$, теперь можно использовать четырехзначные случайные числа: например, числу 5461 соответствует +0,1—, а числу 3500 соответствует —0,4—.

Если бы мы взяли таблицу, в которой числа приводятся с точностью до n -го знака после запятой, то нам потребовались бы n -значные случайные числа. Конечно, можно получить более точную аппроксимацию, беря меньшие интервалы. Такие вопросы следует решать с точки зрения степени нужной аппроксимации.

*) Таблицы 400 случайных перестановок чисел от 1 до 20 приведены в книге Кендалла *Rank Correlation Methods*, изд. 3-е (Griffin, 1962). Перестановки меньшего, чем 20, количества чисел могут быть получены отбрасыванием ненужных чисел. С помощью очевидных приемов из них можно получить также 800 перестановок всех чисел от 1 до 10, 1600 перестановок всех чисел от 1 до 5 и т. д. Кокрэн и Кокс в книге *Experimental Designs* (Wiley, 1950) приводят по 1000 перестановок 9 и 16 целых чисел.

x	$F(x)$	x	$F(x)$
$-\infty$	0,00000	0,0	0,5000
.	.	.	.
.	.	.	.
-3,7	0,00011	0,1	0,5398
-3,6	0,00016	0,2	0,5793
-3,5	0,00023	0,3	0,6179
.	.	.	.
.	.	.	.
-0,4	0,3085	3,5	0,99977
-0,3	0,3821	3,6	0,99984
-0,2	0,4207	3,7	0,99989
.	.	.	.
.	.	.	.
-0,1	0,4602	∞	1,00000

9.16 Чтобы случайные числа могли использоваться, они должны удовлетворять некоторым условиям. Любое множество чисел является случайным в том смысле, что оно могло появиться в результате случайного выбора, однако не всякое такое множество подходит в качестве таблицы случайных чисел. Из вышеприведенных примеров ясно, что желательно иметь таблицу, обладающую большой гибкостью. Она должна давать случайные результаты в возможно большем числе случаев вне зависимости от того, используется она частично или целиком.

Однако невозможно построить таблицу случайных чисел, которая полностью удовлетворяла бы этому требованию. Предположим, беря крайний случай, что построена таблица из $10^{10^{10}}$ цифр. Вероятность любой цифре оказаться нулем равна $\frac{1}{10}$, и поэтому вероятность того, что любое заданное множество из миллиона цифр состоит из нулей, есть 10^{-10^6} . Таких групп, следовательно, было бы довольно много в совокупности $10^{10^{10}-6}$ групп, по миллиону цифр в каждой. В противном случае все множество не было бы удовлетворительным для некоторых выборочных экспериментов. Ясно, однако, что множество из миллиона нулей не подходит для получения выборок в эксперименте, требующем меньше, чем миллион цифр.

Таким образом, следует ожидать, что в таблице случайных чисел будут появляться куски, не подходящие для использования сами по себе. Редкое должно иметь возможность появляться

в соответствующей малой пропорции. Кендалл и Бэбингтон Смит пытались принимать во внимание это обстоятельство, указывая части их таблицы (5 тысяч из 100 тысяч), которые лучше не использовать в выборочных экспериментах, требующих меньше 1000 цифр.

9.17 Если таблица случайных чисел используется для извлечения элементов из генеральной совокупности объема 10, то мы ожидаем появления элементов в приблизительно одинаковых пропорциях. Иными словами, мы рассчитываем, что такая таблица содержит десять цифр 0—9 в приблизительно одинаковых пропорциях. Точно так же мы рассчитываем на появление в приблизительно одинаковых пропорциях ста пар 00—99 и т. д. Можно придумать различные тесты этого типа, основываясь на сравнении действительных частот с частотами, которые должны быть согласно законам теории вероятностей. Ни одна таблица не может удовлетворять им всем, однако если она удовлетворяет тестам: (а) обеспечивающим случайность чисел для тех общих типов выборочных исследований, для которых она предназначена, и (б) способным обнаружить любой вид пристрастия, которому числа подвержены в силу способа их построения, — то она может иметь общее применение.

9.18 Развитие быстродействующих вычислительных машин сделало возможным ввести в некоторые из них «рандомизирующее устройство», которое производит случайные выборки. Использование таких механизмов требует исследования; представляется трудным произвести множества чисел, удовлетворяющие обычным тестам на случайность.

Еще один из применявшихся методов состоит в возведении чисел в очень высокие степени и в выборе цифр из середины так полученных очень больших чисел. Созданные этим путем «псевдослучайные числа» также могут не выдерживать тестов на случайность. С другой стороны, при некоторых обстоятельствах они могут быть полезны в выборочных экспериментах по одной из следующих двух различных причин: они могут быть «приблизительно» случайными; если их взять в достаточном количестве, то ими можно пронумеровать все возможные выборки, хотя и в нерегулярном порядке. Широкое использование этого метода небезопасно без значительного опыта.

Выбор из непрерывной генеральной совокупности

9.19 Использование случайных чисел является наилучшим известным в настоящее время методом для получения случайных выборок из перечислимой генеральной совокупности. Как видно из примера 9.8, этот метод может быть также использован

для получения случайных выборок из непрерывной генеральной совокупности с распределением известной математической формы. Однако иногда встречаются случаи, когда случайные числа использовать нельзя. Например, если мы хотим взять выборку молока или муки, то мы практически не можем перенумеровать каждую частицу и выбрать ее из генеральной совокупности для изучения. В таких случаях мы обычно вынуждены использовать более интуитивные методы. Например, чтобы взять случайную выборку из маслостойки, можно тщательно взболтать ее содержимое и затем зачерпнуть пробу наудачу. В некоторых случаях, когда генеральная совокупность не очень велика, можно действовать по единой схеме: разбить генеральную совокупность на несколько групп и выбирать эти группы с помощью обычной техники, используя случайные числа. Многие науки имеют свои собственные, специфические задачи, связанные с выбором, и нет возможности обсуждать здесь их все. На этом заканчиваем рассмотрение техники случайного выбора и в дальнейшем предполагаем, если не оговорено противное, что обсуждаемый материал получен с помощью некоторого случайного процесса. К сравнению эффективности различных методов выбора мы вернемся в третьем томе.

Выборочное изучение качественных признаков

9.20 В качестве введения в общие проблемы выбора мы рассмотрим выбор с целью изучения качественных признаков, который содержит в себе многие принципиальные трудности и в то же время не затемняется математической техникой.

Предположим, что имеется случайная выборка из генеральной совокупности, все элементы которой обладают либо качеством A , либо противоположным ему качеством — не- A . Пусть объем выборки равен n и доля элементов, обладающих качеством A , равна p , т. е. имеется pn таких элементов; доля и количество элементов, обладающих качеством не- A , равны тогда соответственно q и qn ($p+q=1$). Мы будем предполагать, что генеральная совокупность велика, или что выбор производится с возвращением. Тогда вероятность выбрать элемент с качеством A в каждом отдельном случае не зависит от других выборов и, следовательно, постоянна, скажем равна ω .

Нам нужно рассмотреть три типа задач:

(а) Пусть имеется основание полагать долю элементов, обладающих качеством A , известной и равной ω . Подтверждает наблюдаемая доля p эту гипотезу или отклонение от ω столь велико, что вызывает сомнение в ее правильности? Пусть, например, скрещивая растения, обладающие двумя различными

качественными признаками, скажем высокий и низкий горох, мы хотим проверить простой закон Менделя о наследовании доминантных и рецессивных признаков. Если скрестить между собой два чистых вида высокого и низкого гороха, а затем скрестить между собой растения этого первого поколения, то согласно простому закону Менделя во втором поколении относительные частоты «низких» и «высоких» будут $3/4$ и $1/4$ при доминировании признака «низкий» и $1/4$ и $3/4$ при доминировании признака «высокий». Предположим, что мы производим такой эксперимент и для 400 растений частоты оказываются равными 70 и 330. Могло ли это отклонение от теоретических значений 100 и 300 возникнуть случайно или оно достаточно велико, чтобы вызвать сомнение в справедливости простого закона Менделя?

(б) В задаче предыдущего типа имеется определенное основание для проверки значения ω , данного *a priori*. Но мы можем ничего не знать об ω , и в этом случае основной задачей будет оценка ω по выборке.

(в) Оценив ω , мы хотим узнать степень надежности этой оценки. Насколько сильно оценка может отклоняться от истинного значения ω ?

9.21 Рассмотрим первый тип задачи, когда ω задано *a priori*. Если производятся повторные выборки объема n из рассматриваемой генеральной совокупности, то распределение числа элементов с качеством A в выборке будет задаваться членами разложения биномиального выражения $(\chi + \omega)^n$, где $\chi = 1 - \omega$. (Это есть не что иное, как биномиальное распределение, рассмотренное в 5.2; мы используем здесь греческие буквы вместо p и q для выделения того факта, что они суть параметры генеральной совокупности). Вероятность P иметь в выборке np или меньше элементов с качеством A равна сумме первых $np + 1$ членов этого распределения. Гипотеза проверяется следующим образом: P сравнивается с заранее заданной (обычно малой) вероятностью ошибки α , которую мы готовы допустить при отвержении гипотезы; если P больше α , то гипотеза не отвергается, если α больше P , то гипотеза отвергается.

Внешне эта процедура представляется разумной: гипотеза отвергается, если наблюдаемое значение статистики оказывается в наиболее «неправдоподобной» области в смысле выборочного распределения. Удовлетворительный анализ логики этого теста, однако, будет сделан лишь в результате более детального обсуждения, содержащегося во втором томе.

Следует отметить, что определенное выше P относится только к одному из «хвостов» выборочного распределения. В практике для тестов более часто используются P , учитывающие оба хвоста распределения. Иначе говоря, часто гипотеза отвергается,

если в результате теста получается значение либо меньшее, либо большее по сравнению с ожидаемым числом элементов со свойством A в выборке, в то время как в других случаях нам бывают интересны отклонения только в одну сторону. Детальное обсуждение рациональности этих тестов опять-таки придется отложить до второго тома. Отметим лишь, что в рассмотренном выше примере с законом Менделя следует использовать интуитивно ясный подход, считая только наличие малого количества «высоких» растений противоречащим гипотезе доминантности признака «высокий».

Пример 9.9

При некотором проведении опытов с подбрасыванием монеты в 20 бросаниях 15 раз выпал «герб». Противоречит ли это гипотезе о симметричности монеты?

Здесь нам нужно проверить гипотезу $\omega = \frac{1}{2}$. Несимметричность монеты может вызывать преимущественное выпадение как герба, так и решетки, и поэтому мы вычислим P , учитывая оба конца выборочного распределения. Вероятность того, что при 20 бросаниях выпадет 15 или более гербов, есть сумма первых шести членов разложения $\left(\frac{1}{2} + \frac{1}{2}\right)^{20}$; согласно таблице 5.2 она равна 0,0207. В силу симметрии это же число дает вероятность выпадения 15 или более решеток. Следовательно, $P=0,0414$. Если бы допускался риск ложного отвержения гипотезы, меньший чем 0,0414, то гипотеза не была бы отвергнута. В противном случае гипотеза отвергалась бы. Обычно допускают риск, равный 0,05; в этом случае нашу гипотезу следовало бы отвергнуть.

9.22 В только что приведенном примере мы намеренно взяли небольшое значение n так, чтобы элементы разложения биномиального выражения могли быть непосредственно вычислены. В практике n часто довольно велико — 100 или более — и вычисление и суммирование отдельных членов было бы громоздким и утомительным. Если требуется полная точность, то можно использовать метод суммирования, предложенный в 5.7, основывающийся на применении неполной В-функции, или воспользоваться указанными там таблицами. Однако для обычных целей вполне достаточно использовать аппроксимацию биномиального распределения нормальным. В примере 4.6 мы видели, что при $n \rightarrow \infty$ биномиальное распределение стремится к нормальному, имеющему, в наших теперешних обозначениях, среднее $n\omega$ и дисперсию $n\omega\chi$. Следовательно, нужные вероятности можно вычислять с помощью нормального интеграла. На самом деле для

многих целей вообще нет необходимости проводить действительные вычисления. Согласно таблицам нормального интеграла (приложение, таблица 2) вероятность того, что абсолютное отклонение не менее стандартного отклонения, равна 0,3173; точно так же удвоенному стандартному отклонению соответствует вероятность 0,0455, утроенному 0,0027, учетверенному *) 0,00006. Следовательно, если в качестве области отклонения гипотезы принять дополнение к интервалу $n\omega \pm 2\sqrt{n\omega\chi}$, то вероятность ложного отклонения будет равна 0,0455. Аналогично определяются вероятности для других интервалов с концами, кратными $\sqrt{n\omega\chi}$.

Нет никакой разницы, сравниваем ли мы действительные или относительные частоты, поскольку при делении варианты на константу среднее и стандартное отклонение меняются одинаково.

Пример 9.10

В одном из экспериментов с игральными костями Уэлдон подбросил кости 49 152 раза. В 25 145 случаях выпадали 4, 5 или 6. Согласуется ли это с гипотезой о симметричности костей?

Если кости симметричны, то вероятность выпадения 4, 5 или 6 равна $1/2$. Следовательно, $n\omega = 24\,576$ и наблюдаемое np отличается от $n\omega$ на 569. Далее, стандартное отклонение равно

$$\sqrt{n\omega\chi} = \sqrt{49\,152 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 110,9.$$

Наблюдаемое отклонение здесь в пять раз превосходит стандартное. Соответственно у нас имеется очень сильное подозрение в несимметричности костей.

Стандартная ошибка

9.23 В биномиальном случае величина $\sqrt{n\omega\chi}$ является стандартной ошибкой статистики np . Стандартная ошибка — важное статистическое понятие. В общем случае стандартная ошибка определяется как стандартное отклонение статистики. Оно особенно важно в том, относительно широком, классе случаев, когда исходное распределение может считаться нормальным либо точно, либо с соответствующей степенью приближения.

9.24 Обратимся теперь к случаю (б), в котором значение ω а priori не задано. Если относительная частота элементов со

*) Приводимые здесь значения вероятностей более точны, чем те, которые можно получить из таблицы 2 приложения. (Прим. перев.)

свойством A в выборке равна p , то какую величину взять в качестве оценки для ω ? Наиболее очевидное решение — взять само p ; это же решение диктуется более тонкими соображениями, описанными в главе 8.

Прежде всего рассмотрим метод максимального правдоподобия. Вероятность получить np элементов со свойством A и nq — со свойством не- A равна

$$\binom{n}{np} \omega^{np} \chi^{nq}. \quad (9.1)$$

Эта величина, рассматриваемая как функция от ω , есть функция правдоподобия. Нам нужно, следовательно, максимизировать по ω функцию

$$L = k \omega^{np} (1 - \omega)^{nq}.$$

Так как L неотрицательна, то $\frac{\partial L}{\partial \omega}$ и $\frac{\partial}{\partial \omega} (\log L) = \frac{1}{L} \frac{\partial L}{\partial \omega}$ имеют одни и те же нули и, стало быть, достаточно максимизировать $\log L$. Имеем

$$\frac{\partial}{\partial \omega} (\log L) = \frac{np}{\omega} - \frac{nq}{1-\omega} = 0.$$

Отсюда непосредственно получаем значение оценки $\hat{\omega}$:

$$\hat{\omega} = p. \quad (9.2)$$

Метод Байеса дает тот же результат, если предположить непрерывную равномерную распределенность ω между 0 и 1. В самом деле, имеем

$$P(\omega | p) \propto P(\omega | H) P(p | \omega H), \quad (9.3)$$

$$\propto \omega^{np} \chi^{nq} d\omega. \quad (9.4)$$

Максимум и здесь достигается при $\omega = p$.

Имеется иной аспект задачи отыскания оценки. Предположим, что из генеральной совокупности, в которой элементы со свойством A составляют ω -ю часть, а со свойством не- A составляют χ -ю, взято большое число выборок объема n . Нашей оценкой для ω в каждом случае является p , причем p меняется от выборки к выборке. Среднее значение этой оценки равно

$$\begin{aligned} M(p) &= \sum_{p=0}^1 p \binom{n}{np} \omega^{np} \chi^{nq} = \frac{1}{n} \sum \left\{ np \binom{n}{np} \omega^{np} \chi^{nq} \right\} = \\ &= \frac{1}{n} \omega n \sum \left\{ \binom{n-1}{np-1} \omega^{np-1} \chi^{nq} \right\} = \omega (\omega + \chi)^{n-1} = \omega. \end{aligned} \quad (9.5)$$

Подобного рода оценки, средние значения которых совпадают с оцениваемыми параметрами, называют *несмещенными*. Несмещенные оценки в известном смысле можно считать надежными.

9.25 Все подходы в этом случае, следовательно, ведут к одному и тому же заключению (счастливое обстоятельство, которое, как будет видно в дальнейшем, не всегда имеет место). Рассмотрим теперь следующую нашу задачу (в): какова надежность оценки? Иными словами, насколько сильно оценка может отличаться от истинного значения?

Мы знаем, что вероятность выборочному значению отличаться от ω на $t \sqrt{(\omega\chi)/n}$ уменьшается при возрастании t . Следовательно, с заданной степенью надежности можно сказать, что абсолютной величине разности между p и ω невероятно оказаться больше некоторой заданной величины. Но чтобы указать эту заданную величину, нужно знать ω , а ω мы как раз и ищем.

Так поставленную задачу можно решить только приближенно. При больших n стандартная ошибка величины p имеет порядок $n^{-\frac{1}{2}}$. Положим

$$\omega = p + \frac{k}{n^{\frac{1}{2}}}.$$

Имеем, пренебрегая величинами порядка n^{-1} ,

$$\begin{aligned} \sqrt{\frac{\omega\chi}{n}} &= \sqrt{\frac{1}{n} \left(p + \frac{k}{n^{\frac{1}{2}}} \right) \left(q - \frac{k}{n^{\frac{1}{2}}} \right)} = \sqrt{\frac{pq}{n} \left(1 + \frac{k(q-p)}{pq n^{\frac{1}{2}}} \right)} = \\ &= \sqrt{\frac{pq}{n}} \left\{ 1 + \frac{1}{2} \frac{k(q-p)}{pq n^{\frac{1}{2}}} \right\}. \quad (9.6) \end{aligned}$$

Следовательно, при больших n стандартная ошибка величины p равна приблизительно $\sqrt{\frac{pq}{n}}$. Мы получили весьма важный результат — при больших выборках из генеральных совокупностей рассматриваемого типа стандартную ошибку можно вычислять, используя вместо (неизвестных) значений оцениваемых параметров их оценки.

Во втором томе будет рассмотрен иной подход, свободный от необходимости подобного рода аппроксимации.

Пример 9.11

Произведена выборка объема 600, и в ней оказалось 240 элементов, обладающих качеством A . Имеем: $p=0,40$, $np=240$, $\sqrt{npq}=12$. Можно считать, следовательно, маловероятным отклонение $n\omega$ от 240 больше чем на 24 и очень маловероятным

отклонение, большее 36. С определенной степенью уверенности можно утверждать, стало быть, что величина $\rho\omega$ лежит в интервале 240 ± 24 , и с большой степенью уверенности можно утверждать ее принадлежность интервалу 240 ± 36 .

9.26 Обратимся теперь к общему рассмотрению проблем выбора, проиллюстрированных выше на примерах. Отметим прежде всего ту роль, которую играет здесь распределение выборочной характеристики. По наблюдениям мы строим некоторую статистику t . Распределение этой статистики обычно (но не всегда) зависит от некоторых параметров исходного распределения. По вероятности наблюденного значения t с помощью одного из методов — байесовского, максимума правдоподобия и т. д. — можно делать некоторые заключения об этих параметрах; иными словами, мы получаем информацию об исходном распределении. Распределения выборочных статистик, таким образом, очень важны для всей теории. Несколько последующих глав целиком будут посвящены методам нахождения выборочных распределений, отвечающих заданным генеральным совокупностям.

Проверка гипотезы относительно генеральной совокупности, выражаемой в терминах некоторых заданных а priori параметров, довольно проста. При фиксированных значениях параметров по соответствующему выборочному распределению мы можем определить вероятность наблюденного значения статистики и затем использовать эту вероятность для оценки приемлемости гипотезы. Но и здесь, однако, могут возникать осложнения, поскольку в общем случае несколько статистик можно построить по одной и той же выборке, и не все они необходимо ведут к одинаковому заключению относительно гипотезы; например, выборка может иметь среднее, дающее основание сомневаться в гипотезе, и дисперсию, которая такого основания не дает. Более полно эта трудность будет обсуждаться во втором томе.

9.27 В том случае, когда параметры генеральной совокупности не заданы а priori, мы имеем двойную задачу — оценить параметры по выборке и установить вероятные границы для полученных оценок. Мы уже касались некоторых принципов отыскания оценок. Этот вопрос будет рассмотрен более систематически в соответствующем месте. Когда оценка получена (она сама является статистикой), ищется ее распределение, с помощью которого можно указать вероятные границы истинного значения параметра. В специальный класс выделяются случаи, когда можно найти статистику, распределение которой зависит только от одного параметра генеральной совокупности (как в случае генеральной совокупности, каждый из элементов которой обладает одним из двух противоположных качеств).

9.28 Если выборка велика, то при рассмотрении этих последних типов задач можно пользоваться некоторой важной аппроксимацией. В главе 7 мы видели, что при весьма широких условиях распределения сумм n независимых произвольно распределенных случайных величин приближаются к нормальному с ростом n . Но многие из обычно используемых статистик (например, все моменты) можно представить в виде суммы случайных величин.

Для относительно большого числа других статистик также можно показать, что они распределены приблизительно нормально при больших выборках.

Таким образом, при изучении рассматриваемого круга вопросов используются следующие аппроксимации:

(а) статистика, вычисленная по выборке, берется в качестве оценки соответствующего параметра генеральной совокупности; например, дисперсию выборки можно взять в качестве оценки дисперсии генеральной совокупности;

(б) среднее и дисперсия распределения статистик вычисляются с использованием вместо неизвестных значений параметров самих значений статистик, определенных в соответствии с (а);

(в) распределение статистики предполагается нормальным и нужные вероятности находятся с помощью нормального интеграла с использованием выборочных среднего и дисперсии (последняя равна квадрату стандартной ошибки).

9.29 Ответить на вопрос, как велико должно быть n , чтобы можно было пользоваться этими аппроксимациями, вовсе не всегда просто. Для некоторых распределений, в частности для распределения среднего, вполне удовлетворительная аппроксимация получается при небольших значениях n , скажем $n > 30$. В случае других распределений хорошая аппроксимация получается лишь при значительно больших n ; например, для распределения коэффициента корреляции при выборках из нормальной генеральной совокупности даже столь большие значения n , как 500, недостаточно еще хороши. Скорость приближения распределения статистики к нормальному в действительности зависит как от вида исходного распределения, так и от рассматриваемой статистики.

Пример 9.12

Возвращаясь к биномиальному распределению, посмотрим, какие появляются отличия в случае выбора без возвращения.

Если из генеральной совокупности размера N , содержащей N_ω элементов с качеством A , извлекается случайная выборка

без возвращения объема n , то распределение доли p элементов в выборке, обладающих качеством A , является гипергеометрическим распределением (5.18), общий член которого в наших те- перешних обозначениях имеет вид

$$\binom{n}{np} \frac{(N\omega)^{[np]} (N\chi)^{[nq]}}{N^{[n]}}.$$

Среднее значение p для этого распределения согласно (5.53) равно ω , так что доля элементов в выборке с качеством A по- прежнему остается несмещенной оценкой доли таких элементов в генеральной совокупности. В рассматриваемом случае p не является уже ни байесовской оценкой, ни оценкой максималь- ного правдоподобия (см. упражнение 9.10), однако обе эти оценки (имеющие здесь довольно сложный вид) отличаются от p только на величину порядка N^{-1} .

Дисперсия p , равная согласно (5.55)

$$Dp = \frac{N-n}{N-1} \frac{\omega\chi}{n}, \quad (9.7)$$

отличается от дисперсии p , соответствующей обычному бино- мальному случаю (выбор с возвращением), множителем $(N-n)/(N-1)$, который меньше единицы при $n > 1$. Диспер- сия p при выборе без возвращения меньше, следовательно, дис- персии p при выборе с возвращением; это и интуитивно ясно, поскольку экстремальных возможностей больше при выборе с возвращением.

Пример 9.13. Последовательный выбор

Иногда встречаются случаи, особенно при выборочном изу- чении качественных признаков, когда объем выборки не фикси- руется заранее. Точнее, вместо извлечения выборки размера n и наблюдения числа $r (=np)$ «успехов», иначе — появлений каче- ства, выбор продолжается до тех пор, пока не появится заре- нее заданное число успехов r . В этом случае r фиксировано и n является случайной величиной.

Легко найти распределение n . Вероятность того, что среди первых $n-1$ извлечений имеется $r-1$ успехов, равна $\binom{n-1}{r-1} \omega^{r-1} \chi^{n-r}$. Вероятность того, что следующее извлечение доставляет успех, равна ω . Следовательно, распределение n имеет вид (поскольку выбор не может прекратиться, пока не извлечено по меньшей мере r элементов)

$$P(n=s) = \binom{s-1}{r-1} \omega^r \chi^{s-r}, \quad s = r, r+1, \dots, \infty \quad (9.8)$$

Последовательные его члены, равные

$$\omega^r \left\{ 1, r\chi, \frac{1}{2} r(r+1)\chi^2, \dots \right\},$$

совпадают с последовательными членами отрицательного биномиального распределения $[\omega/(1-\chi)]^r$. Это — иная форма результата, полученного в 5.14. Согласно (5.38) среднее и дисперсия этого распределения равны соответственно $r\chi/\omega$ и $r\chi/\omega^2$.

Метод максимального правдоподобия в качестве оценки ω дает здесь, как и в случае обычного биномиального распределения, значение p . Однако теперь эта оценка не является несмещенной. В самом деле, покажем, что $(r-1)/(n-1)$ есть несмещенная оценка для ω . Имеем

$$\begin{aligned} M\left(\frac{r-1}{n-1}\right) &= \omega^r \sum_{n=r}^{\infty} \frac{r-1}{n-1} \chi^{n-r} \left(\frac{n-1}{r-1}\right) = \\ &= \omega^r \sum \binom{n-2}{r-2} \chi^{n-r} = \omega^r (1-\chi)^{-(r-1)} = \omega. \end{aligned}$$

Отсюда следует смещенность оценки r/n , поскольку, будучи всегда больше (за исключением лишь случая $r=n$), чем $(r-1)/(n-1)$, она не может иметь одинаковое с $(r-1)/(n-1)$ среднее.

Любопытно, что тем не менее n/r является несмещенной оценкой $1/\omega$, так как $M(n) = r/\omega$.

На первый взгляд это представляется довольно парадоксальным. Ведь n/r есть несмещенная оценка для $1/\omega$, но r/n — смещенная оценка для ω . Причина этого, конечно, в довольно произвольном определении смещенности как отклонения от среднего значения. Не следует удивляться тому, что одновременно $M(x) = \alpha$ и $M(1/x)$ не равно $1/\alpha$. Для положительных случайных величин это, как правило, так (см. упражнение 9.13).

Метод выбора, которым мы пользовались, производя наблюдения до тех пор, пока не появится определенное число «успехов», иногда называется «обратным выбором» (против этого термина имеется ряд возражений). Он представляет собой частный случай изучающегося во втором томе метода, известного под названием *последовательный выбор* (для изучения качественных признаков). Характерной чертой этого более общего метода является то, что в каждый момент вопрос о продолжении испытаний решается на основе данных, полученных до этого момента.

УПРАЖНЕНИЯ

9.1 Из 10 000 детей, родившихся в некоторой стране, 5100 оказались мальчиками. Рассматривая это как случайную выборку рождений в этой стране, показать, что она вызывает значительные сомнения в равенстве частот рождения детей обоих полов.

Выяснить, насколько сильно изменился бы вывод, если бы выборка состояла из 1000 рождений, 510 из которых дали мальчиков.

9.2 В выборке объема n_1 из генеральной совокупности π_1 доля элементов, обладающих качеством A , оказалась равной p_1 . В другой, не зависящей от первой, выборке объема n_2 из генеральной совокупности π_2 соответствующая доля равна p_2 . Чтобы проверить, одинаковы ли доли элементов с качеством A в обеих генеральных совокупностях, предлагается рассматривать разность $p_1 - p_2$. Показать, что в предположении равенства исходных долей указанная разность имеет среднее нуль, а в качестве оценки для ее дисперсии можно взять

$$(p_1 n_1 + p_2 n_2) (q_1 n_1 + q_2 n_2) / \{n_1 n_2 (n_1 + n_2)\},$$

где $q_1 = 1 - p_1$, $q_2 = 1 - p_2$.

9.3. Пусть нужно оценить ω по простой случайной выборке, в которой доля успехов p . Если f — априорная вероятность ω , то апостериорная вероятность ω согласно теореме Байеса пропорциональна величине

$$f \omega^{np} (1 - \omega)^{nq}.$$

Показать, что величина эта достигает максимума, когда

$$\frac{1}{f} \frac{\partial f}{\partial \omega} + n \frac{p - \omega}{\omega (1 - \omega)} = 0.$$

Вывести отсюда, что вообще, какова бы ни была априорная вероятность ω , решение этого уравнения стремится к $\omega = p$ при возрастании n .

9.4 Используя неравенство Бьенэме — Чебышева, показать, что при выборе для изучения качественных признаков из большой генеральной совокупности вероятность того, что наблюдаемая доля p в выборке объема n отличается от доли ω в генеральной совокупности более чем на k , не превосходит $1/(4nk^2)$.

(Это — точный результат. При его выводе не делается допущений о предельной нормальности рассматриваемого биномиального распределения и не используются оценки в вычислении стандартных ошибок. Однако указываемые здесь границы обычно слишком широки.)

9.5 В большой генеральной совокупности доля ω элементов, обладающих качеством A , мала. Показать, что метод максимального правдоподобия в качестве оценки для ω по выборке объема n даст долю элементов с качеством A в этой выборке. Установить, что оценка эта несмещенная и что для дисперсии числа наблюдаемых элементов с качеством A , скажем n_1 , оценкой может служить n_1 .

9.6 В k различных генеральных совокупностях доли элементов, обладающих различными качественными признаками, равны $\omega_1, \omega_2, \dots, \omega_k$. С вероятностью $\frac{1}{k}$ случайно выбирается одна из генеральных совокупностей и из нее делается случайная выборка с возвращением объема n , содержащая p -ю долю элементов, обладающих качественным признаком. Показать, что при повторении этого процесса (включая повторный выбор генеральной совокупности)

среднее значение p будет равно $\omega = \sum_{i=1}^k \omega_i/k$, а дисперсия равна

$$\omega(1-\omega)/n + (n-1) \sum (\omega_i - \omega)^2/nk.$$

Получить соответствующие формулы, когда вероятность выбора i -й генеральной совокупности равна α_i , $\sum_{i=1}^k \alpha_i = 1$.

9.7 В условиях предыдущего упражнения производится выборка объема kn , полученная посредством извлечения n элементов из каждой генеральной совокупности. Показать, что доля p элементов, обладающих качественным признаком, среди взятых вместе kn выбранных элементов имеет среднее ω и дисперсию

$$\omega(1-\omega)/(nk) - \sum (\omega_i - \omega)^2/(nk^2).$$

9.8 Из генеральной совокупности, в которой ω — доля элементов, обладающих некоторым качественным признаком, производится выборка с возвращением объема n . Из нее производится подвыборка с возвращением объема n_1 . Пусть p и p_1 — доли элементов с качественным признаком в выборке и подвыборке соответственно. Показать, что разность $p - p_1$ имеет среднее нуль и дисперсию

$$\omega(1-\omega) \frac{n-1}{nn_1}.$$

9.9 Показать, что если в предыдущем упражнении выбор с возвращением заменить на выбор без возвращения и объем генеральной совокупности положить равным N , то $p - p_1$ будет иметь среднее нуль и дисперсию

$$\omega(1-\omega) \frac{N}{N-1} \cdot \frac{n-n_1}{nn_1}.$$

9.10 Показать, что с точностью до величин порядка N^{-1} оценка максимального правдоподобия доли ω при выборе без возвращения имеет вид

$$p + \frac{1}{2N} (p - q)$$

и что с точностью до величин того же порядка дисперсия этой оценки равна

$$\frac{N-n}{N-3} \frac{\omega\chi(1-\omega)}{n}.$$

9.11 Вывести соотношение (9.7), используя случайные величины, равные единице, когда интересующий нас элемент обладает качеством A , и нулю в противном случае.

9.12 Показать, что в условиях примера 9.13

$$M(r/n) = r\omega^r \chi^{-r} \int_0^1 t^{r-1} (1-t)^{-r} dt = \omega \sum_{j=0}^{\infty} \frac{r! j!}{(r+j)!} \chi^j.$$

9.13 Пусть x — случайная величина, определенная в интервале $0 \leq x \leq \infty$. Показать, что если существуют $M(x)$ и $M(1/x)$, то

$$M(x) M(1/x) \geq 1,$$

причем равенство здесь достигается лишь в том случае, когда распределение x целиком сосредоточено в некоторой точке. Вывести отсюда, что если t — несмещенная оценка для θ , то $1/t$ не может быть несмещенной оценкой для $1/\theta$.

9.14 Доказать, что в случае обычного биномиального распределения $M(n/r)$ не существует. Показать, что

$$M\left(\frac{n+1}{r+1}\right) = (1 - \chi^{n+1})/\omega$$

и что, следовательно, $(n+1)/(r+1)$ является асимптотически по n несмещенной оценкой для $1/\omega$.

СТАНДАРТНЫЕ ОШИБКИ

10.1 В конце предыдущей главы мы обсудили некоторые вопросы, связанные с оценками статистических параметров, получаемыми по выборкам большого объема, и рассмотрели степень надежности этих оценок с точки зрения теории стандартных ошибок. Там было отмечено, что оценку нужного параметра генеральной совокупности можно получить по выборке большого объема, найдя соответствующее значение параметра для этой выборки, рассматриваемой как генеральная подсовокупность. Там же было установлено, что для выборок объема n стандартная ошибка служит вполне удовлетворительной мерой точности при условии, что: (а) имеет место асимптотическая нормальность выборочного распределения статистик и (б) n достаточно велико в том смысле, как это было там определено. Было также отмечено, что достаточно точные оценки для самих стандартных ошибок можно получить, заменяя значения истинных параметров их эмпирическими значениями.

Поскольку в большинстве случаев статистики имеют асимптотически нормальное распределение, оценка степени их точности сводится к нахождению стандартных ошибок. В этой главе мы остановимся на основных методах нахождения стандартных ошибок и фактически их найдем для различных статистик, рассмотренных нами в предыдущих главах. Чтобы избежать трудностей, возникающих при обращении с квадратными корнями, удобнее пользоваться выборочными дисперсиями и ковариациями. Так, для статистики t дисперсию ее выборочного распределения обозначим Dt . Ковариацию статистик t и u , т. е. смешанный момент их совместного выборочного распределения, мы обозначим $\text{cov}(t, u)$.

10.2 По определению r -й момент статистики t , т. е. r -й момент ее выборочного распределения, есть среднее значение величины t^r , взятое по всевозможным выборкам. Оно будет записываться $M(t^r)$ (см. 2.27). Если совместное распределение величин $x_1, x_2, \dots, x_n$, на основе которых вычислена статистика t , есть $dF(x_1, \dots, x_n)$, то r -й момент от t является интегралом от

$t dF$, взятым по всевозможным значениям x и y . Если, в частности, рассматривается простой случайный выбор и теоретическое распределение есть dF , то

$$M(t^r) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} t^r dF(x_1) \dots dF(x_n). \quad (10.1)$$

В этой главе мы будем главным образом интересоваться первыми и вторыми выборочными моментами t , т. е. средними и дисперсиями выборочных распределений. Напомним, что среднее значение суммы есть сумма средних значений, и если величины независимы, то среднее значение произведения есть произведение средних значений (2.28). Эти два результата в дальнейшем будут постоянно использоваться.

Стандартные ошибки моментов

10.3 В следующих пунктах мы будем придерживаться обычного соглашения: величины, относящиеся к генеральной совокупности, обозначаются греческими буквами, а относящиеся к статистикам, — латинскими. В соответствии с этим r -й момент теоретического распределения обозначается μ'_r , выборочный момент — статистика относительно начала отсчета —

$$m'_r = \frac{1}{n} \sum_{j=1}^n x_j^r, \quad (10.2)$$

а выборочный момент — статистика относительно среднего —

$$m_r = \frac{1}{n} \sum_{j=1}^n (x_j - m'_1)^r. \quad (10.3)$$

10.4 Рассмотрим среднее значение статистики m'_r . Имеем

$$M(m'_r) = \frac{1}{n} \sum M(x^r) = M(x^r) = \mu'_r. \quad (10.4)$$

Выборочная дисперсия статистики m'_r по определению равна $M(m'_r - \mu'_r)^2$, откуда, в предположении существования нужного числа моментов, следует, что

$$\begin{aligned} Dm'_r &= M\left\{\frac{1}{n} \sum x^r - \mu'_r\right\}^2 = \frac{1}{n^2} M\left[\left\{\sum x^r\right\}^2 - 2n\mu'_r \sum x^r + n^2\mu_r'^2\right] = \\ &= \frac{1}{n^2} M\left[\left\{\sum x^r\right\}^2\right] - \mu_r'^2 = \frac{1}{n^2} M\left\{\sum x_j^r + \sum (x_j^r x_k^r)\right\} - \mu_r'^2, \end{aligned}$$

где второе суммирование распространено по всем $n(n-1)$ перестановкам индексов, в которых $j \neq k$. Поскольку иксы независимы, среднее значение произведения равно произведению средних значений, откуда

$$Dm'_r = \frac{1}{n^2} \{n\mu'_{2r} + n(n-1)\mu'^2_r\} - \mu'^2_r = \frac{1}{n} (\mu'_{2r} - \mu'^2_r). \quad (10.5)$$

Заметим, что этот результат является точным. Тем же самым путем, если у нас есть два момента, m'_q и m'_r , мы получим формулу для выборочной ковариации

$$\begin{aligned} \text{cov}(m'_q, m'_r) &= M\{(m'_q - \mu'_q)(m'_r - \mu'_r)\} = \\ &= M\left\{\left(\frac{1}{n} \sum x^q - \mu'_q\right)\left(\frac{1}{n} \sum x^r - \mu'_r\right)\right\} = \\ &= \frac{1}{n^2} M\left(\sum x^{q+r}\right) + \frac{1}{n^2} M \sum_{j \neq k} (x^q_j x^r_k) - \frac{1}{n} \mu'_q M\left(\sum x^r\right) - \\ &\quad - \frac{1}{n} \mu'_r M\left(\sum x^q\right) + M(\mu'_q \mu'_r) = \frac{1}{n} (\mu'_{q+r} - \mu'_q \mu'_r). \end{aligned} \quad (10.6)$$

В случае $q=r$ это сводится к (10.5), как и должно быть в силу того, что ковариация самой случайной переменной есть ее дисперсия.

10.5 Формулы для центральных моментов не столь уже просты, поскольку среднее само подвержено выборочным флуктуациям. Из (10.5), полагая $r=1$, мы получим

$$Dm'_1 = (\mu'_2 - \mu'^2_1)/n = \mu_2/n. \quad (10.7)$$

Отсюда видно, что флуктуации в m'_1 имеют порядок $n^{-1/2}$. Мы можем поэтому с этой точностью считать, что эмпирическое среднее совпадает с истинным, и пренебрегать в m'_1 членами порядка выше первого. С точностью до членов порядка $n^{-1/2}$

$$\begin{aligned} M(m_r) &= M\left\{\sum (x - m'_1)^r / n\right\} = M\left\{\sum x^r - r m'_1 \sum x^{r-1}\right\} / n = \\ &= M\left\{(1 - r/n) \sum x^r - (r/n) \sum x_j x^{r-1}_k\right\} / n, \quad j \neq k. \end{aligned}$$

Выражение справа содержит моменты $\mu'_1 \mu'_{r-1}$, которые равны нулю, если положить $\mu'_1 = 0$. Следовательно, с точностью до членов порядка $n^{-1/2}$

$$M(m_r) = \mu_r. \quad (10.8)$$

Еще раз отметим, что, в то время как результат (10.4) является точным, формула (10.8) верна только с точностью до членов

порядка $n^{-1/2}$. Аналогично можно получить

$$\begin{aligned} Dm_r &= M(m_r^2) - \mu_r^2 = M\left\{\sum x^r - r \sum x_j x_k^{r-1}/n\right\}^2/n^2 - \mu_r^2 = \\ &= \frac{1}{n^2} M\left\{\sum x^{2r} + \sum x_j^r x_k^r + \frac{r^2}{n^2} \sum x_j^2 x_k^{2r-2} + \right. \\ &\quad \left. + \frac{r^2}{n^2} \sum x_j^2 x_k^{r-1} x_l^{r-1} - \frac{2r}{n} \sum x_j^{r+1} x_k^{r-1}\right\} - \mu_r^2, \end{aligned}$$

где $j \neq k \neq l$. Средние значения других членов, возникающих при возведении m_r в квадрат, обращаются в нуль, поскольку они содержат μ'_1 . Среднее значение суммы $\sum x_j^2 x_k^{2r-2}$ имеет порядок $n(n-1)$, и поэтому общий ее вклад в дисперсию имеет порядок n^{-2} и ею можно пренебречь. С учетом сказанного мы получим с точностью до порядка n^{-1}

$$Dm_r = \frac{1}{n} (\mu_{2r} - \mu_r^2 + r^2 \mu_2 \mu_{r-1}' - 2r \mu_{r-1} \mu_{r+1}), \quad (10.9)$$

и аналогично

$$\begin{aligned} \text{cov}(m_r, m_q) &= \frac{1}{n} (\mu_{r+q} - \mu_r \mu_q + r q \mu_2 \mu_{r-1} \mu_{q-1} - \\ &\quad - r \mu_{r-1} \mu_{q+1} - q \mu_{r+1} \mu_{q-1}). \end{aligned} \quad (10.10)$$

Пример 10.1

Для распределения мужчин по их росту (таблица 1.7) мы нашли (см. примеры 2.1 и 2.7), что $m'_1 = 67,46$, $\sqrt{m'_2} = 2,57$. Предположим теперь, что мы рассматриваем это распределение как возникшее в результате простого случайного выбора из генеральной совокупности взрослых мужчин, проживающих в Соединенном Королевстве в тот период, когда производился сбор этих данных. Что можно сказать относительно среднего значения всей генеральной совокупности?

Стандартная ошибка среднего зависит от величины μ_2 , точное значение которой нам неизвестно. Но в соответствии с общими принципами в случае большого объема наблюдаемых данных вместо μ_2 можно использовать его эмпирическое значение m_2 . Тогда найдем, что стандартная ошибка среднего m'_1 равна $2,57/\sqrt{8585} \sim 0,028$. Поэтому можно сказать, что среднее значение генеральной совокупности лежит, вероятно, в интервале $67,46 \pm 2(0,028)$ и, очень вероятно, в интервале $67,46 \pm 3(0,028)$. Ошибка в определении среднего значения будет, почти наверное, меньше десятой доли дюйма.

Пример 10.2

Из уравнения (10.9) с $r=4$ следует, что

$$Dm_4 = \frac{1}{n} (\mu_8 - \mu_4^2 - 8\mu_5\mu_3 + 16\mu_2\mu_3^2).$$

В главе 12 мы получим этот результат с помощью более точных методов. В частности, отсюда будет следовать, что эта формула действительно с точностью до членов порядка n^{-1} является верной.

Пример 10.3

Показать, что в случае симметричной генеральной совокупности первый смешанный момент между эмпирическими моментами m'_r и m_r (r — четное число) с точностью до членов порядка n^{-1} равен нулю.

Так же как в 10.5, мы имеем по определению

$$\begin{aligned} \text{cov}(m'_r, m_r) &= \frac{1}{n^2} M \left[\left\{ \sum x^r \left\{ \sum x^r - \frac{r}{n} \sum x_j x_k^{r-1} \right\} \right\} \right] = \\ &= \frac{1}{n^2} M \left\{ \sum x^{r+1} - \frac{r}{n} \sum x_j^2 x_k^{r-1} \right\} = \frac{1}{n} (\mu_{r+1} - r\mu_2\mu_{r-1}). \end{aligned}$$

Если r четно, то μ_{r+1} и μ_{r-1} будут моментами нечетного порядка, которые для симметричных распределений обращаются в нуль. Отсюда

$$\text{cov}(m'_r, m_r) = 0.$$

На языке теории корреляции это означает, что выборочное среднее и выборочный центральный момент (относительно выборочного среднего) с точностью до членов порядка n^{-1} некоррелированы.

Стандартные ошибки функций от случайных величин

10.6 Если $\xi_1, \xi_2, \dots, \xi_k$ — случайные величины, принимающие значения $x_1, x_2, \dots, x_k$, то определим функцию $g(\xi_1, \dots, \xi_k)$ как случайную величину, принимающую значение $g(x_1, \dots, x_k)$ при $\xi_1 = x_1, \dots, \xi_k = x_k$.

Предположим, что случайные величины ξ_i имеют средние θ_i и дисперсии порядка n^{-1} , а случайная величина g имеет конечное среднее и дисперсию. Если функция $g(x_1, \dots, x_k)$ дифференцируема в точках $x_i = \theta_i$, то, обозначая символом $\partial g / \partial \theta_i$

значение $\partial g / \partial x_i$ при $x_i = \theta_i$, мы получим

$$\Delta g = \sum_{i=1}^k \frac{\partial g}{\partial \theta_i} \Delta x_i + O[(\Delta x)^2], \quad (10.11)$$

где $\Delta x_i = x_i - \theta_i$, $(\Delta x)^2 = \sum (\Delta x_i)^2$ и $\Delta g = g(x) - g(0)$.

Для случая выборок большого объема среднее значение $(\Delta x)^2$ имеет порядок $o(1)$ и, с точностью до величин первого порядка малости, $M(\Delta g) = 0$, поскольку $M(\Delta x_i) = 0$. Следовательно, среднее значение случайной величины $g(\xi_1, \dots, \xi_k)$ равно $g(\theta_1, \dots, \theta_k)$. Точно так же, с точностью до величин первого порядка малости, в предположении, что производные не равны нулю, имеем

$$Dg = M\left(\sum \frac{\partial g}{\partial \theta_i} \Delta x_i\right)^2 = \sum_{i,j=1}^k \left\{ \frac{\partial g}{\partial \theta_i} \frac{\partial g}{\partial \theta_j} \text{cov}(x_i, x_j) \right\}, \quad (10.12)$$

где индексы i и j могут быть и равными (в последнем случае $\text{cov}(x_i, x_j) = Dx_i$).

Полученная формула имеет чрезвычайно важное значение. Она дает нам возможность представлять дисперсию функции от случайных величин через их дисперсии и ковариации.

Аналогичным путем устанавливается, что для двух функций от случайных величин, скажем g и h ,

$$\text{cov}(g, h) = \sum_{i,j=1}^k \left\{ \frac{\partial g}{\partial \theta_i} \frac{\partial h}{\partial \theta_j} \text{cov}(x_i, x_j) \right\}. \quad (10.13)$$

10.7 Отметим три интересных частных случая формулы (10.12).

(а) Пусть g — функция от одной переменной. Тогда

$$Dg(x) = \left(\frac{dg}{d\theta} \right)^2 Dx. \quad (10.14)$$

В частности, если $g(x)$ — линейная функция $cx + e$, то $Dg = c^2 Dx$, что, конечно, очевидно и так. Заметим, что в этом случае результат оказывается не приближенным, а точным.

(б) Пусть g является линейной функцией от случайных величин

$$g(x_1, \dots, x_k) = \sum_{i=1}^k a_i x_i. \quad (10.15)$$

Тогда в силу (10.12)

$$Dg = \sum a_i^2 Dx_i + \sum_{i \neq j} a_i a_j \text{cov}(x_i, x_j), \quad (10.16)$$

что также является точным результатом. В частности, когда иксы независимы, дисперсия g равна взвешенной сумме отдельных дисперсий.

(в) Пусть функция g является отношением x_1/x_2 . Чтобы избежать возможных здесь трудностей (см. следующую главу), мы предположим, что $x_2 > 0$, если распределение x_2 дискретно, и $x_2 \geq 0$, если оно непрерывно. Тогда из (10.12) следует, что

$$D(x_1/x_2) = \frac{Dx_1}{\theta_2^2} + \frac{\theta_1^2 Dx_2}{\theta_2^4} - \frac{2\theta_1 \text{cov}(x_1, x_2)}{\theta_2^3} = \\ = \left\{ \frac{M(x_1)}{M(x_2)} \right\}^2 \left\{ \frac{Dx_1}{M^2(x_1)} + \frac{Dx_2}{M^2(x_2)} - \frac{2 \text{cov}(x_1, x_2)}{M(x_1)M(x_2)} \right\}. \quad (10.17)$$

Второй множитель в (10.17) есть сумма квадратов коэффициентов вариации за вычетом удвоенной величины, модуль которой по аналогии можно было бы назвать квадратом коэффициента ковариации.

Пример 10.4

Чтобы найти выборочную дисперсию четвертого семиинварианта, будем для семиинварианта выборки вместо k пользоваться обозначением κ (причины этого станут ясными в главе 12). Итак, требуется найти дисперсию величины

$$\kappa_4 = m_4 - 3m_2^2.$$

Из (10.12) с $x_1 = m_4$, $x_2 = m_2$ и $g = x_1 - 3x_2^2$ мы получим

$$D\kappa_4 = Dm_4 + 36\mu_2^2 Dm_2 - 12\mu_2 \text{cov}(m_4, m_2).$$

Пользуясь формулами (10.9) и (10.10), найдем

$$D\kappa_4 = \frac{1}{n} \{ \mu_8 - 12\mu_6\mu_2 - 8\mu_5\mu_3 - \mu_4^2 + 48\mu_4\mu_2^2 + 64\mu_3^2\mu_2 - 36\mu_2^4 \}.$$

В случае нормальных совокупностей справедливы соотношения $\mu_4 = 3\mu_2^2 = 3\sigma^4$, $\mu_6 = 15\sigma^6$, $\mu_8 = 105\sigma^8$, откуда

$$D\kappa_4 = 24\sigma^8/n.$$

Пример 10.5

Найти выборочную дисперсию коэффициента вариации Карла Пирсона

$$V = 100 \sqrt{m_2/m'_1}.$$

Используя формулу (10.17) и обозначая V коэффициент вариации генеральной совокупности (как и μ'_1 , он предполагается

отличным от нуля), мы получим

$$D(V/100) = \left(\frac{V}{100}\right)^2 \left\{ \frac{D(\sqrt{m_2})}{M^2(\sqrt{m_2})} + \frac{Dm'_1}{M^2(m'_1)} - \frac{2 \operatorname{cov}(\sqrt{m_2}, m'_1)}{M(\sqrt{m_2}) M(m'_1)} \right\}. \quad (10.18)$$

Согласно (10.7) $Dm_1 = \mu_2/n$. И точно так же из (10.14) и (10.9)

$$D(\sqrt{m_2}) = \frac{Dm_2}{4\mu_2} = \frac{\mu_4 - \mu_2^2}{4n\mu_2}.$$

Наконец, в примере 10.3 мы видели, что, с точностью до членов порядка n^{-1} ,

$$\operatorname{cov}(m_2, m'_1) = \mu_3/n,$$

и согласно (10.13)

$$\operatorname{cov}(\sqrt{m_2}, m'_1) = \frac{1}{2\sqrt{\mu_2}} \operatorname{cov}(m_2, m'_1) = \mu_3/(2n\sqrt{\mu_2}).$$

Подставляя эти значения в (10.18), находим

$$DV = \frac{V^2}{n} \left\{ \frac{\mu_4 - \mu_2^2}{4\mu_2^2} + \frac{\mu_2}{\mu_1'^2} - \frac{\mu_3}{\mu_2\mu_1'} \right\}, \quad (10.19)$$

откуда для нормального случая ($\mu_3 = 0$, $\mu_4 = 3\mu_2^2$) получаем

$$DV = \frac{V^2}{n} \left\{ \frac{1}{2} + \frac{\mu_2}{\mu_1'^2} \right\} = \frac{V^2}{2n} \left\{ 1 + 2 \left(\frac{V}{100} \right)^2 \right\} \sim \frac{V^2}{2n}. \quad (10.20)$$

10.8 Сделаем несколько замечаний об употреблении стандартных ошибок.

(а) Использование стандартных ошибок, строго говоря, законно только в том случае, когда распределение статистик стремится к нормальному. В других случаях в силу неравенства Бьенэме — Чебышева их можно использовать для оценки сходимости по вероятности. Для получения более точных утверждений обычно все же приходится рассматривать нормальный интеграл.

(б) К вопросу об асимптотической нормальности надо подходить с известной осторожностью, поскольку она не всегда имеет место. В самом деле, распределение многих статистик (таких как, например, эмпирические моменты, которые зависят от суммы случайных величин) в силу центральной предельной теоремы стремится к нормальному. Однако распределение ряда других статистик (примерами которых служат статистики, зависящие от крайних членов) не является асимптотически нормальным.

(в) Вспомним в связи с этим, что распределение одних статистик стремится к нормальному быстрее, чем других, и

в одном случае заданное число наблюдений n может оказаться достаточно большим, для других же случаев — нет. Обычно (но не всегда) можно считать, что значения $n \geq 500$ являются «большими». Значения $n \geq 100$ также оказываются для многих целей достаточно «большими». С этой точки зрения к значениям, меньшим 100, надо относиться очень осторожно. Значения, меньшие 30, очень редко можно считать «большими». Распределения некоторых статистик, построенные по выборочным моментам, сходятся к нормальному очень медленно. В этом отношении примером могут служить статистики $\sqrt{b_1}$ или b_1 , являющиеся выборочными значениями соответственно $\sqrt{\beta_1}$ и β_1 . В этом случае методу стандартных ошибок следует предпочесть более точные методы, излагаемые в главе 12.

(г) Выборочное распределение дисперсии зависит от тех статистик, с помощью которых эта дисперсия получена. Так, в нормальном случае стандартное отклонение в выборке можно подсчитывать, например, как взятое $\sqrt{\frac{1}{2} \pi}$ раз значение среднего отклонения вместо обычно используемого корня квадратного из среднего квадратичного относительно среднего. Но при этом формула $Ds = (\mu_4 - \mu_2^2)/(4n\mu_2)$, которую можно получить из (10.9), будет неприменима.

(д) Из формул (10.5) и (10.9) видно, что выборочная дисперсия моментов порядка r зависит от моментов генеральной совокупности до порядка $2r$ включительно и становится очень большой для моментов высоких порядков даже тогда, когда n велико. Это объясняет, почему такие моменты имеют ограниченное практическое применение.

(е) Нужно не забывать о том, что многие полученные нами выше формулы являются приближенными. Поэтому если, например, коэффициент вариации выборки $V=0$, то согласно формулам примера 10.5 $DV=0$; если забыть, что наши формулы приближенные, это будет означать, что коэффициент вариации генеральной совокупности также равен нулю. Но на самом деле это, конечно, не так, хотя это и весьма правдоподобно, когда n велико.

(ж) Интересно выборочные флуктуации, отражаемые величиной выборочной дисперсии, сравнить с поправками Шеппарда для моментов. Пусть, на время, s_1^2 — выборочная дисперсия без учета поправок, s_2^2 — с их учетом. Тогда

$$\frac{s_2^2}{s_1^2} = 1 - \frac{1}{12} \frac{h^2}{s_1^2},$$

где h — ширина интервала группирования. Во многих практических случаях, если d — число интервалов, то dh примерно равно $6s_1$, и поэтому

$$\frac{s_2^2}{s_1^2} = 1 - \frac{3}{d^2}, \quad \frac{s_2}{s_1} \sim 1 - \frac{3}{3d^2}.$$

Обозначая буквой s корень квадратный из дисперсии выборки m_2 , мы получим

$$Ds = (\mu_4 - \mu_2^2)/(4n\mu_2),$$

что для нормального распределения равно $\sigma^2/(2n)$.

Поэтому если n равно, скажем, 1000, то стандартная ошибка в s примерно равна $0,0224\sigma$, что составляет $2,24\%$ σ . Поправки Шеппарда в случае $d=20$ дают только $0,375\%$ s_1 , т. е. равны примерно одной шестой стандартной ошибки σ . Поправки желательно делать и в том случае, когда n меньше 1000, чтобы избежать систематических ошибок. Предостережем, однако, читателя от мысли о том, что поправки обеспечивают более высокую степень надежности в оценке истинных значений параметров.

Заметим, что аналогичные рассуждения а fortiori применимы и к моментам высших порядков.

Стандартные ошибки смешанных моментов

10.9 Распространение вышеприведенных формул на случай смешанных моментов не представляет особого труда, за исключением разве что несколько более сложных алгебраических преобразований. Читателю предоставляется самостоятельно проверить следующие формулы:

$$D(m'_{r,s}) = \frac{1}{n} (\mu'_{2r,2s} - \mu'^2_{r,s}), \quad (10.21)$$

$$\text{cov}(m'_{r,s}, m'_{u,v}) = \frac{1}{n} (\mu'_{r+u,s+v} - \mu'_{r,s} \mu'_{u,v}), \quad (10.22)$$

$$\begin{aligned} D(m_{r,s}) = & \frac{1}{n} (\mu_{2r,2s} - \mu_{r,s}^2 + r^2 \mu_{2,0} \mu_{r-1,s}^2 + \\ & + s^2 \mu_{0,2} \mu_{r,s-1}^2 + 2rs \mu_{1,1} \mu_{r-1,s} \mu_{r,s-1} - \\ & - 2r \mu_{r+1,s} \mu_{r-1,s} - 2s \mu_{r,s+1} \mu_{r,s-1}), \end{aligned} \quad (10.23)$$

$$\begin{aligned} \text{cov}(m_{r,s}, m_{u,v}) = & \frac{1}{n} (\mu_{r+u,s+v} - \mu_{r,s} \mu_{u,v} + r \mu_{2,0} \mu_{r-1,s} \mu_{u-1,v} + \\ & + s \mu_{0,2} \mu_{r,s-1} \mu_{u,v-1} + r \mu_{1,1} \mu_{r-1,s} \mu_{u,v-1} + \\ & + s \mu_{1,1} \mu_{r,s-1} \mu_{u-1,v} - \mu_{r+1,s} \mu_{u-1,v} - \mu_{r,s+1} \mu_{u,v-1} - \\ & - r \mu_{r-1,s} \mu_{u+1,v} - s \mu_{r,s-1} \mu_{u,v+1}). \end{aligned} \quad (10.24)$$

Не опасаясь путаницы, далее вместо $\mu_{r,s}$ мы будем писать μ_{rs} .

Пример 10.6

Коэффициент корреляции в выборке определяется формулой $r = m_{11}/(m_{20}m_{02})^{1/2}$. Теоретический коэффициент корреляции равен $\rho = \mu_{11}/(\mu_{20}\mu_{02})^{1/2}$. Из (10.7) имеем

$$Dr = \rho^2 \left\{ \frac{Dm_{11}}{\mu_{11}^2} + \frac{D(m_{20}m_{02})^{1/2}}{\mu_{20}\mu_{02}} - 2 \frac{\text{cov}\{m_{11}, (m_{20}m_{02})^{1/2}\}}{\mu_{11}(\mu_{20}\mu_{02})^{1/2}} \right\},$$

откуда, подсчитывая $D(m_{20}m_{02})^{1/2}$ и $\text{cov}\{m_{11}, (m_{20}m_{02})^{1/2}\}$ по формулам (10.12) и (10.13), найдем

$$Dr = \rho^2 \left\{ \frac{Dm_{11}}{\mu_{11}^2} + \frac{1}{4} \left(\frac{Dm_{20}}{\mu_{20}^2} + \frac{Dm_{02}}{\mu_{02}^2} + \frac{2 \text{cov}(m_{20}, m_{02})}{\mu_{20}\mu_{02}} \right) - \right. \\ \left. - \left(\frac{\text{cov}(m_{11}, m_{02})}{\mu_{11}\mu_{02}} + \frac{\text{cov}(m_{11}, m_{20})}{\mu_{11}\mu_{20}} \right) \right\}.$$

В силу формул (10.23) и (10.24) окончательно получим

$$Dr = \frac{\rho^2}{n} \left\{ \frac{\mu_{22}}{\mu_{11}^2} + \frac{1}{4} \left(\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} \right) - \left(\frac{\mu_{31}}{\mu_{11}\mu_{20}} + \frac{\mu_{13}}{\mu_{11}\mu_{02}} \right) \right\}.$$

Для двумерного нормального распределения, воспользовавшись данными из примера 3.17, находим

$$Dr = \frac{1}{n} (1 - \rho^2)^2.$$

Не рекомендуется, однако, применять стандартную ошибку для проверки гипотетического ненулевого значения ρ , так как выборочное распределение r очень медленно сходится к нормальному. К этому вопросу мы еще вернемся в 16.23—16.33.

Стандартные ошибки квантилей

10.10 Среди различных характеристик, служащих мерами расположения и рассеяния, рассмотренных в главе 2, была одна группа величин, названных квантилями. Эти величины не являются алгебраическими функциями от результатов наблюдений, и их выборочная дисперсия не может быть получена с помощью изложенных выше методов. Рассмотрим их теперь.

Предположим, что распределение генеральной совокупности представимо в виде $dF = f dx$. Вероятность того, что в выборке объема n имеется $(l-1)$ наблюдений меньше x_1 , одно значение лежит в интервале $x_1 \pm \frac{1}{2} dx_1$, а оставшиеся $(n-l)$ значений больше x_1 , пропорциональна

$$\{F(x_1)\}^{l-1} f(x_1) dx_1 \{1 - F(x_1)\}^{n-l} = F_1^{l-1} (1 - F_1)^{n-l} dF_1, \quad (10.25)$$

где $F_1 = F(x_1)$. Это выражение представляет собой плотность распределения l -го выборочного квантиля.

Положим $l = nq$, $n - l = n(1 - q) = np$. Распределение (10.25) имеет модальное значение, которое можно вычислить путем дифференцирования соответствующей плотности по x_1 . Предварительно логарифмируя, а потом дифференцируя, получим

$$(l-1) \frac{f_1}{F_1} - (n-l) \frac{f_1}{(1-F_1)} + \frac{f_1'}{f_1} = 0. \quad (10.26)$$

Этому уравнению удовлетворяет модальное значение $\bar{x}$. Для больших n отношение f_1'/f_1 мало по сравнению с другими членами в (10.26), когда l и $(n-l)$ велики. Пренебрегая этим слагаемым, с точностью до величин порядка n^{-1} получим

$$\frac{q}{F} - \frac{p}{1-F} = 0$$

или $F(\bar{x}) = q$, что согласуется с нашими общими представлениями. С точностью до величин порядка n^{-1} квантили выборки совпадают с теоретическими квантилями.

Изучим теперь распределение (10.25) в окрестности модального значения. Положим

$$F_1 = q + \xi.$$

Тогда с точностью до констант вместо (10.25) получим

$$(q + \xi)^{nq} (p - \xi)^{np}.$$

Логарифмируя и разлагая логарифм в ряд, получим (с точностью до констант)

$$\begin{aligned} nq \log \left(1 + \frac{\xi}{q} \right) + np \log \left(1 - \frac{\xi}{p} \right) = \\ = nq \left(+ \frac{\xi}{q} - \frac{1}{2} \frac{\xi^2}{q^2} \dots \right) + np \left(- \frac{\xi}{p} - \frac{1}{2} \frac{\xi^2}{p^2} \dots \right) = - \frac{n\xi^2}{2pq} + O(\xi^3). \end{aligned}$$

Для случая больших выборок ξ мало по сравнению с q . Пренебрегая членами высокого порядка малости, убеждаемся, что распределение ξ выражается формулой

$$dF \propto \exp \left(\frac{-n\xi^2}{2pq} \right) d\xi,$$

или с учетом нормирующего множителя

$$dF = \frac{1}{\sqrt{2\pi} \sqrt{pq/n}} \exp \left(\frac{-n\xi^2}{2pq} \right) d\xi. \quad (10.27)$$

Отсюда следует, что ξ асимптотически нормальна с дисперсией

$$D\xi = \frac{pq}{n}. \quad (10.28)$$

Для получения дисперсии l -го выборочного квантиля x_l заметим, что $d\xi = dF_1 = f_1 dx_1$, откуда

$$Dx_1 = \frac{pq}{nf_1^2}. \quad (10.29)$$

На практике эта формула часто применяется к сгруппированным частотным распределениям. В этом случае надо помнить, что в формуле (10.29) ордината исходного распределения f_1 является частотой, приходящейся на *единичный интервал*, соответствующий среднему значению x_1 , поскольку она служит наилучшей оценкой этой ординаты.

Пример 10.7

Если x_1 — медиана, то $p = q = \frac{1}{2}$ и $Dx_1 = 1/4\pi f_1^2$, где f_1 — ордината медианы. Например, если исходное распределение нормально с дисперсией σ^2 , то $f_1 = 0,39894/\sigma$ (см. таблицу 1 в приложении). Следовательно, стандартная ошибка выборочной медианы равна

$$\frac{\sigma}{\sqrt{n}} \frac{1}{2 \cdot 0,39894} = 1,2533 \frac{\sigma}{\sqrt{n}}.$$

Стандартная ошибка среднего выборки объема n из нормальной совокупности равна $\sigma/\sqrt{n}$, т. е. меньше, чем стандартная ошибка выборочной медианы.

10.11 Для нахождения ковариации двух квантилей обобщим уравнение (10.25). Если мы имеем случайную выборку объема n , то вероятность того, что $(l-1)$ ее значений будут меньше x_1 , одно лежит в интервале $x_1 \pm \frac{1}{2} dx_1$, $(n-l-m)$ лежат между x_1 и x_2 , одно попадает в интервал $x_2 \pm \frac{1}{2} dx_2$, и $(m-1)$ значений окажутся больше x_2 , пропорциональна

$$dF \propto F_1^{l-1} (F_2 - F_1)^{n-l-m} (1 - F_2)^{m-1} dF_1 dF_2, \quad (10.30)$$

где $F_1 = F(x_1)$, $F_2 = F(x_2)$. Положим $l = q_1 n$, $m = p_2 n$. Уравнения, аналогичные (10.26) и предназначенные для нахождения модальных значений, имеют вид

$$\frac{q_1}{F_1} - \frac{(q_2 - q_1)}{F_2 - F_1} = 0, \quad \frac{q_2 - q_1}{F_2 - F_1} - \frac{p_2}{1 - F_2} = 0,$$

откуда для предельных модальных значений получаем

$$\left. \begin{aligned} F(\check{x}_1) &= q_1, \\ F(\check{x}_2) &= q_2. \end{aligned} \right\} \quad (10.31)$$

(Условия относительной малости $f'(\bar{x})/\bar{f}(\bar{x})$ обычно выполняются.) Положим

$$\begin{aligned} F_1 &= q_1 + \xi_1, \\ F_2 &= q_2 + \xi_2. \end{aligned}$$

Тогда совместное распределение ξ_1 и ξ_2 можно записать как

$$dF \propto (q_1 + \xi_1)^{q_1 n} (q_2 - q_1 + \xi_2 - \xi_1)^{(q_2 - q_1) n} (p_2 - \xi_2)^{p_2 n} d\xi_1 d\xi_2.$$

Так же как и в предыдущем пункте, беря логарифмы, разлагая их в ряды и пренебрегая членами порядка ξ^3 и выше, получим

$$dF \propto \exp \left\{ -\frac{n}{2(q_2 - q_1)} \left(\frac{q_2}{q_1} \xi_1^2 - 2\xi_1 \xi_2 + \frac{p_1}{p_2} \xi_2^2 \right) \right\} d\xi_1 d\xi_2. \quad (10.32)$$

Тем самым совместное распределение ξ_1 и ξ_2 стремится к двумерному нормальному распределению. Сравнивая (10.32) со стандартной формой двумерного нормального распределения (пример 3.17), мы видим, что

$$\begin{aligned} \frac{1}{(1 - \rho^2) D\xi_1} &= \frac{nq_2}{(q_2 - q_1) q_1}, \\ \frac{1}{(1 - \rho^2) D\xi_2} &= \frac{np_1}{(q_2 - q_1) p_2}, \\ \frac{\rho^2}{(1 - \rho^2) \text{cov}(\xi_1, \xi_2)} &= \frac{\rho}{(1 - \rho^2) \sqrt{D\xi_1 D\xi_2}} = \frac{1}{(q_2 - q_1)}. \end{aligned}$$

Отсюда легко можно найти

$$\left. \begin{aligned} D\xi_1 &= \frac{p_1 q_1}{n}, \\ D\xi_2 &= \frac{p_2 q_2}{n}, \\ \text{cov}(\xi_1, \xi_2) &= \frac{p_2 q_1}{n}. \end{aligned} \right\} \quad (10.33)$$

Заметим, что несимметричность формулы для ковариации отражает тот факт, что число p_2 необходимо относится к *верхнему* квантилю. Соответствующее выражение для ковариации x_1 и x_2 имеет вид

$$\text{cov}(x_1, x_2) = \frac{p_2 q_1}{n f_2 \bar{f}_1}. \quad (10.34)$$

Уравнения (10.33) и (10.34) могут быть использованы для нахождения дисперсии интерквантильной широты и других аналогичных статистик.

Пример 10.8

Дисперсия разности δ двух квантилей x_1 и x_2 в силу (10.16) задается формулой

$$D\delta = D x_1 + D x_2 - 2 \operatorname{cov}(x_1, x_2) = \frac{1}{n} \left\{ \frac{p_1 q_1}{f_1^2} + \frac{p_2 q_2}{f_2^2} - \frac{2 p_2 q_1}{f_1 f_2} \right\}.$$

Возьмем в качестве квантилей квартили $p_2 = q_1 = \frac{1}{4}$, $p_1 = q_2 = \frac{3}{4}$. Тогда для дисперсии семиинтерквартильной широты (с. и. ш.) мы получим формулу

$$D(\text{с. и. ш.}) = \frac{1}{64n} \left(\frac{3}{f_1^2} + \frac{3}{f_2^2} - \frac{2}{f_1 f_2} \right),$$

где f_1 и f_2 — частоты в единичных интервалах, отвечающих этим двум квартилям, причем f_2 соответствует большему квартилю.

Например, для нормального распределения $f_1 = f_2$ и поэтому

$$D(\text{с. и. ш.}) = \frac{1}{16n f_1^2}.$$

Из таблиц нормального распределения можно найти, что значение большего квартиля равно 0,6745 и значение плотности в этой точке равно $f_1 = 0,3178/\sigma$. Таким образом, стандартная ошибка выборочной семиинтерквартильной широты равна

$$\frac{\sigma}{\sqrt{n(4 \cdot 0,3178)}} = 0,7867 \frac{\sigma}{\sqrt{n}}.$$

10.12 Остановимся еще раз на важности вопроса о способах получения оценок неизвестных параметров, поскольку от этих способов зависят величины стандартных ошибок (ср. 10.8 (г)). Например, стандартное отклонение нормального распределения по результатам наблюдений можно оценить многими способами: с помощью второго момента; рассматривая среднее отклонение, умноженное на $\sqrt{\pi/2}$; беря семиинтерквартильную широту, деленную на 0,6745, и т. д. Каждый из этих методов будет давать свою стандартную ошибку. Для первого из них согласно 10.8 она равна $\sigma/\sqrt{2\pi}$, для третьего $1,64956/\sqrt{2\pi}$, как это нетрудно получить из примера 10.8. Естественно поэтому поставить вопрос о том, каковы те оценки, которые в случае их существования обладают минимальной выборочной дисперсией. Здесь мы не будем останавливаться на этом вопросе, а подчеркнем лишь важность того обстоятельства, что при использовании различных оценок неизвестных параметров надо не забывать о том, каким способом они получены.

Стандартная ошибка среднего отклонения

10.13 Предположим, что имеется некоторая выборка из генеральной совокупности со средним μ и дисперсией σ^2 .

Пусть d_α — среднее отклонение выборки, измеряемое относительно фиксированной точки α , т. е.

$$d_\alpha = \frac{1}{n} \sum_{i=1}^n |x_i - \alpha|.$$

Тогда

$$M(d_\alpha) = M|x_i - \alpha| = \delta_\alpha, \quad (10.35)$$

где δ_α — теоретическое значение среднего отклонения. Поскольку результаты наблюдений предполагаются независимыми, то

$$\begin{aligned} Dd_\alpha &= \frac{1}{n} D(|x - \alpha|) = \frac{1}{n} \{M(|x - \alpha|^2) - M^2(|x - \alpha|)\} = \\ &= \frac{1}{n} \{\sigma^2 + (\alpha - \mu)^2 - \delta_\alpha^2\}. \end{aligned} \quad (10.36)$$

Если α равно истинному среднему значению, т. е. $\alpha = \mu$, то

$$Dd_\mu = \frac{1}{n} (\sigma^2 - \delta_\mu^2). \quad (10.37)$$

Рассмотрим теперь среднее отклонение d_α относительно значения a , которое само подвержено выборочным флуктуациям и обладает тем свойством, что

$$M(a) = \mu, \quad Da = O(n^{-1}).$$

Тогда, так же как и в 10.5, можно установить, что приближенно

$$M(d_a) = \delta_a, \quad Dd_a = Dd_\alpha.$$

В частности, если $a = \bar{x}$, то для среднего отклонения от среднего выборки

$$Dd = \frac{1}{n} (\sigma^2 - \delta_\mu^2). \quad (10.38)$$

Для случая нормального распределения в (5.75) мы видели, что

$$\delta_\mu = \rho \sqrt{(2/\pi)},$$

и, следовательно,

$$Dd = \sigma^2 \left(1 - \frac{2}{\pi}\right) / n. \quad (10.39)$$

В нормальном случае выражение для Dd известно точно: Хельмерт (1876) и Фишер (1920) независимо установили, что

$$Dd = \frac{2\sigma^2(n-1)}{\pi n^2} \left[\frac{1}{2} \pi + \sqrt{n(n-1)} - n - \arcsin \{1/(n-1)\} \right]. \quad (10.40)$$

При больших n это выражение близко к (10.39). Годвин (1945) нашел точное выражение ф. р. статистики d . Эта ф. р. табулирована Хартли (1945) для $n=2(1) 10$.

Стандартная ошибка средней разности

10.14 Пусть g — выборочное значение средней разности Джини (без повторений), теоретическое значение которой (согласно 2.21) обозначено Δ . Тогда

$$g = \frac{1}{n(n-1)} \sum_{i,j=1}^n |x_i - x_j| \quad (10.41)$$

и

$$M(g) = M|x_i - x_j| = \Delta. \quad (10.42)$$

Аналогично

$$\begin{aligned} M(g^2) &= \frac{1}{n^2(n-1)^2} M\left\{\sum |x_i - x_j|\right\}^2 = \\ &= \frac{1}{n^2(n-1)^2} M\left\{\sum (x_i - x_j)^2 + \sum' |x_i - x_j| |x_i - x_k| + \right. \\ &\quad \left. + \sum'' |x_i - x_j| |x_k - x_l|\right\}, \quad i \neq j \neq k \neq l, \end{aligned}$$

где в суммировании $\sum$ содержится $2n(n-1)$ членов, в $\sum'$ — $4n(n-1)(n-2)$ и в $\sum''$ — $n(n-1)(n-2)(n-3)$ членов. После некоторых преобразований получим

$$\begin{aligned} M(g^2) &= \frac{1}{n(n-1)} \{2M(x_i - x_j)^2 + 4(n-2)M(|x_i - x_j| |x_i - x_k|) + \\ &\quad + (n-2)(n-3)M(|x_i - x_j| |x_k - x_l|)\}. \end{aligned}$$

Первое математическое ожидание равно $2\sigma^2$, третье — Δ^2 . Обозначим $\mathcal{J}$ второе математическое ожидание. Тогда

$$M(g^2) = \frac{1}{n(n-1)} \{4\sigma^2 + 4(n-2)\mathcal{J} + (n-2)(n-3)\Delta^2\}$$

и, следовательно,

$$Dg = \frac{1}{n(n-1)} \{4\sigma^2 + 4(n-2)\mathcal{J} - 2(2n-3)\Delta^2\}. \quad (10.43)$$

Что касается величины $\mathcal{J}$, то ее можно представить различными способами. Если $f(x)$ — плотность распределения, то

$$\mathcal{J} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x-y| |x-z| f(x) f(y) f(z) dx dy dz. \quad (10.44)$$

Пусть

$$Z(x) = \int_{-\infty}^x t f(t) dt$$

— неполный момент, тогда

$$\mathcal{J} = 4 \int_{-\infty}^{\infty} \{(xF - Z)^2 + (\mu - x)(xF - Z)\} f(x) dx. \quad (10.45)$$

(Переход от (10.44) к (10.45) предоставляется читателю в качестве упражнения.) Указанный вывод был дан Ломницки (1952). Более сложный вывод впервые был указан Нэйром (1936), который нашел формулы для дисперсии в следующих трех случаях:

(а) *Нормальное распределение*: любое среднее, дисперсия σ^2 ,

$$Mg = 2\sigma/\sqrt{\pi},$$

$$Dg = \frac{4\sigma^2}{n(n-1)} \left\{ \frac{1}{3}(n+1) + 2(n-2)\sqrt{3}/\pi - 2(n-3)/\pi \right\} \sim \sigma^2(0,8068)^2/n. \quad (10.46)$$

(б) *Экспоненциальное распределение*: $dF = \exp(-x/\sigma) \times \times d(x/\sigma)$ ($0 \leq x \leq \infty$).

$$Mg = \sigma,$$

$$Dg = \sigma^2 \frac{2(2n-1)}{3n(n-1)} \sim 4\sigma^2/(3n). \quad (10.47)$$

(в) *Равномерное распределение*: $dF = dx/k$ ($0 \leq x \leq k$).

$$Mg = \frac{1}{3}k,$$

$$Dg = \frac{k^2(n+3)}{45n(n-1)} \sim k^2/(45n). \quad (10.48)$$

Некоторые наиболее употребительные стандартные ошибки

10.15 По-видимому, полезно дать сводку наиболее часто встречающихся стандартных ошибок. Часть результатов в приводимой ниже таблице получена в этой главе, другие же можно непосредственно получить, используя развитые здесь методы. Некоторые из этих результатов относятся к статистикам, которые встретятся нам позже.

Для удобства записи в таблице, как правило, приводятся дисперсии, а не стандартные ошибки. Некоторые коэффициенты являются квадратами, так что переход к квадратным корням легко осуществим.

Статистики	Дисперсия, умноженная на n	Примечания
1. Среднее, m'_1	$\mu_2 (= \sigma^2)$	Результат верен для любого распределения с конечным вторым моментом μ_2
2. Дисперсия выборки, m_2	$\mu_4 - \mu_2^2$	Для нормального распределения $= 2\sigma^4$
3. Стандартное отклонение выборки, s	$(\mu_4 - \mu_2^2)/(4\mu_2)$	Для нормального распределения $= \sigma^2/2$
4. Третий момент, m_3	$\mu_6 - \mu_3^2 - 6\mu_4\mu_2 + 9\mu_2^3$	Для нормального распределения $= 6\sigma^6$
5. Четвертый момент, m_4	$\mu_8 - \mu_4^2 - 8\mu_5\mu_3 + 16\mu_2\mu_3^2$	Для нормального распределения $= 96\sigma^8$
6. b_1	$\beta_1 (4\beta_4 - 24\beta_2 + 36 + 9\beta_1\beta_2 - 12\beta_3 + 35\beta_1)$	Для нормального распределения дисперсия равна σ/n ; β_1 — коэффициент, определенный в уравнении (3.88). Отметим несостоятельность этой формулы для симметричных распределений с $\beta_1 = 0$ (ср. 10.8 и 12.18)
7. b_2	$\beta_6 - 4\beta_2\beta_4 + 4\beta_2^3 - \beta_2^2 + 16\beta_2\beta_1 - 8\beta_3 + 16\beta_1$	Для нормального распределения дисперсия равна $24/n$. См. также 12.18 и упражнение 12.12
8. Коэффициент вариации, V	См. пример 10.5	Для нормального распределения $\sim V^2/2$
9. Пирсоновская мода	См. Ясукава (1926) — формулы и таблицы	Распределение для умеренных значений n асимметрично
10. Среднее отклонение	См. 10.13	Для нормального распределения $= \sigma^2 \left(1 - \frac{2}{\pi}\right)$
11. Среднее разностей (Джини)	См. 10.14	Для нормального распределения $= \sigma^2 (0.8068)^2$
12. Медиана	$1/(4y_0^2)$, где y_0 — ордината плотности, отвечающая медиане	По поводу нормальных распределений и малых выборок см. Хойо (1931) и К. Пирсон (1931). Для случая нормального распределения и больших выборок $= (1.2533)^2 \sigma^2$

Статистики	Дисперсия, умноженная на n	Примечание
13. Квартили	$3/(16y^2)$, где y — ордината плотности, отвечающая квартилю	Для нормального распределения $= (1,3626)^2 \sigma^2$. См. также Хойо (1931)
14. Децили	См. 10.10	Для нормального распределения (децили 4,6) $= (1,2680)^2 \sigma^2$ (децили 3,7) $= (1,3180)^2 \sigma^2$ (децили 2,8) $= (1,4288)^2 \sigma^2$ (децили 1,9) $= (1,7094)^2 \sigma^2$
15. Семинтерквартильная ширина	$\frac{1}{4} \{ 3/16 y_1^2 + 3/16 y_2^2 - 1/8 y_1 y_2 \} \sigma^2$, где y_1 и y_2 — ординаты плотности в точках квартилей	Для нормального распределения $= (1,7867)^2 \sigma^2$
16. Коэффициент корреляции, r (смешанный момент)	См. пример 10.6	Для двумерного нормального распределения $= (1-r^2)^2$. Надежность оценки для коэффициента корреляции может быть гарантирована только для случая очень больших n . См. главу 16

УПРАЖНЕНИЯ

10.1 Для распределения ростов в таблице 1.7 было найдено

$$m_2 = 6,16, \quad m_3 = -0,207, \quad m_4 = 137,689.$$

Предполагая, что данные этой таблицы отвечают выборке из генеральной совокупности, имеющей приближенно нормальное распределение, показать, что m_3 незначительно отличается от нуля (что, конечно, так и должно быть, если распределение в действительности нормально) и что стандартная ошибка в определении m_4 составляет примерно 4% от его значения.

10.2 Пусть для некоторого многомерного распределения вероятности каких-то k событий равны $\omega_1, \omega_2, \dots, \omega_k$. Предположим, что в результате n наблюдений получены частоты $f_1, \dots, f_k$ появления этих событий. Показать, что для каждой функции T от них асимптотически

$$\frac{1}{n} DT = \sum_{i=1}^k \omega_i \left(\frac{\partial T}{\partial f_i} \right)^2 - \left(\frac{\partial T}{\partial n} \right)^2,$$

где производные взяты в точках $f_i = n\omega_i$. (Fisher, *Statistical Methods for Research Workers*.)

10.3 Известно, что для выборки объема 100 шансы того, что некоторая статистика отличается от соответствующего теоретического значения больше чем на две единицы, равны 1 к 99. Найти, как велика должна была бы быть выборка, чтобы уменьшить эти шансы до 1 к 199. Найти шансы того, что для выборки объема 100 статистика будет отличаться от соответствующего значения генеральной совокупности более чем на одну единицу.

10.4 Пусть некоторый признак A для данной генеральной совокупности имеет вероятность ω и пусть имеется k выборок разного объема. Показать, что стандартная ошибка средней частоты появления признака A равна

$$\sqrt{\frac{\omega(1-\omega)}{kH}}, \text{ где } H — \text{гармоническое среднее из объемов выборок.}$$

10.5 Пусть область значений некоторой варианти разбита на ряд подобластей. И пусть f_p — частота попадания в p -ю область за n испытаний, где n велико. Показать, что

$$Df_p = f_p \left(1 - \frac{f_p}{n}\right), \quad \text{cov}(f_p, f_q) = -\frac{1}{n} f_p f_q.$$

Найти выражение для выборочной дисперсии r -го момента относительно произвольной точки.

10.6 Показать, что для случая выборок большого объема из генеральной совокупности с симметричным распределением каждый момент выборки нечетного порядка не коррелирован (имеет нулевую ковариацию) с каждым моментом выборки четного порядка.

10.7 Пользуясь обозначениями этой главы, показать, что с точностью до n^{-1}

$$\text{cov}(m_r, m'_q) = (\mu_{q+r} - \mu_q \mu_r - r \mu_{q+1} \mu_{r-1})/n.$$

10.8 Показать, что для симметричного распределения ковариация двух моментов выборки четного порядка в случае выборок большого объема не может быть отрицательной.

10.9 В случае выборок большого объема дисперсия статистики m_2 равна $(\mu_4 - \mu_2^2)/n$. Возьмем в качестве оценки этой дисперсии статистику $(m_4 - m_2^2)/n$. Показать, что дисперсия $(m_4 - m_2^2)/n$ для случая нормального распределения равна $56\sigma^8/n^2$; сравнить этот результат со значением $Dm_2 = 2\sigma^4/n$.

10.10 Доказать формулу для Dh_1 , данную в 10.15.

10.11 Проверить соотношение (10.46).

10.12 Показать, что для первых четырех семиквариантов, полученных из моментов, с точностью до членов порядка n^{-1}

$$D\kappa_1 = \frac{1}{n} \kappa_2,$$

$$D\kappa_2 = \frac{1}{n} (\kappa_4 + 2\kappa_2^2),$$

$$D\kappa_3 = \frac{1}{n} (\kappa_6 + 9\kappa_4\kappa_2 + 9\kappa_3^2 + 6\kappa_2^3),$$

$$D\kappa_4 = \frac{1}{n} (\kappa_8 + 16\kappa_6\kappa_2 + 48\kappa_5\kappa_3 + 34\kappa_4^2 + 72\kappa_4\kappa_2^2 + 144\kappa_3^2\kappa_2 + 24\kappa_2^4).$$

(З а м е ч а н и е. Вопреки обычному соглашению, в левых частях этих формул используются греческие буквы для обозначения отвечающих выбор-

кам величин k ; если придерживаться обычных обозначений, то это могло бы привести к несогласованности с обозначениями k -статистик главы 12.)

10.13 Показать, что среднее значение дисперсии выборки равно

$$M(m_2) = \frac{n-1}{n} \mu_2$$

и

$$Dm_2 = \left(\frac{n-1}{n}\right)^2 \frac{\mu_4 - \mu_2^2}{n} + \frac{2(n-1)}{n^3} \mu_2^2.$$

С помощью этого установить, что формулы настоящей главы, применяемые к дисперсии, имеют точность порядка n^{-1} .

10.14 Доказать, что стандартная ошибка первого дециля в случае нормальной совокупности равна $1,7094\sigma/\sqrt{n}$.

10.15 Показать, что в случае нечетного n выборочная дисперсия медианы для равномерного распределения на $[0, 1]$ в точности равна $1/\{4(n+2)\}$.

10.16 Пусть две случайные переменные независимы, распределение каждой из них симметрично и $(r+s)$ — четное число. Показать, что тогда в обозначениях этой главы

$$n \operatorname{cov}(m_r, s, m_u, v) = \mu_{r+u, 0} \mu_{0, s+v} - \mu_{r0} \mu_{u0} \mu_{0s} \mu_{0v}.$$

10.17 Пусть τ — некоторая числовая характеристика теоретического распределения и t — ее оценка. Предположим, что с точностью до n^{-1} $D_t = f(\tau)$. Показать, что если

$$u(t) = \int_0^t \{f(t)\}^{-1/2} dt,$$

то для случая выборок большого объема $Du = 1$. Проверить это для коэффициента корреляции, для которого при больших n

$$Dr = (1 - \rho^2)^2/n.$$

10.18 В случае выборок большого объема и нормального распределения в качестве оценки четвертого момента μ_4 можно взять: (а) m_4 и б) $3m_2^2$. Показать, что дисперсия оценки (а) равна $4/3$ дисперсии оценки (б).

10.19 Показать, что для выборок объема n из равномерного распределения $dF = kdx$, $0 \leq x \leq 1/k$, дисперсия среднего отклонения равна примерно $1/(48k^2n)$.

10.20 Пусть $x_1, \dots, x_n$ — результаты наблюдений над одной случайной величиной и $y_1, \dots, y_m$ — над другой, имеющей ту же самую дисперсию μ_2 , что и первая. Пусть среднее значение из $m+n$ наблюдений равно z . Показать, что стандартная ошибка разности между z и $\bar{x}$ равна

$$\sqrt{\frac{m\mu_2}{n(m+n)}}.$$

10.21 Две группы состоят из M и N предметов каждая, причем D предметов у них совпадают. Независимым и случайным образом из каждой группы берется (без возвращения) соответственно m и n предметов ($m \geq n \geq 2$). Обнаружено, что d предметов у них общие. Показать, что несмещенная оценка для D дается формулой

$$\hat{D} = \frac{MN}{mn} d$$

и что

$$D(\hat{D}) = \frac{MND}{mn} \left\{ 1 + \frac{m-1}{M-1} \frac{n-1}{N-1} (D-1) \right\} - D^2.$$

Показать, что при $d \rightarrow n$ $\hat{D}$ превышает N и что, следовательно, оценка принимает «невозможные» значения. (Ср. Гудмэн (L. A. Goodman) (1952), *Ann. Math. Statist.* 23, 632), см. также Деминг и Глэссер (W. E. Deming, G. J. Glaser) (1959), *J. Amer. Statist. Ass.* 54, 403.)

10.22 В упражнении 10.21 показать, что несмещенная оценка для $D(\hat{D})$ равна

$$\hat{V} = \hat{D} \left(\frac{M-1}{m-1} \frac{N-1}{n-1} - 1 \right) + \hat{D}^2 \left(1 - \frac{mn}{MN} \frac{M-1}{m-1} \frac{N-1}{n-1} \right)$$

и что, следовательно, $\hat{V}$ может принимать отрицательные значения. Тем самым $\hat{V}$ является «невозможной» оценкой для $D(\hat{D})$ в том случае, когда m , n , M и N малы.

10.23 Пусть x и y имеют совместное распределение такое, что $M(x) = \mu'_{10} \neq 0$ и $M(y) = \mu'_{01} \neq 0$. Показать, что

$$D(x, y) = \mu'^2_{10}\mu_{02} + \mu_{20}\mu'^2_{01} + 2\mu'_{10}\mu_{11}\mu'_{01} + 2\mu'_{10}\mu_{12} + 2\mu_{21}\mu'_{01} + \mu_{22} - \mu'^2_{11}$$

(это точный результат) и что, следовательно, в случае независимых x и y

$$D(xy) = \mu'^2_{10}\mu_{02} + \mu_{20}\mu'^2_{01} + \mu_{20}\mu_{02}. \quad (A)$$

Установить, что с точностью до членов порядка n^{-1} этот результат можно получить из (10.12) (отличие будет состоять только в последнем члене формулы (A)), если считать, что μ_{20} и μ_{02} имеют порядок n^{-1} .

При условии, что m_{20} , m_{02} являются несмещенными оценками для μ_{20} и μ_{02} , показать, что несмещенная оценка для (A) равна

$$x^2 m_{02} + y^2 m_{20} - m_{20} m_{02}.$$

(Гудмэн (L. A. Goodman) (1960),
J. Amer. Statist. Ass. 55, 708.)

ТОЧНЫЕ ВЫБОРОЧНЫЕ РАСПРЕДЕЛЕНИЯ

11.1 Роль выборочного распределения в статистических выводах была уже отмечена в главе 7. Настоящая глава посвящается обзору основных методов нахождения этих распределений при заданной генеральной совокупности, из которой производится выборка. Как обычно, выбор будет предполагаться простым и случайным, так что если $dF(x)$ — элемент исходного распределения, то элементом совместного распределения n значений $x_1, \dots, x_n$ будет $dF(x_1) dF(x_2) \dots dF(x_n)$, и если z -статистика

$$z = z(x_1, \dots, x_n), \quad (11.1)$$

то функция распределения z выражается формулой

$$F(z_0) = \int \dots \int dF(x_1) \dots dF(x_n), \quad (11.2)$$

где интегрирование распространяется на область тех значений z , для которых $z(x_1, \dots, x_n) \leq z_0$.

Формула (11.2) является формальным решением нашей задачи: она сводит ее к чисто математической задаче вычисления некоторых многомерных интегралов или сумм. Рассматриваемые ниже различного рода методы предназначены для облегчения процесса интегрирования. Их можно расклассифицировать на четыре группы:

(а) непосредственное вычисление интеграла (11.2) с помощью обычных аналитических средств, таких как, например, подходящая замена переменных;

(б) использование для вычисления (11.2) геометрической терминологии с целью избежать громоздких аналитических формул;

(в) использование характеристических функций;

(г) применение других аналитических методов, включая математическую индукцию.

Вынося общие множители столбцов, находим, что детерминант здесь равен

$$\cos^{n-1} \theta_1 \cos^{n-2} \theta_2 \dots \cos \theta_{n-1} \sin \theta_1 \sin \theta_2 \dots \sin \theta_{n-1} \times$$

$$\times \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ -\operatorname{tg} \theta_1 & -\operatorname{tg} \theta_1 & -\operatorname{tg} \theta_1 & \dots & \operatorname{ctg} \theta_1 \\ -\operatorname{tg} \theta_2 & -\operatorname{tg} \theta_2 & -\operatorname{tg} \theta_2 & & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\operatorname{tg} \theta_{n-2} & -\operatorname{tg} \theta_{n-2} & \operatorname{ctg} \theta_{n-2} & \dots & 0 \\ -\operatorname{tg} \theta_{n-1} & \operatorname{ctg} \theta_{n-1} & 0 & \dots & 0 \end{vmatrix}.$$

Вычитая теперь каждый столбец из предыдущего, получаем окончательное выражение для детерминанта $\cos^{n-2} \theta_1, \cos^{n-3} \theta_2, \dots, \cos \theta_{n-2}$.

Вычисляемый нами интеграл примет, следовательно, вид

$$\int \dots \int \frac{1}{(2\pi)^{\frac{1}{2}n}} e^{-\frac{1}{2}z} \frac{1}{2} z^{\frac{1}{2}n-1} \cos^{n-2} \theta_1 \dots$$

$$\dots \cos \theta_{n-2} dz d\theta_1 \dots d\theta_{n-1}. \quad (11.6)$$

Преимущество новых переменных состоит в том, что пределы интегрирования стали значительно проще: z меняется от 0 до z , θ_{n-1} — от 0 до 2π , а остальные θ — от $-\frac{1}{2}\pi$ до $\frac{1}{2}\pi$. Интеграл (11.6) разбивается на произведение интегралов, причем интегралы по θ оказываются равными постоянным. Следовательно, функция распределения z равна

$$F(z) = k \int_0^z e^{-\frac{1}{2}z} z^{\frac{1}{2}n-1} dz. \quad (11.7)$$

Значение константы k можно получить интегрированием от 0 до ∞ . Вычисляя ее, находим, что искомое распределение имеет вид

$$dF = \frac{1}{2^{\frac{1}{2}n} \Gamma\left(\frac{1}{2}n\right)} e^{-\frac{1}{2}z} z^{\frac{1}{2}n-1} dz, \quad 0 \leq z \leq \infty, \quad (11.8)$$

и, следовательно, принадлежит к III типу Пирсона.

11.3 Существенным в замене переменных является переход к иной, более простой, области интегрирования, определяемой пределами изменения новых переменных. Обычно статистика, выборочное распределение которой ищется, берется в качестве

одной из новых переменных, а остальные $n-1$ из них выбираются каким-либо удобным для рассматриваемой задачи образом. Интеграл (11.2) принимает при этом вид

$$F(z) = \int \dots \int f(x_1) \dots \dots f(x_n) \frac{\partial(x_1, \dots, x_n)}{\partial(z, \theta_1, \dots, \theta_{n-1})} dz d\theta_1 \dots d\theta_{n-1}, \quad (11.9)$$

где $J = \frac{\partial(x_1, \dots, x_n)}{\partial(z, \theta_1, \dots, \theta_{n-1})}$ — соответствующий якобиан, $f(x_j)$ — функция плотности исходного распределения, а x_j рассматриваются как функции от z и θ . Далее производится интегрирование по θ , причем обычно переменные θ можно выбрать так, чтобы пределы этого интегрирования не зависели от z . Тем самым (11.2) заменяется более легко вычислимым интегралом.

Как обычно в подобных случаях, здесь имеется неопределенность знака J ; знак следует выбирать так, чтобы преобразованный интеграл был положительным. Законность применяемого преобразования зависит от известных условий, обеспечивающих возможность замены переменных в многомерных интегралах. Например, достаточно, чтобы новые переменные были непрерывны по x вместе со своими первыми производными и чтобы якобиан J не менял знака в области интегрирования. Несколько дальнейших примеров сделают эти общие рассуждения более ясными.

Пример 11.1

Найти распределение среднего значения выборки $x_1, \dots, x_n$ объема n из генеральной совокупности с распределением

$$dF = \frac{dx}{\pi(1+x^2)}, \quad -\infty \leq x \leq \infty.$$

Элемент совместного распределения $x_1, \dots, x_n$ имеет вид

$$\frac{1}{\pi^n} \prod_{j=1}^n \frac{dx_j}{(1+x_j^2)}, \quad (11.10)$$

а рассматриваемая статистика z определяется соотношением

$$nz = \sum_{j=1}^n x_j. \quad (11.11)$$

Нам нужно проинтегрировать (11.10) по множеству тех x , для которых $\sum x \leq nz$. Перейдем к новым переменным

$u_1 = x_1, u_2 = x_2, \dots, u_{n-1} = x_{n-1}, u_n = \frac{1}{n} \sum_{j=1}^n x_j = z$. Якобиан

равен здесь, очевидно, постоянной n . Новые переменные $u_1, \dots, u_{n-1}$ могут меняться от $-\infty$ до $+\infty$, а u_n — от $-\infty$ до z . Имеем, следовательно,

$$F(z) = \int_{-\infty}^z du_n \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{n}{\pi^n} \prod_{j=1}^{n-1} \frac{du_j}{(1+u_j^2)} \frac{1}{1+(nu_n-u_1-\dots-u_{n-1})^2}. \quad (11.12)$$

В этой формуле $(n-1)$ -кратный интеграл по $u_1, \dots, u_{n-1}$ представляет собой плотность распределения z . Его можно вычислить последовательно. Имеем

$$\frac{1}{(1+x^2)\{r^2+(a-x)^2\}} = \frac{1}{\{a^2+(r+1)^2\}\{a^2+(r-1)^2\}} \times \\ \times \left[\frac{2ax}{x^2+1} + \frac{a^2+r^2-1}{x^2+1} + \frac{2a^2-2ax}{r^2+(a-x)^2} + \frac{a^r-r^2+1}{r^2+(a-x)^2} \right].$$

Проинтегрируем обе части этого равенства по x от $-\infty$ до $+\infty$. Справа получим

$$\frac{1}{\{a^2+(r+1)^2\}\{a^2+(r-1)^2\}} \left[a \log(x^2+1) - a \log\{r^2+(a-x)^2\} + \right. \\ \left. + (a^2+r^2-1) \operatorname{arctg} x + \frac{a^2-r^2+1}{r} \operatorname{arctg} \frac{x-a}{r} \right]_{-\infty}^{\infty},$$

что сводится к

$$\pi \left(\frac{r+1}{r} \right) \left\{ \frac{1}{a^2+(r+1)^2} \right\}. \quad (11.13)$$

Производя в (11.12) интегрирование по u_{n-1} и пользуясь найденной формулой при $x=u_{n-1}$, $r=1$, $a=nu_n-u_1-\dots-u_{n-2}$, находим, что $(n-1)$ -кратный интеграл, задающий плотность z , равен

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{n}{\pi^{n-1}} \prod_{j=1}^{n-2} \frac{du_j}{(1+u_j^2)} \frac{2}{2^2+(nu_n-u_1-\dots-u_{n-2})^2}.$$

Далее, интегрируя последовательно по $u_{n-2}, u_{n-3}, \dots$, в конце концов, получим следующее выражение для плотности z :

$$\frac{n^2}{\pi \{n^2+(nu_n)^2\}} = \frac{1}{\pi (1+u_n^2)}. \quad (11.14)$$

Таким образом, элемент распределения z равен

$$dF = \frac{dz}{n(1+z^2)}, \quad -\infty \leq z \leq \infty, \quad (11.15)$$

т. е. z имеет то же распределение, что и отдельное наблюдение.

Здесь мы имеем интересный пример невыполнения центральной предельной теоремы: распределение среднего значения выборки из n наблюдений не стремится к нормальному распределению при возрастании n . Отметим, что исходное распределение не имеет моментов.

Пример 11.2

Найти распределение линейной функции от n независимых нормально распределенных случайных величин $x_1, \dots, x_n$; предполагается, что величина x_i имеет среднее нуль и дисперсию σ_i^2 .

Пусть линейная функция равна

$$z = a_1 x_1 + \dots + a_n x_n. \quad (11.16)$$

Переходя к переменным $u_j = x_j / \sigma_j$, получим

$$z = \sum a_j \sigma_j u_j. \quad (11.17)$$

Величины u_j имеют среднее нуль и единичную дисперсию. Наша задача, следовательно, сводится к нахождению распределения линейной функции от n независимых нормально распределенных случайных величин с нулевым средним и единичной дисперсией.

Рассмотрим преобразование вида

$$v_j = \sum_{i=1}^n c_{ij} u_i \quad (11.18)$$

с ортогональными коэффициентами

$$\sum_{i=1}^n c_{ij} c_{ik} = \begin{cases} 0, & j \neq k, \\ 1, & j = k. \end{cases} \quad (11.19)$$

Имеем, очевидно,

$$\sum_{i=1}^n v_i^2 = \sum (\sum c_{ij} u_j)^2 = \sum_{j=1}^n u_j^2.$$

Якобиан преобразования (11.18) равен

$$J = \left| \frac{\partial v}{\partial u} \right| = |c_{ij}|.$$

Если матрицу $\|c_{ij}\|$ умножить на транспонированную матрицу $\|c_{ji}\|$, то в силу (11.19) получится матрица, у которой на главной диагонали стоят единицы, а остальные элементы — нули. Детерминант такой матрицы равен единице. Следовательно, $J^2=1$ и $J=1$. Элемент совместного распределения величин u согласно предположению имеет вид

$$dF = \frac{1}{(2\pi)^{\frac{1}{2}n}} \exp\left(-\frac{1}{2} \sum_{j=1}^n u_j^2\right) \prod_j du_j. \quad (11.20)$$

По сказанному выше величины v распределены в точности так же:

$$dF = \frac{1}{(2\pi)^{\frac{1}{2}n}} \exp\left(-\frac{1}{2} \sum_{j=1}^n v_j^2\right) \prod_j dv_j. \quad (11.21)$$

Из (11.21) вытекает, что случайная величина v_1 независима от остальных величин v и что распределение ее имеет вид

$$dF = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} v_1^2\right) dv_1, \quad (11.22)$$

т. е. v_1 нормально распределена со средним нуль и единичной дисперсией. По симметрии то же самое верно для любой величины v_j . Величина v_1 с точностью до постоянной — любая ненулевая линейная функция от u_j , поскольку для всякой последовательности $c_{11}, \dots, c_{1n}$, удовлетворяющей условию $c_{11}^2 + \dots + c_{1n}^2 = 1$, можно подобрать числа c_{ij} , $i \geq 2$, $1 \leq j \leq n$, так, чтобы получилась ортогональная матрица. Для нахождения распределения величины z , заданной соотношением (11.17), положим $c_{1j} = a_j \sigma_j / (a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2)^{\frac{1}{2}}$. Тогда $z = (a_1^2 \sigma_1^2 + \dots + a_n^2 \sigma_n^2)^{\frac{1}{2}} v_1$. Следовательно, z имеет нормальное распределение с нулевым средним и дисперсией

$$Dz = \sum_j a_j^2 \sigma_j^2. \quad (11.23)$$

Пример 11.3. Преобразование Хельмерта

Преобразования типа, использованного в предыдущем примере, особенно полезны при выборках из нормальной генеральной совокупности, поскольку якобиан такого преобразования постоянен и функция плотности в этом случае сохраняет свою

Порядковые статистики

11.4 Не всегда можно найти такое преобразование, чтобы интересующая нас статистика стала одной из новых переменных и область интегрирования по остальным $n-1$ новым переменным не зависела бы от z . Особенно интересен случай так называемых «порядковых статистик», теория которых будет подробно изучаться в главе 14 и во втором томе. Пусть имеется выборка объема n . Перенумеруем выборочные значения в порядке возрастания их величины: $y_1, y_2, \dots, y_n$. В частности, наименьшим значением будет тогда y_1 . Полученное после перенумерации r -е значение y_r называется r -й порядковой статистикой. В этом смысле медиана и квантили являются порядковыми статистиками (медиана, например, равна $\frac{1}{2}(n+1)$ -му наблюдению, или среднему значению $\frac{1}{2}n$ -го и $\frac{1}{2}(n+2)$ -го наблюдений). Порядковыми статистиками являются также наименьшее ($r=1$) и наибольшее ($r=n$) значения в выборке.

Элемент совместного распределения компонент $x_1, x_2, \dots, x_n$ выборки объема n из генеральной совокупности, в которой распределение задается ф. р. $F(x)$, равен

$$dG = dF(x_1) dF(x_2) \dots dF(x_n). \quad (11.26)$$

Если к иксам применить преобразование

$$y = \int_{-\infty}^x dF(t) = F(x),$$

то распределение полученных игреков будет иметь вид *)

$$dH = dy_1 dy_2 \dots dy_n \quad (0 \leq y_i \leq 1 \text{ для всех } i). \quad (11.27)$$

Порядок расположения игреков и иксов один и тот же, поскольку y — неубывающая функция от x . Если игреки перенумеровать так, чтобы y_1 был наименьшим, y_2 — наименьшим из остальных и т. д., то область их изменения будет задаваться неравенствами

$$0 \leq y_1 \leq y_2 \leq \dots \leq y_n \leq 1. \quad (11.28)$$

Интеграл от (11.27) по этой области сводится к

$$\int_0^1 \dots \left[\int_0^{y_1} \left\{ \int_0^{y_3} \left(\int_0^{y_2} dy_1 \right) dy_2 \right\} dy_3 \right] \dots dy_n. \quad (11.29)$$

*) Здесь предполагается, конечно, что ф. р. $F(x)$ непрерывна. (Прим. перев.)

Последовательно интегрируя, находим, что интеграл (11.29) равен $1/n!$. Следовательно, элемент совместного распределения упорядоченных игреков имеет вид

$$dK = n! \prod_{i=1}^n dy_i. \quad (11.30)$$

Найдем распределение y_r . Чтобы несколько упростить вычисления, произведем еще одно преобразование:

$$z_i = \begin{cases} y_i, & i = 1, 2, \dots, r, \\ 1 - y_i, & i = r+1, r+2, \dots, n. \end{cases} \quad (11.31)$$

Модуль якобиана этого преобразования равен единице, элемент совместного распределения величин z равен

$$dL = n! \prod_{i=1}^n dz_i, \quad (11.32)$$

и само их распределение сосредоточено в области

$$0 \leq z_1 \leq z_2 \leq \dots \leq z_r \leq 1 - z_{r+1} \leq \dots \leq 1 - z_n \leq 1,$$

или, в иной записи,

$$0 \leq z_1 \leq z_2 \leq \dots \leq z_r, \quad 0 \leq z_n \leq z_{n-1} \leq \dots \leq z_{r+1} \leq 1 - z_r.$$

Чтобы найти распределение z_r , проинтегрируем по всем остальным z , разбив их на две группы:

$$\begin{aligned} dM &= dz_r n! \int_0^{z_r} \dots \int_0^{z_3} \int_0^{z_2} dz_1 dz_2 \dots \\ &\quad \dots dz_{r-1} \int_0^{1-z_r} \int_0^{z_{r+1}} \dots \int_0^{z_{n-1}} dz_n dz_{n-1} \dots dz_{r+1} = \\ &= dz_r n! \frac{z_r^{r-1}}{(r-1)!} \frac{(1-z_r)^{n-r}}{(n-r)!} = \frac{n!}{(r-1)!(n-r)!} z_r^{r-1} (1-z_r)^{n-r} dz_r, \\ &\quad 0 \leq z_r \leq 1. \end{aligned} \quad (11.33)$$

Таким образом, z_r имеет бета-распределение первого рода. Для нахождения распределения r -й порядковой статистики x_r достаточно теперь вспомнить, что $F(x_r) = z_r$. Элемент этого распределения равен

$$dN = \frac{n!}{(r-1)!(n-r)!} [F(x_r)]^{r-1} [1-F(x_r)]^{n-r} f(x_r) dx_r. \quad (11.34)$$

Это выражение можно было бы получить более прямым способом (ср. (10.25)), однако мы предпочли изложенное выше рассуждение с целью иллюстрации применяемых методов.

Пример 11.4

Если n нечетно, скажем $n=2m+1$, то $(m+1)$ -й элемент (перепорядоченной) выборки является медианой. Элемент распределения z_{m+1} согласно (11.33) равен

$$dF = \frac{(2m+1)!}{(m!)^2} z^m (1-z)^m dz.$$

Легко убедиться, что дисперсия z_{m+1} равна $1/\{4(2m+3)\} = 1/\{4(n+2)\}$.

Так как медиана x_{m+1} связана с z_{m+1} соотношением

$$z_{m+1} = \int_{-\infty}^{x_{m+1}} dF(x), \quad \text{то} \quad \frac{dz_{m+1}}{dx_{m+1}} = f(x_{m+1}).$$

Отсюда и из (10.14) следует, что для больших выборок

$$Dx_{m+1} = Dz_{m+1} \{f(x_{m+1})\}^{-2} = 1/\{4nf^2(x_{m+1})\},$$

что согласуется с (10.29).

Геометрический метод

11.5 Значительной части громоздких рассуждений обычно можно избежать, используя геометрическую интерпретацию области интегрирования. Значения $x_1, \dots, x_n$ любой выборки можно представлять себе как координаты точки n -мерного евклидова пространства. Произведение $dF(x_1) \dots dF(x_n)$ можно тогда рассматривать как элемент массы в точке $x_1, \dots, x_n$, а полную частоту значений статистики z между z_1 и z_2 — как интеграл от элемента массы (т. е. как массу) по области, лежащей между двумя геометрическими местами точек $z(x_1, \dots, x_n) = z_1$ и $z(x_1, \dots, x_n) = z_2$, которые обычно являются гиперповерхностями в евклидовом n -мерном пространстве, т. е. представляют собой $(n-1)$ -мерные пространства. Значение функции распределения величины z в точке z_1 будет полной массой, сосредоточенной между гиперповерхностями, соответствующими $z = -\infty$ и $z = z_1$; значение элемента распределения будет представлять собой элемент массы, заключенной между гиперповерхностями $z - \frac{1}{2}dz$ и $z + \frac{1}{2}dz$.

Пример 11.5

Рассмотрим еще раз задачу примера 11.2. В n -мерном u -пространстве плотность задана формулой

$$\frac{1}{(2\pi)^{\frac{1}{2}n}} \exp\left(-\frac{1}{2} \sum u^2\right).$$

Статистика $z (= \sum a_i x_i)$ определяет гиперплоскость

$$z = \sum a_j \sigma_j u_j. \quad (11.35)$$

Требуется найти полную массу между этой гиперплоскостью и гиперплоскостью, соответствующей $z = -\infty$, т. е. массу, находящуюся по одну («нижнюю») сторону от гиперплоскости (11.35).

Величина $\sum u^2$ представляет собой квадрат расстояния от точки $u_1, \dots, u_n$ до начала координат и, следовательно, не меняется при поворотах координатных осей. Повернем оси так, чтобы одна из них стала перпендикулярной гиперплоскости (11.35), пересекая ее в точке Q . Пусть $P = (u_1, \dots, u_n)$ — точка, соответствующая выборке, и O — начало координат. Тогда

$$\sum u^2 = OP^2 = OQ^2 + QP^2$$

и плотность в точке P принимает вид

$$\frac{1}{(2\pi)^{\frac{1}{2}n}} \exp\left(-\frac{1}{2} OQ^2\right) \exp\left(-\frac{1}{2} QP^2\right).$$

При перемещениях по гиперплоскости (11.35) OQ^2 остается неизменным, так что интеграл от $\exp\left(-\frac{1}{2} QP^2\right)$ есть постоянная, не зависящая от OQ . Следовательно, функция плотности в точке z имеет вид

$$f(z) = k \exp\left(-\frac{1}{2} OQ^2\right),$$

где k — постоянная.

Квадрат расстояния OQ от O до гиперплоскости равен

$$OQ^2 = \frac{z^2}{\sum a_j^2 \sigma_j^2},$$

откуда

$$f(z) = k \exp\left\{-\frac{1}{2} \frac{z^2}{\sum a_j^2 \sigma_j^2}\right\}.$$

т. е. z распределена нормально с дисперсией $\sum a_j^2 \sigma_j^2$ и нулевым средним.

Поучительно сравнить этот пример с примером 11.2. И там и здесь речь идет об одном и том же, только на разных языках.

Пример 11.6

Рассмотрим опять задачу из 11.2. Изящество геометрического подхода отчетливо выступает при сравнении данного там решения и приводимого ниже.

Как в предыдущем примере, функция плотности здесь равна $k \exp\left(-\frac{1}{2} OP^2\right)$. Нам нужно найти распределение статистики $z = OP^2$. Заметим, что плотность постоянна на поверхности $z = \text{const}$, т. е. на поверхности n -мерного шара. Элемент распределения z равен, следовательно, интегралу от этой постоянной плотности по области между поверхностями, отвечающими z и $z + dz$, т. е. пропорционален произведению $\exp\left(-\frac{1}{2} OP^2\right)$ на элемент приращения объема n -мерного шара. Поскольку объем n -мерного шара пропорционален n -й степени его радиуса OP , то

$$dF = k \exp\left(-\frac{1}{2} OP^2\right) \frac{d}{dz} OP^n dz \propto e^{-\frac{1}{2} z} z^{\frac{1}{2}(n-2)} dz.$$

Вычисляя константу, получаем

$$dF = \frac{1}{2^{\frac{1}{2}n} \Gamma\left(\frac{1}{2}n\right)} e^{-\frac{1}{2} z} z^{\frac{1}{2}(n-2)} dz,$$

что согласуется с (11.8).

Предположим теперь, что величины $x_1, \dots, x_n$ по-прежнему нормальны с нулевым средним и единичной дисперсией, но подчинены p однородным линейным условиям вида

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0.$$

Это значит, что величины x должны лежать на p гиперплоскостях в n -мерном пространстве. Первая из гиперплоскостей высекает в гиперсфере постоянной плотности гиперсферу на единицу меньшей размерности; на этой гиперсфере плотность, конечно, тоже постоянна; в высеченной гиперсфере вторая гиперплоскость в свою очередь высекает гиперсферу размерности еще на единицу меньшей и т. д.; сечение последней гиперсферы гиперплоскостью приведет к гиперсфере размерности на p меньше исходной. Следовательно, распределение z здесь имеет такую же

форму, как и выше, только с n , замененным на $n - p$:

$$dF = \frac{1}{2^{\frac{1}{2}(n-p)} \Gamma\left\{\frac{1}{2}(n-p)\right\}} e^{-\frac{1}{2} z^2} z^{\frac{1}{2}(n-p-2)} dz. \quad (11.36)$$

Пример 11.7. Выборочное распределение среднего и дисперсии выборок из нормальной генеральной совокупности

Среднее выборки обозначим $\bar{x}$. Дисперсия выборки s^2 равна

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

Для выборок из нормальной генеральной совокупности с нулевым средним и единичной дисперсией плотность в точке $x_1, \dots, \dots, x_n$ пропорциональна

$$\exp\left(-\frac{1}{2} \sum x^2\right) = \exp\left\{-\frac{1}{2}(ns^2 + n\bar{x}^2)\right\}. \quad (11.37)$$

Найдем выборочные распределения величин s и $\bar{x}$. Выражение (11.37) показывает, что функция плотности может быть просто выражена в терминах s и $\bar{x}$, и нам нужно теперь найти некоторое преобразование элемента объема $dx_1 \dots dx_n$. Задача нахождения распределений $\bar{x}$ и s уже рассматривалась в примере 11.3 с аналитической точки зрения. Теперь нас интересует геометрический подход.

Рассмотрим в n -мерном пространстве (O — начало координат) вектор OQ с направляющими косинусами $\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}$. Пусть PM обозначает перпендикуляр, опущенный из точки P , отвечающей выборке, на OQ . Длина OM равна

$$\frac{x_1}{\sqrt{n}} + \frac{x_2}{\sqrt{n}} + \dots + \frac{x_n}{\sqrt{n}} = \bar{x} \sqrt{n},$$

а длина OP равна $\sqrt{\sum x^2}$. Следовательно, PM имеет длину $(\sum x^2 - n\bar{x}^2)^{\frac{1}{2}} = s \sqrt{n}$.

Элемент объема в точке P можно рассматривать как произведение элементарного приращения OM , равного $d\bar{x} \sqrt{n}$, и элементарного объема в гиперплоскости, проходящей через M и перпендикулярной OM . В этой гиперплоскости поверхностями одинаковой плотности, аналогично тому как это было в предыдущем примере, являются гиперболы радиуса $s \sqrt{n}$ с центром в точке M . Следовательно, элемент объема равен произведению

$k d\bar{x} s^{n-2} ds$ и некоторых других элементов, нас не интересующих, поскольку они не зависят от $\bar{x}$ и s . Для элемента распределения dF величин $\bar{x}$ и s можем, стало быть, написать

$$dF \propto \exp \left\{ -\frac{1}{2} (ns^2 + n\bar{x}^2) \right\} s^{n-2} d\bar{x} ds, \quad (11.38)$$

откуда вытекает, что dF разлагается в произведение двух множителей

$$dF_1 \propto \exp \left(-\frac{1}{2} n\bar{x}^2 \right) d\bar{x}, \quad (11.39)$$

и

$$dF_2 \propto \exp \left(-\frac{1}{2} ns^2 \right) s^{n-2} ds. \quad (11.40)$$

Таким образом, мы получаем новое доказательство уже известного нам факта, а именно: для выборок из нормальной генеральной совокупности выборочные значения среднего и дисперсии независимы. Из (11.40) следует, что для распределения s^2

$$dF \propto \exp \left(-\frac{1}{2} ns^2 \right) s^{n-3} ds^2.$$

Вычисляя здесь константу, находим

$$dF = \frac{n^{\frac{1}{2}(n-1)}}{2^{\frac{1}{2}(n-1)} \Gamma \left\{ \frac{1}{2}(n-1) \right\}} \exp \left(-\frac{1}{2} ns^2 \right) (s^2)^{\frac{1}{2}(n-3)} d(s^2), \quad (11.41)$$

$$0 \leq s^2 \leq \infty,$$

что согласуется с (11.25).

Интересно сравнить распределение s^2 с распределением из предыдущего примера. В предыдущем примере было найдено распределение суммы квадратов отклонений от *некоторой фиксированной точки*. В рассматриваемом примере нами найдено распределение суммы квадратов отклонений от *выборочного среднего*, деленной на n . Сравнение формул (11.41) и (11.36) показывает, что с точностью до констант распределение дисперсии совпадает с распределением суммы квадратов при наличии одного линейного условия — факт, уже отмечавшийся нами в примере 11.3.

Пример 11.8. Распределение «Стьюдента»

В обозначениях предыдущего примера имеем

$$\frac{\bar{x}\sqrt{n}}{s\sqrt{n}} = \frac{OM}{PM} = \operatorname{ctg} \varphi,$$

где φ — угол POM .

Определим статистику z , положив $z = \bar{x}/s$. Очевидно, z не меняется на конусе, образованном вращением OP вокруг единичного вектора при постоянном угле φ . Распределение z будет, стало быть, найдено, если определить массу, заключенную между конусами, определяемыми углами φ и $\varphi + d\varphi$.

Рассмотрим пересечение этих конусов с гиперсферой радиуса OP . Они высекают на гиперсфере кольцо, площадь которого пропорциональна

$$OP d\varphi P M^{n-2} = OP^{n-1} \sin^{n-2} \varphi d\varphi.$$

Функция плотности на гиперсфере постоянна и пропорциональна $\exp\left(-\frac{1}{2} OP^2\right)$. Следовательно, масса, заключенная между конусами, пропорциональна

$$\int_0^\infty e^{-\frac{1}{2} OP^2} OP^{n-1} \sin^{n-2} \varphi d\varphi d(OP) \propto \sin^{n-2} \varphi d\varphi, \quad 0 \leq \varphi \leq \pi.$$

Отсюда для распределения $z (= \text{ctg } \varphi)$ получаем

$$dF \propto \frac{k dz}{(1+z^2)^{\frac{1}{2}n}}$$

и, вычисляя константу.

$$dF = \frac{1}{B\left\{\frac{1}{2}(n-1), \frac{1}{2}\right\}} \frac{dz}{(1+z^2)^{\frac{1}{2}n}}. \quad (11.42)$$

Поскольку z есть отношение двух функций одинаковой размерности, то его распределение одно и то же для выборок из нормальной генеральной совокупности независимо от шкалы. иными словами, независимо от дисперсии исходного распределения.

Обычно рассматривают не статистику z , а отличающуюся от нее множителем статистику

$$t = \frac{\bar{x} \sqrt{n}}{\sqrt{\frac{1}{n-1} \sum (x - \bar{x})^2}} = \sqrt{n-1} z.$$

Для распределения t из (11.42) получаем

$$dF = \frac{1}{\sqrt{n-1} B\left\{\frac{1}{2}(n-1), \frac{1}{2}\right\}} \frac{dt}{\left(1 + \frac{t^2}{n-1}\right)^{\frac{1}{2}n}} = \quad (11.43)$$

$$= \frac{\Gamma\left\{\frac{1}{2}(v+1)\right\}}{\sqrt{v\pi} \Gamma\left(\frac{1}{2}v\right)} \frac{dt}{(1+t^2/v)^{\frac{1}{2}(v+1)}}, \quad (11.44)$$

где $v = n - 1$.

Это знаменитое распределение известно как распределение «Стьюдента», названное так в честь открывшего его ученого, писавшего свои работы под этим псевдонимом (1908). Оно будет подробно обсуждаться в главе 16.

Пример 11.9. Распределение среднего выборки из равномерно распределенной совокупности

(Холл (P. Hall), 1927, *Biometrika* 19, 240)

Рассмотрим выборку объема n из совокупности с равномерным законом распределения

$$dF = dx, \quad 0 \leq x \leq 1.$$

Соответствующая функция плотности в n -мерном пространстве будет равна постоянной внутри гиперкуба

$$0 \leq x_j \leq 1, \quad j = 1, \dots, n, \quad (11.45)$$

и нулю в остальной части пространства. Единичный вектор направлен по диагонали этого куба. Если P — соответствующая выборке точка $(x_1, \dots, x_n)$ и PM — перпендикуляр на диагональ, то, как было показано в примере 11.7, $OM = \bar{x}\sqrt{n}$. Таким образом, чтобы найти распределение $\bar{x}$, нужно найти элемент массы между гиперплоскостями x и $x + d\bar{x}$. Последнее эквивалентно нахождению площади области, вырезаемой в гиперплоскости $\bar{x}$ различными сторонами гиперкуба. Здесь имеется некоторая сложность, обусловленная тем, что при изменении $\bar{x}$ эта область меняет свою форму в соответствии с числом ребер гиперкуба, пересекаемых гиперплоскостью.

Рассмотрим «квадранты»

$$\left. \begin{array}{l} x_j \geq r_j, \\ r_j = 0 \text{ или } 1, \end{array} \right\} \quad j = 1, 2, \dots, n, \quad (11.46)$$

углы которых совпадают с углами гиперкуба. Каждый угол имеет какое-нибудь число 0, 1, 2, ... или n координат, равных единице, а остальные его координаты равны нулю. Разобьем совокупность всех квадрантов на $(n+1)$ множеств в соответствии с количеством 0, 1, ... или n координат, равных единице, т. е. в соответствии со значениями 0, 1, ... или n сумм

$$r = \sum_{j=1}^n r_j.$$

Квадрант, принадлежащий t -му множеству, обозначим Q_t . Имеется $\binom{n}{t}$ различных квадрантов Q_t . Пусть S — любая точка из Q_0 , т. е. любая точка с неотрицательными координатами, и пусть s из ее координат ≥ 1 . Тогда S принадлежит $\binom{s}{0} (= 1)$ квадрантам Q_0 , $\binom{s}{1}$ квадрантам Q_1 , $\binom{s}{2}$ квадрантам Q_2 и т. д. Если $s > 0$, то

$$\sum_{t=0}^s (-1)^t \binom{s}{t} = (1 - 1)^s = 0. \quad (11.47)$$

Следовательно, если всякий раз, когда точка принадлежит Q_t , придать ей плотность $(-1)^t$, а затем просуммировать по всем Q , то в результате

получится 1 или 0 в соответствии с тем, принадлежит точка гиперкубу или нет *).

Пусть часть гиперплоскости

$$z = \sum x, \quad (11.48)$$

лежащая в Q_0 , имеет площадь $V_n(z)$. Тогда часть ее, лежащая в любом квадранте (11.46), будет иметь площадь $V_n(z-r)$, причем, очевидно, $V_n(z-r)=0$ при $r \geq z$, а площадь части ее, лежащей в гиперкубе, будет равна

$$\sum_{r=0}^k (-1)^r \binom{n}{r} V_n(z-r), \quad (11.49)$$

где $k=[z]$ — наибольшее целое число, не превосходящее z .

Найдем теперь $V_n(z)$. Пусть $\bar{V}_{n-1}(z)$ — площадь проекции части гиперплоскости (11.48), лежащей в Q_0 и спроектированной параллельно одной из осей. Имеем

$$V_n(z) = \sqrt{n} \bar{V}_{n-1}(z).$$

Величина $\bar{V}_n(z)$ является объемом n -мерной области, ограниченной координатными гиперплоскостями и гиперплоскостью (11.48), т. е. является объемом области, основание которой имеет площадь $V_n(z)$. Длина перпендикуляра, опущенного из точки O на это основание, равна $\frac{z}{\sqrt{n}}$. Поэтому

$$\bar{V}_n(z) = \frac{1}{n} \frac{z}{\sqrt{n}} V_n(z)$$

и

$$\bar{V}_{n-1}(z) = \frac{1}{n-1} \frac{z}{\sqrt{n-1}} V_{n-1}(z),$$

следовательно,

$$V_n(z) = \frac{1}{n-1} \sqrt{\frac{n}{n-1}} z V_{n-1}(z). \quad (11.50)$$

Пользуясь тем, что $V_2(z) = z\sqrt{2}$, и применяя последовательно эту формулу, находим

$$V_n(z) = \frac{\sqrt{n}}{(n-1)!} z^{n-1}.$$

Подставляя значения V_n в (11.49), получим следующее выражение для площади общей части гиперкуба и гиперплоскости, соответствующей значениям z , лежащим между k и $k+1$:

$$f(z) = \frac{\sqrt{n}}{(n-1)!} \sum_{r=0}^k (-1)^r \binom{n}{r} (z-r)^{n-1}. \quad (11.51)$$

Так как

$$\int_0^n f(z) \frac{dz}{\sqrt{n}} = 1,$$

*) Это верно с точностью до точек, лежащих на поверхности гиперкуба, которые, однако, можно не принимать во внимание. (Прим. перев.)

то плотность распределения среднего $\bar{x} = \frac{z}{n}$ имеет вид

$$g(\bar{x}) = \frac{n^n}{(n-1)!} \sum_{r=0}^k (-1)^r \binom{n}{r} \left(\bar{x} - \frac{r}{n}\right)^{n-1}, \quad \frac{k}{n} \leq \bar{x} \leq \frac{k+1}{n}. \quad (11.52)$$

Весьма любопытно, что график этой плотности состоит из n дуг парабол степени $n-1$ относительно $\bar{x}$, соединяющихся в $(n-1)$ точках $\frac{k}{n}$ ($k = 1, 2, \dots, \dots, n-1$).

Очевидно, найденное распределение симметрично относительно $\bar{x} = \frac{1}{2}$.

Для значений $n=2, 3, 4$ формула (11.52) даст следующие функции плотности:

$$\begin{aligned} n=2: \quad & 4\bar{x}, & 0 \leq \bar{x} \leq \frac{1}{2}, \\ & 4(1-\bar{x}), & \frac{1}{2} \leq \bar{x} \leq 1. \\ n=3: \quad & \frac{27\bar{x}^2}{2}, & 0 \leq \bar{x} \leq \frac{1}{3}, \\ & \frac{27}{2} \left\{ \bar{x}^2 - 3 \left(\bar{x} - \frac{1}{3} \right)^2 \right\}, & \frac{1}{3} \leq \bar{x} \leq \frac{2}{3}, \\ & \frac{27}{2} (1-\bar{x})^2, & \frac{2}{3} \leq \bar{x} \leq 1. \\ n=4: \quad & \frac{128}{3} \bar{x}^3, & 0 \leq \bar{x} \leq \frac{1}{4}, \\ & \frac{128}{3} \left\{ \bar{x}^3 - 4 \left(\bar{x} - \frac{1}{4} \right)^3 \right\}, & \frac{1}{4} \leq \bar{x} \leq \frac{1}{2}, \\ & \frac{128}{3} \left\{ (1-\bar{x})^3 - 4 \left(\frac{3}{4} - \bar{x} \right)^3 \right\}, & \frac{1}{2} \leq \bar{x} \leq \frac{3}{4}, \\ & \frac{128}{3} (1-\bar{x})^3, & \frac{3}{4} \leq \bar{x} \leq 1. \end{aligned}$$

При $n=2$ получается «треугольное» распределение. В силу центральной предельной теоремы распределения будут приближаться к нормальному при возрастании n .

Метод характеристических функций

11.6 В 7.18 уже отмечалось, что характеристическая функция суммы n независимых случайных величин есть произведение их характеристических функций. Это простое свойство позволяет найти выборочные распределения широкого класса статистик, выражаемых в виде сумм, в частности, средних значений.

Если имеется выборка объема n из генеральной совокупности с характеристической функцией $\varphi(t)$, то характеристической

функцией суммы выборочных значений будет φ^n . Для функции распределения $F(z)$ и функции плотности $f(z)$ этой суммы имеем согласно (4.4) и (4.33)

$$F(z) - F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-izt}}{it} \varphi^n dt \quad (11.53)$$

и

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itz} \varphi^n dt. \quad (11.54)$$

Следующие примеры иллюстрируют мощность этих простых результатов.

Пример 11.10. Распределение среднего выборки в случае биномиального распределения

Характеристическая функция биномиального распределения, задаваемого символическим выражением $(q+p)^n$, равна

$$(q + pe^{it})^n.$$

Х.ф. выборочного распределения суммы n значений имеет вид

$$(q + pe^{it})^{rn},$$

а х.ф. распределения среднего выборки (суммы, деленной на n) есть

$$(q + pe^{it/n})^{rn}.$$

Но последнее выражение является х.ф. биномиального распределения, задаваемого символическим выражением

$$(q+p)^{rn}, \quad (11.55)$$

но только с шагом $1/n$ вместо единицы; это распределение, следовательно, и является распределением среднего.

Пример 11.11. Распределение среднего выборки в случае пуассоновского распределения

Характеристическая функция пуассоновского распределения с общим членом $e^{-\lambda} \lambda^r / r!$ равна

$$\exp \{ \lambda (e^{it} - 1) \}.$$

Х.ф. среднего имеет вид

$$\exp \{ n \lambda (e^{it/n} - 1) \}.$$

Следовательно, распределение среднего является пуассоновским распределением с общим членом

$$e^{-n\lambda} \frac{(n\lambda)^r}{r!} \quad (11.56)$$

и с шагом $1/n$ вместо единицы.

Пример 11.12. Распределение среднего выборки из нормальной генеральной совокупности

Согласно примеру 3.4 характеристическая функция нормального распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right\} dx$$

равна

$$\exp \left(-\frac{1}{2} t^2 \sigma^2 + it\mu \right).$$

Х. ф. распределения среднего n значений имеет вид

$$\exp \left\{ n \left(-\frac{1}{2} \frac{t^2 \sigma^2}{n^2} + \frac{it\mu}{n} \right) \right\} = \exp \left\{ -\frac{1}{2} \frac{t^2 \sigma^2}{n} + it\mu \right\}. \quad (11.57)$$

Выражение (11.57) есть х. ф. нормального распределения со средним μ и дисперсией σ^2/n , которое и является искомым распределением.

Пример 11.13. Распределение среднего выборки из генеральной совокупности с распределением III типа

Характеристическая функция распределения III типа

$$dF = \frac{1}{\Gamma(\gamma)} e^{-x/a} \left(\frac{x}{a} \right)^{\gamma-1} \frac{dx}{a}, \quad a > 0,$$

равна (пример 3.6)

$$\varphi(t) = \frac{1}{(1 - ita)^\gamma}.$$

Х. ф. распределения среднего n значений имеет вид

$$\varphi(t) = \frac{1}{\left(1 - \frac{ita}{n} \right)^{\gamma n}}.$$

Это последнее выражение является х. ф. распределения

$$dF = \frac{1}{\Gamma(\gamma n)} e^{-nx/a} \left(\frac{nx}{a} \right)^{\gamma n - 1} \frac{n dx}{a}. \quad (11.58)$$

Пример 11.14. Распределение среднего выборки для случая равномерного распределения

Характеристическая функция равномерного распределения $dF = dx$ ($0 \leq x \leq 1$) равна

$$\int_0^1 e^{itx} dx = \frac{e^{it} - 1}{it}.$$

Х. ф. среднего n значений имеет вид $\left(\frac{e^{it/n}-1}{it/n}\right)^n$, так что функция плотности распределения этого среднего равна

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \left(\frac{e^{it/n}-1}{it/n}\right)^n dt. \quad (11.59)$$

Поскольку подынтегральная функция всюду аналитична, вместо интегрирования по действительной прямой можно интегрировать по контуру Γ , состоящему из полупрямой $-\infty < t \leq -c$, маленькой полуокружности радиуса c с центром в начале координат и полупрямой $c \leq t < \infty$. Имеем

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{\Gamma} e^{-itx} \left(\frac{e^{it/n}-1}{it/n}\right)^n dt = \\ &= \frac{(-1)^n}{2\pi} \int_{\Gamma} e^{-itx} \sum_{j=0}^n (-1)^j \binom{n}{j} \frac{e^{itj/n}}{(it/n)^n} dt. \end{aligned} \quad (11.60)$$

Далее,

$$\int_{\Gamma} \frac{e^{igz}}{z^n} dz = \begin{cases} 0 & \text{при } g > 0, \\ -2\pi i \frac{g^{n-1}}{(n-1)!} & \text{при } g \leq 0. \end{cases}$$

В этом можно убедиться, интегрируя по контуру, образованному контуром Γ и полуокружностью бесконечного радиуса, лежащей соответственно в верхней или нижней полуплоскости, при $g > 0$ или $g \leq 0$.

Пользуясь найденным значением интеграла, из (11.60) получаем

$$\begin{aligned} f(x) &= \frac{(-1)^n}{2\pi} \int_{\Gamma} \sum_{j=0}^n (-1)^j \binom{n}{j} \frac{e^{itj/n-itx}}{(it/n)^n} dt = \\ &= \frac{(-1)^{n-1}}{(n-1)!} \sum_{j \leq nx} (-1)^j \binom{n}{j} \frac{\left(\frac{j}{n} - x\right)^{n-1}}{\left(\frac{1}{n}\right)^n} = \\ &= \frac{n^n}{(n-1)!} \sum_{j=0}^{[nx]} (-1)^j \binom{n}{j} \left(x - \frac{j}{n}\right)^{n-1}. \end{aligned}$$

Эта формула лишь обозначениями отличается от (11.52).

11.7 Можно найти также общие выражения для распределений геометрических средних и моментов относительно фиксированных точек.

В самом деле, если $y = \log x$, то характеристическая функция y имеет вид

$$\alpha(t) = \int_{-\infty}^{\infty} e^{it \log x} dF = \int_{-\infty}^{\infty} x^{it} dF.$$

Обозначим nz сумму n независимых значений игроков. Для функции распределения nz имеем представление

$$F(nz) - F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-itnz}}{it} \alpha^n dt. \quad (11.61)$$

По распределению nz определяется распределение среднего значения z . Зная распределение z , с помощью формулы $z = \log u$ можно найти распределение среднего геометрического u .

Функция плотности распределения nz , если она существует, равна

$$f(nz) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itnz} \alpha^n dt.$$

Подобным образом можно найти распределение моментов выборки. Характеристическая функция степени x^r , где x имеет распределение F , выражается формулой

$$\beta(t) = \int_{-\infty}^{\infty} \exp(itx^r) dF,$$

и для распределения r -го момента z имеем

$$F(nz) - F(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1 - e^{-itnz}}{it} \beta^n dt. \quad (11.62)$$

Пример 11.15. Распределение геометрического среднего выборки в случае равномерного распределения

Если выбор производится из генеральной совокупности с распределением

$$dF = dx/a, \quad 0 \leq x \leq a,$$

то характеристическая функция величины $\log x$ имеет вид

$$\int_0^a x^{it} \frac{dx}{a} = \frac{a^{it}}{1+it},$$

и плотность распределения суммы $u = \sum \log x$ равна, следовательно,

$$f(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itu} a^{nit}}{(1+it)^n} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{it(n \log a - u)}}{(1+it)^n} dt, \quad n \log a - u \geq 0.$$

Вычисляя этот последний интеграл (его нелегко вычислить, переходя к другому контуру интегрирования), находим

$$f(u) = \frac{(n \log a - u)^{n-1} e^{-(n \log a - u)}}{\Gamma(n)}.$$

Отсюда, подставляя $z = e^{u/n}$, получаем следующее выражение для плотности геометрического среднего:

$$f(z) = \frac{n^n z^{n-1}}{a^n \Gamma(n)} \left(\log \frac{a}{z} \right)^{n-1}. \quad (11.63)$$

Пример 11.16. Распределение момента второго порядка относительно среднего генеральной совокупности для выборок из нормальной генеральной совокупности

Для распределения

$$dF = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2} x^2 / \sigma^2\right) dx$$

характеристическая функция величины x^2 равна

$$\frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(itx^2) \exp\left(-\frac{1}{2} x^2 / \sigma^2\right) dx = \frac{1}{(1 - 2\sigma^2 it)^{1/2}}.$$

Обозначим m_2 среднее n независимых значений x^2 . Х. ф. распределения величины m_2 имеет вид

$$\frac{1}{(1 - 2\sigma^2 it/n)^{\frac{1}{2}n}}, \quad (11.64)$$

и, следовательно, функция плотности этого распределения равна

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(-itm_2)}{(1 - 2\sigma^2 it/n)^{\frac{1}{2}n}} dt.$$

Последний интеграл можно вычислить, переходя к другому контуру интегрирования. Но его можно и прямо выписать, используя то обстоятельство, что (11.64) является характеристической

функцией распределения,

$$dF = \frac{n^{\frac{1}{2}n}}{(2\sigma^2)^{\frac{1}{2}n} \Gamma\left(\frac{1}{2}n\right)} \exp(-nm_2/2\sigma^2) m_2^{\frac{1}{2}n-1} dm_2. \quad (11.65)$$

Полезно сравнить это распределение с распределением (11.8). Последнее получается из него заменой $z = nm_2/\sigma^2$.

Метод индукции

11.8 Распределение суммы двух независимых случайных величин можно вычислить непосредственно следующим образом. Элемент их совместного распределения равен

$$dF(x_1, x_2) = dF_1(x_1) dF_2(x_2) = f_1(x_1) f_2(x_2) dx_1 dx_2.$$

Для преобразованных величин $z = x_1 + x_2$, $y = x_2$ (якобиан здесь равен 1) имеем

$$dG(z, y) = f_1(z - y) f_2(y) dz dy. \quad (11.66)$$

Интегрируя по y , находим, что для распределения z

$$dH(z) = \left\{ \int_{-\infty}^{\infty} f_1(z - y) f_2(y) dy \right\} dz,$$

так что функция плотности распределения z есть

$$h(z) = \int_{-\infty}^{\infty} f_1(z - y) f_2(y) dy. \quad (11.67)$$

Функция распределения z получается интегрированием $h(z)$ от $-\infty$ до z . Она равна

$$H(z) = \int_{-\infty}^{\infty} F_1(z - y) f_2(y) dy. \quad (11.68)$$

Отметим, что если $x_1 \geq 0$, то интегрировать по y следует от $-\infty$ до z , а если также и $x_2 \geq 0$, то интегрирование ведется от 0 до z .

Формулу (11.67) можно использовать для последовательного нахождения распределений сумм любого числа случайных величин, распределение каждой из которых известно. Если все слагаемые распределены одинаково, то общая формула для функции распределения суммы может быть подсказана формулами для распределения суммы двух или трех слагаемых. В правильности общей формулы можно затем убедиться с помощью индукции. Следующие примеры иллюстрируют этот метод.

Пример 11.17

Рассмотрим еще раз распределение из примера 11.1

$$dF = \frac{dx}{\pi(1+x^2)}, \quad -\infty \leq x \leq \infty.$$

Согласно (11.67) функция плотности для суммы двух независимых случайных величин, каждая из которых имеет такое распределение, равна

$$\int_{-\infty}^{\infty} \frac{1}{\pi^2} \left(\frac{1}{1+(z-x)^2} \right) \left(\frac{1}{1+x^2} \right) dx = \frac{2}{\pi(z^2+2^2)}.$$

Это подсказывает общую формулу

$$\frac{n}{\pi(z^2+n^2)}.$$

Если она правильна, то для распределения суммы $(n+1)$ величин будем иметь

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{n}{\pi^2} \left(\frac{1}{n^2+(z-x)^2} \right) \left(\frac{1}{1+x^2} \right) dx = \\ = \frac{n}{\pi} \left(\frac{n+1}{n} \right) / [z^2 + (n+1)^2] = \frac{n+1}{\pi[z^2 + (n+1)^2]}. \end{aligned}$$

Следовательно, поскольку формула верна для $n=1, 2$, она верна для всех n .

Пример 11.18

В примере 11.6 мы обнаружили, что распределение суммы квадратов n независимых нормально распределенных случайных величин имеет вид

$$dF = \frac{1}{2^{\frac{1}{2}n} \Gamma\left(\frac{1}{2}n\right)} e^{-\frac{1}{2}z} z^{\frac{1}{2}n-1} dz. \quad (11.69)$$

Представим себе, что на основе изучения формы распределения для малых n мы предположили, что формула (11.69) верна для всех n . Чтобы убедиться в справедливости сделанного предположения, допустим, что эта формула верна для n . Пусть x — нормально распределенная случайная величина со средним нуль и единичной дисперсией, не зависящая от n других (независимых) так же распределенных случайных величин, сумму квадратов

которых обозначим z . Для распределения величины $v = x^2$ имеем

$$dF = \frac{1}{2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right)} e^{-\frac{1}{2}v} v^{-\frac{1}{2}} dv, \quad 0 \leq v \leq \infty.$$

Поскольку обе величины z и v неотрицательны, то согласно (11.67) плотность распределения величины $u = z + v$ равна

$$\begin{aligned} \int_0^u \frac{1}{2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right)} e^{-\frac{1}{2}(u-y)} (u-y)^{-\frac{1}{2}} \frac{1}{2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}n\right)} e^{-\frac{1}{2}y} y^{\frac{1}{2}(n-2)} dy = \\ = \frac{e^{-\frac{1}{2}u}}{2^{\frac{1}{2}(n+1)} \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}n\right)} \int_0^u (u-y)^{-\frac{1}{2}} y^{\frac{1}{2}(n-2)} dy = \\ = \frac{B\left(\frac{1}{2}, \frac{1}{2}n\right)}{2^{\frac{1}{2}(n+1)} \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}n\right)} e^{-\frac{1}{2}u} u^{\frac{1}{2}(n-1)} = \frac{1}{2^{\frac{1}{2}(n+1)} \Gamma\left\{\frac{1}{2}(n+1)\right\}} e^{-\frac{1}{2}u} u^{\frac{1}{2}(n-1)} \end{aligned}$$

т. е. имеет тот же вид, что и в (11.69), если там n заменить на $n+1$. Тем самым формула (11.69) верна для всех n .

Распределение отношения случайных величин

11.9 В некоторых ситуациях нужно знать выборочное распределение отношения двух независимых случайных величин x_1 и x_2 . Их совместное распределение выражается формулой

$$dF(x_1, x_2) = dF_1(x_1) dF_2(x_2) = f_1(x_1) f_2(x_2) dx_1 dx_2. \quad (11.70)$$

Перейдем к новым величинам

$$u = x_1/x_2, \quad v = x_2.$$

Якобиан этого преобразования равен

$$J = v. \quad (11.71)$$

Как указывалось в 11.3, замена переменных должна удовлетворять некоторым условиям. По существу, нужно, чтобы внутри области интегрирования якобиан не обращался в нуль; нулевые значения якобиана на границе допускаются. Из (11.71) вытекает, следовательно, что, за исключением границы, должно быть

$$x_2 = v \neq 0. \quad (11.72)$$

Таким образом, в первую очередь мы должны ограничиться рассмотрением тех отношений случайных величин, у которых *знаменатель* может принимать значения либо только от 0 до ∞ , либо только от $-\infty$ до 0. В первом случае согласно (11.70) и (11.71),

$$dG(u, v) = f_1(uv)f_2(v)v du dv. \quad (11.73)$$

Следовательно, распределение u имеет вид

$$dH(u) = \left\{ \int_{-\infty}^{\infty} f_1(uv)f_2(v)v dv \right\} du, \quad (11.74)$$

и функция распределения u равна

$$H(u) = \int_{-\infty}^{\infty} F_1(uv)f_2(v)dv. \quad (11.75)$$

Пример 11.19

Рассмотрим опять распределение отношения $\bar{x}/s$, о котором шла речь в примере 11.8. Здесь $\bar{x}$ и s — соответственно среднее и дисперсия выборки объема n из нормальной совокупности. Как мы уже видели ((11.39) и (11.40); в этих формулах $\sigma^2=1$), для распределений $\bar{x}$ и s соответственно

$$\begin{aligned} dF_1 &\propto \exp(-n\bar{x}^2/2\sigma^2) d\bar{x}, \\ dF_2 &\propto \exp(-ns^2/2\sigma^2) s^{n-2} ds. \end{aligned}$$

Из (11.74) следует, что плотность распределения $z(=\bar{x}/s)$ с точностью до константы равна

$$\int_0^{\infty} \exp(-nz^2s^2/2\sigma^2) \exp(-ns^2/2\sigma^2) s^{n-2} ds \propto \frac{1}{(1+z^2)^{\frac{1}{2}n}}.$$

Вычисляя константу, получаем результат, совпадающий с (11.42).

11.10 Если величина x_2 , введенная в 11.19, неотрицательна и имеет конечное среднее, то распределение отношения u можно также найти с помощью характеристических функций. При неотрицательном v уравнение (11.75) переходит в

$$H(u) = \int_0^{\infty} F_1(uv) dF_2(v).$$

Далее, обозначая $\varphi_1(t)$, $\varphi_2(t)$ соответственно х. ф. величин x_1 , x_2 и используя (4.4), получаем

$$\begin{aligned} H(u) &= \int_0^\infty \left\{ F_1(0) + \frac{1}{2\pi} \int_{-\infty}^\infty \frac{1 - e^{-ituv}}{it} \varphi_1(t) dt \right\} dF_2(v) = \\ &= F_1(0) + \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{\varphi_1(t)}{t} \left\{ \int_0^\infty (1 - e^{-ituv}) dF_2(v) \right\} dt = \\ &= F_1(0) + \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{\varphi_1(t)}{t} \{1 - \varphi_2(-tu)\} dt. \quad (11.76) \end{aligned}$$

Дифференцируя по u , находим, что ф. п. величины u при условии, что она существует, равна

$$h(u) = \frac{1}{2\pi i} \int_{-\infty}^\infty \varphi_1(t) \varphi_2'(-tu) dt, \quad (11.77)$$

если только интеграл сходится. Этот результат принадлежит Крамёру (1937). Даниэлсом (1954) было отмечено, что конечность среднего величины x_2 достаточна для абсолютной сходимости интеграла в правой части формулы (11.77) при всех u . Гири (Geary, 1944, *J. Roy. Statist. Soc.* **107**, 56) обобщил (11.77) на случай зависимых x_1 , x_2 с х. ф. совместного распределения $\varphi(t_1, t_2)$, предполагая неотрицательность величины x_2 и конечность ее среднего. Им получена формула

$$h(u) = \frac{1}{2\pi i} \int_{-\infty}^\infty \left[\frac{\partial \varphi(t_1, t_2)}{\partial t_2} \right]_{t_2 = -t_1 u} dt_1. \quad (11.78)$$

Набросок доказательства (11.78) приводится ниже в упражнении 11.24, из которого видна важность условия конечности $M(x_2)$

Пример 11.20. Распределение дисперсионного отношения Фишера

Пусть имеются две независимые выборки объема n_1 и n_2 из нормальных генеральных совокупностей с дисперсиями σ_1^2 и σ_2^2 соответственно. Определяя дисперсию выборки как $\sum (x - \bar{x})^2/n$,

обозначим дисперсии рассматриваемых выборок s_1^2 и s_2^2 . Элементы распределений этих величин согласно (11.40) выражаются формулами

$$dF_1 \propto \exp\left(-\frac{n_1 s_1^2}{2\sigma_1^2}\right) s_1^{n_1-3} ds_1^2,$$

$$dF_2 \propto \exp\left(-\frac{n_2 s_2^2}{2\sigma_2^2}\right) s_2^{n_2-3} ds_2^2.$$

Из (11.74) для функции плотности отношения $t^2 = s_1^2/s_2^2$ получаем

$$\begin{aligned} f(t^2) &\propto \int_0^\infty \exp\left(-\frac{n_1 t^2 s_2^2}{2\sigma_1^2}\right) (s_2 t)^{n_1-3} \exp\left(-\frac{n_2 s_2^2}{2\sigma_2^2}\right) s_2^{n_2-3} ds_2^2 \propto \\ &\propto \int_0^\infty \exp\left\{-\frac{s_2^2}{2} \left(\frac{n_1 t^2}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)\right\} s_2^{n_1+n_2-4} t^{n_1-3} ds_2^2 \propto \\ &\propto \frac{t^{n_1-3}}{\left(\frac{n_1 t^2}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}\right)^{\frac{1}{2}(n_1+n_2-2)}}, \quad 0 \leq t^2 \leq \infty. \end{aligned} \quad (11.79)$$

Это — функция плотности бета-распределения второго рода (ср. (6.17)). Иногда вместо t^2 рассматривают величину

$$z = \frac{1}{2} \log \frac{n_1 (n_2 - 1) s_1^2 / \sigma_1^2}{n_2 (n_1 - 1) s_2^2 / \sigma_2^2} = \frac{1}{2} \log \frac{n_1 (n_2 - 1)}{n_2 (n_1 - 1)} t^2 / \left(\frac{\sigma_1^2}{\sigma_2^2}\right). \quad (11.80)$$

Используя (11.79), легко найти функцию плотности распределения z . Имеем

$$f(z) \propto \frac{e^{(n_1-1)z}}{\{(n_1-1)e^{2z} + (n_2-1)\}^{\frac{1}{2}(n_1+n_2-2)}}, \quad -\infty \leq z \leq \infty, \quad (11.81)$$

и, полагая $v_1 = n_1 - 1$, $v_2 = n_2 - 1$ и вычисляя константу, получаем

$$f(z) = \frac{2v_1^{\frac{1}{2}} v_2^{\frac{1}{2}}}{B\left(\frac{1}{2} v_1, \frac{1}{2} v_2\right)} \frac{e^{v_1 z}}{(v_1 e^{2z} + v_2)^{\frac{1}{2}(v_1+v_2)}}. \quad (11.82)$$

Р. А. Фишер предложил z вместо t^2 по той причине, что интерполировать таблицы для z легче, чем для t . На распределении статистики z основывается большое количество статистических

тестов; позже (в главе 16) мы изучим его свойства подробнее *).

Найдем распределение t^2 с помощью теоремы Крамера о х. ф. отношения. Х. ф. величин s_1^2 и s_2^2 равны соответственно (ср. примеры 11.3 и 11.13)

$$\varphi_1(t) = \left(1 - \frac{2\sigma_1^2 it_1}{n_1}\right)^{-\frac{1}{2}(n_1-1)},$$

$$\varphi_2(t) = \left(1 - \frac{2\sigma_2^2 it_2}{n_2}\right)^{-\frac{1}{2}(n_2-1)}.$$

Отсюда, применяя формулу (11.77), для ф. п. отношения $v = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$ получаем

$$f(v) \propto \frac{1}{2\pi i} \int_{-\infty}^{\infty} \left(1 - \frac{2it}{n_1}\right)^{-\frac{1}{2}(n_1-1)} \left(1 + \frac{2itv}{n_2}\right)^{-\frac{1}{2}(n_2+1)} dt.$$

Интегрируя теперь по контуру, состоящему из мнимой оси и лежащей слева от нее бесконечной полуокружности, находим

$$f(v) \propto \left[\frac{d^\alpha}{dt^\alpha} \left(1 - \frac{2tv}{n_2}\right)^{-\frac{1}{2}(n_2+1)} \right]_{t=-\frac{1}{2}n_1},$$

где $\alpha = \frac{1}{2}(n_1-3)$, и, следовательно,

$$f(v) \propto \frac{v^{\frac{1}{2}(n_1-3)}}{\left(1 + \frac{n_1 v}{n_2}\right)^{\frac{1}{2}(n_1+n_2-2)}}.$$

Полученное выражение подстановкой $v = \frac{t^2}{\sigma_1^2/\sigma_2^2}$ приводится к виду (11.79). Следует отметить, что это доказательство применимо только в случае нечетных n_1 , поскольку α должно быть целым.

11.11 В 11.9 мы указывали на трудность при отыскании распределения отношения, возникающую, когда область значений знаменателя содержит внутри себя нуль. Один из способов обойти

*) На самом деле предложение Фишера большого распространения не получило. К настоящему времени созданы подробные таблицы распределения t^2 , интерполяция и экстраполяция которых осуществляются достаточно просто; см., например, «Таблицы математической статистики». (Прим. ред.)

эту трудность состоит в разбиении области значений знаменателя на две части: одну — состоящую из отрицательных чисел, другую — из неотрицательных. Мы будем считать, что функция распределения знаменателя непрерывна в нуле, поскольку в противном случае отношение, с положительной вероятностью, бесконечно. Из (11.70) находим (отдельно для каждой из двух частей области значений знаменателя), что

$$dG(u, v) = \begin{cases} f_1(uv) f_2(v) |v| du dv, & v < 0, \\ f_1(uv) f_2(v) v du dv, & v \geq 0, \end{cases} \quad (11.83)$$

откуда, аналогично (11.74), получаем

$$dH(u) = \left\{ \int_0^{\infty} f_1(uv) f_2(v) v dv - \int_{-\infty}^0 f_1(uv) f_2(v) v dv \right\} du. \quad (11.84)$$

Как и выше, из этого выражения легко получить формулу для функции распределения.

Практически обычно интересуются лишь теми отношениями, знаменатели которых неотрицательны. В этом случае, как в примерах 11.19 и 11.20, применим метод из 11.9. Но даже когда знаменатель может принимать значения любого знака, задачу иногда можно упростить, рассматривая квадрат интересующего отношения и применяя к нему метод из 11.9.

Пример 11.21

Рассмотрим отношение двух независимых нормально распределенных случайных величин с нулевым средним и единичной дисперсией. Для элемента распределения этого отношения из (11.84) получаем следующее выражение:

$$dH(u) = \frac{du}{2\pi} \left\{ \int_0^{\infty} e^{-\frac{1}{2}(uv)^2} e^{-\frac{1}{2}v^2} v dv - \int_{-\infty}^0 e^{-\frac{1}{2}(uv)^2} e^{-\frac{1}{2}v^2} v dv \right\},$$

которое в силу симметрии нормального распределения сводится к

$$dH(u) = \frac{du}{\pi} \int_0^{\infty} \exp \left\{ -\frac{1}{2} v^2 (u^2 + 1) \right\} v dv.$$

Интеграл справа, как легко видеть, равен $1/(1+u^2)$ (ср. со средним отклонением нормального распределения (5.75)). Следовательно, рассматриваемое отношение имеет распределение Коши

$$dH(u) = \frac{du}{\pi(1+u^2)}, \quad -\infty \leq u \leq \infty. \quad (11.85)$$

В данном случае эффективен предложенный в 11.11 переход к рассмотрению квадрата отношения, поскольку он ведет к отношению квадратов двух нормированных нормально распределенных случайных величин, каждый из которых имеет распределение III типа (11.8) с $n=1$. Распределение квадрата отношения совпадает, стало быть, с фишеровским распределением величины t^2 из примера 11.20. Легко убедиться в том, что (11.79) сводится к полученному нами выше результату, если положить $n_1=n_2=2^*$, $\sigma_1=\sigma_2=1$.

11.12 До сих пор мы имели дело главным образом с распределением одной статистики, построенной по элементам простой случайной выборки. Однако рассмотренные методы без особого труда обобщаются на случай задачи нахождения совместного распределения нескольких статистик. Например, если имеется несколько статистик $z_1, z_2, \dots, z_p$ и $dF(x_1, \dots, x_n)$ задает совместное распределение выборочных значений $x_1, \dots, x_n$, то х. ф. совместного распределения величин z имеет вид

$$\Phi(t_1, \dots, t_p) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp(it_1 z_1 + \dots + it_p z_p) dF(x_1, \dots, x_n), \quad (11.86)$$

а функция плотности (если она существует) выражается формулой

$$f(z_1, \dots, z_p) = \frac{1}{(2\pi)^p} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp(-it_1 z_1 - \dots - it_p z_p) \times \\ \times \Phi(t_1, \dots, t_p) dt_1 \dots dt_p. \quad (11.87)$$

Примеры использования этих результатов встретятся в дальнейшем.

Метод перевала

11.13 Хотя на протяжении этой главы мы занимались выводом точных результатов о выборочных распределениях, нам удобно здесь отметить один развитый Даниэлсом (1954) аппроксимативный метод, совершенно отличный от тех методов, которые будут обсуждаться в главе 12. Существенные черты его состоят в следующем. Функция плотности среднего представляется как обращение n -й степени характеристической функции. Полученный

*) Не 1, поскольку в данном случае берутся квадраты отклонений не относительно среднего выборки, а относительно среднего генеральной совокупности, и поэтому получается лишняя степень свободы. См. по этому поводу пример 11.7.

интеграл вычисляется приближенно, для чего выбирается путь интегрирования, проходящий через седловую точку подынтегральной функции так, чтобы подынтегральная функция была пренебрежимо мала вне малой окрестности этой точки. Это действительно можно сделать, и достигаемая аппроксимация, по крайней мере в простых случаях, очень хороша. В результате получаются разложения функции плотности эджвортовского типа (6.18), которые, однако, лучше, чем само разложение Эджворта.

УПРАЖНЕНИЯ

11.1 Найти распределение среднего выборки, образованной n независимыми наблюдениями из генеральной совокупности с распределением

$$dF = e^{-x} dx, \quad 0 \leq x \leq \infty,$$

(а) посредством индукции, (б) с помощью характеристических функций.

11.2 Пользуясь методом характеристических функций, найти выражение для выборочного распределения среднего выборки из генеральной совокупности с распределением

$$dF = \frac{dx}{\pi(1+x^2)}, \quad -\infty \leq x \leq \infty.$$

11.3 Показать, что распределение геометрического среднего g выборки объема n из генеральной совокупности с распределением III типа

$$dF = \frac{e^{-x} x^{p-1}}{\Gamma(p)} dx, \quad 0 \leq x \leq \infty,$$

имеет вид

$$dF = \frac{ng^{np-1}}{\Gamma(n) \{\Gamma(p)\}^n} \sum_{j=0}^{\infty} (-1)^{n+j+1} \left[\frac{d^{n-1}}{dt^{n-1}} \frac{g^{nt}}{\{\Gamma(t+1)\}^n} \right]_{t=j}. \quad (\text{Каллбэк, 1934.})$$

11.4 Пусть $x_1, \dots, x_n$ представляют собой n независимых нормально распределенных случайных величин с нулевым средним и единичной дисперсией.

Показать, что распределения статистик $x_i^2 / \left(\sum_{j=1}^n x_j^2 \right)$ и $\sum_{j=1}^n x_j^2$ независимы.

Установить также независимость величин $(x_i - \bar{x})^2 / \sum (x_j - \bar{x})^2$ и $\sum (x_j - \bar{x})^2$, где $\bar{x}$ — арифметическое среднее n значений.

11.5 Используя результаты предыдущего упражнения, показать, что если m_2, m_3 и m_4 — моменты порядков 2, 3 и 4 относительно среднего выборки, то

$$M \left(\frac{m_3}{m_2^{3/2}} \right)^k = \frac{M m_3^k}{M m_2^{3k/2}},$$

$$M \left(\frac{m_4}{m_2^2} \right)^k = \frac{M m_4^k}{M m_2^{2k}}.$$

11.6 Пусть $x_1, \dots, x_n$ представляют собой n независимых случайных величин, каждая из которых равномерно распределена между 0 и 1.

Показать, что если

$$Q = \prod_{j=1}^n x_j,$$

то

$$P \left\{ Q \leq e^{-\frac{1}{2}t} \right\} = \int_t^{\infty} z^{n-1} e^{-\frac{1}{2}z} dz / \{2^n (n-1)!\}.$$

11.7 Показать, что совместное распределение двух порядковых статистик x_r и x_s ($r < s$) имеет вид

$$dF(x_r, x_s) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} F_1^{r-1} (F_2 - F_1)^{s-r-1} (1 - F_2)^{n-s} dF_1 dF_2,$$

где F_1 и F_2 — значения функции распределения генеральной совокупности в точках x_r и x_s соответственно.

11.8 Интегрируя формулу предыдущего упражнения по x_s , найти распределение x_r .

11.9 Показать, что распределение геометрического среднего n независимых величин, имеющих распределения с плотностями соответственно

$$\frac{x^{p-1} e^{-x}}{\Gamma(p)}, \quad \frac{x^{p+1/n-1} e^{-x}}{\Gamma(p+1/n)}, \quad \dots, \quad \frac{x^{p+(n-1)/n-1} e^{-x}}{\Gamma(p+(n-1)/n)},$$

совпадает с распределением среднего арифметического n независимых случайных величин, распределение каждой из которых совпадает с распределением первой из предыдущих величин.

(Каллбэк, 1934.)

11.10 Пусть S — сумма n независимых случайных величин, распределенных с функцией плотности, равной сумме некоторого числа первых членов ряда типа А:

$$f(x) = a(x) \left\{ 1 + \frac{a_3}{\sigma^3} H_3 \left(\frac{x}{\sigma} \right) + \dots + \frac{a_k}{\sigma^k} H_k \left(\frac{x}{\sigma} \right) \right\}.$$

Показать, что сумма S имеет функцию плотности

$$f(S) = a(S) \left\{ 1 + \frac{A_3}{\Sigma^3} H_3 \left(\frac{S}{\Sigma} \right) + \dots + \frac{A_k}{\Sigma^k} H_k \left(\frac{S}{\Sigma} \right) \right\},$$

где $\Sigma = \sigma \sqrt{n}$ и

$$A_k = \sum \frac{n!}{v_3! v_4! \dots v_k! (N - v_3 - \dots - v_k)!} a_3^{v_3} \dots a_k^{v_k};$$

суммирование в последней формуле распространяется на те v , для которых

$$3v_3 + 4v_4 + \dots + kv_k = j.$$

(Бэйкер, 1930.)

11.11 Показать, что отношение двух независимых нормально распределенных случайных величин имеет функцию плотности, равную

$$f(v) = \frac{1}{\sqrt{2\pi}} \frac{m_2 \sigma_1^2 + m_1 \sigma_2^2 v}{(\sigma_1^2 + \sigma_2^2 v^2)^{3/2}} \exp \left\{ -\frac{1}{2} \frac{(m_1 - m_2 v)^2}{\sigma_1^2 + \sigma_2^2 v^2} \right\},$$

где m_1 , σ_1 и m_2 , σ_2 — среднее и дисперсия соответственно первой и второй величин. Предполагается, что m_2 столь велико по сравнению с σ_2 , что область значений второй величины можно считать расположенной вправо от нуля.

Вывести отсюда, что величина $\frac{m_1 - m_2 v}{(\sigma_1^2 + \sigma_2^2 v^2)^{1/2}}$ распределена нормально с нулевым средним и единичной дисперсией.

(Гири, 1930.)

11.12 Показать, что отношение двух независимых случайных величин с распределениями соответственно

$$dF \propto e^{-\gamma_1(x-m_1)}(x-m_1)^{p_1-1} dx, \quad 0 < m_1 \leq x < \infty,$$

$$dF \propto e^{-\gamma_2(x-m_2)}(x-m_2)^{p_2-1} dx, \quad 0 < m_2 \leq x < \infty,$$

имеет функцию плотности

$$f(v) = \frac{\gamma_1^{p_1} e^{\xi \gamma_1} m_2}{(p_1 - 1)!} \left\{ \frac{\xi^{p_1-1}}{\left(1 + \frac{\gamma_1 v}{\gamma_2}\right)^{p_2}} - \binom{p_1-1}{1} \left(\frac{v}{\gamma_2}\right) \frac{p_2 \xi^{p_1-2}}{\left(1 + \frac{\gamma_1 v}{\gamma_2}\right)^{p_2+1}} + \right. \\ \left. + \binom{p_1-2}{2} \left(\frac{v^2}{\gamma_2^2}\right) \frac{p_2(p_2+1) \xi^{p_1-3}}{\left(1 + \frac{\gamma_1 v}{\gamma_2}\right)^{p_2+2}} - \dots \right\} + \\ + \frac{\gamma_1^{p_1} e^{\xi \gamma_1} p_2}{(p_1 - 1)! \gamma_2} \left\{ \frac{\xi^{p_1-1}}{\left(1 + \frac{\gamma_1 v}{\gamma_2}\right)^{p_2+1}} - \binom{p_1-1}{1} \left(\frac{v}{\gamma_2}\right) \frac{(p_2+1) \xi^{p_1-2}}{\left(1 + \frac{\gamma_1 v}{\gamma_2}\right)^{p_2+2}} + \dots \right\},$$

где $\xi = m_1 - m_2 v$. (Распределение этого отношения включает в себя фишеровское распределение дисперсионного отношения как частный случай.)

11.13 Пусть x_1 и x_2 — независимые случайные величины и x_1 распределена нормально со средним m_1 и дисперсией σ^2 , а x_2 имеет распределение стандартного отклонения в нормальной выборке, но с измененным началом отсчета, т. е.

$$dF \propto e^{-\gamma(s-m_2)}(s-m_2)^{p-1} ds, \quad 0 \leq m_2 \leq s < \infty.$$

Показать, что функция плотности отношения $v = x_1/x_2$ выражается формулой

$$f(v) = \frac{\gamma^p e^{-\xi^2/2\sigma^2}}{\sigma \sqrt{2\pi} \Gamma(p)} \left\{ m_2 \sum_{j=0}^{\infty} \frac{v^j \xi^j \Gamma\left(p + \frac{j}{2}\right)}{j! \sigma^{2j} \left(\gamma + \frac{1}{2} \frac{v^2}{\sigma^2}\right)^{p + \frac{j}{2}}} + \right. \\ \left. + \sum_{j=0}^{\infty} \frac{v^j \xi^j \Gamma\left(p + \frac{j+1}{2}\right)}{j! \sigma^{2j} \left(\gamma + \frac{1}{2} \frac{v^2}{\sigma^2}\right)^{p + \frac{j+1}{2}}} \right\},$$

где $\xi = m_1 - m_2 v$.

(Распределение Стьюдента является частным случаем распределения этого отношения.)

11.14 Показать, что распределение среднего выборки объема n из генеральной совокупности с распределением II типа

$$dF \propto x^{p-1} (1-x)^{p-1} dx, \quad p > 0, \quad 0 \leq x \leq 1,$$

имеет плотность

$$f(\bar{x}) = \frac{1}{2} n \pi^{\frac{1}{2} n - 1} \{\Gamma(p)\}^n \int_{-\infty}^{\infty} \left\{ \mathcal{J}_{p-\frac{1}{2}} \left(\frac{1}{2} t \right) \right\}^n \cos(n \bar{x} t) dt / t^{n(p-\frac{1}{2})},$$

где $\mathcal{J}_r(z)$ — функция Бесселя порядка r по z .

(Ирвин, 1927.)

11.15 Показать, что функция плотности разности z двух независимых случайных величин, каждая из которых имеет распределение III типа

$$dF = \frac{e^{-x} x^{p-1}}{\Gamma(p)} dx, \quad p > 0, \quad 0 \leq x \leq \infty,$$

равна

$$f(z) = \frac{z^{p-\frac{1}{2}} K_{p-\frac{1}{2}}(z)}{2^{p-\frac{1}{2}} \Gamma(p) \Gamma\left(\frac{1}{2}\right)},$$

где $K_r(x)$ — функция Бесселя второго порядка от мнимого аргумента.

(К. Пирсон и др., 1932.)

11.16 Показать, что разность двух значений, каждое из которых выбрано случайно и независимо из пуассоновской генеральной совокупности, распределение которой имеет общий член вида $e^{-\lambda} \lambda^r / r!$, подчиняется дискретному распределению с общим членом вида $e^{-2\lambda} T_d(2\lambda)$, где d может принимать все целые значения от $-\infty$ до ∞ и $T_d(2\lambda)$ — модифицированная функция Бесселя первого рода порядка d от аргумента 2λ (ср. с примером 4.5).

(Ирвин, 1937.)

11.17 Используя результат упражнения 11.15, найти распределение среднего n независимых величин, каждая из которых имеет распределение

$$dF = \frac{1}{2} e^{-|x|} dx, \quad -\infty \leq x \leq \infty,$$

11.18 Найти характеристическую функцию совместного распределения среднего и дисперсии выборки объема n из нормальной генеральной совокупности и показать, что эта х. ф. распадается в произведение х. ф. среднего и х. ф. дисперсии, так что среднее и дисперсия независимы.

11.19 Пусть $\varphi(t_1, t_2)$ — х. ф. совместного распределения среднего и дисперсии выборки объема n из генеральной совокупности, распределение которой

имеет плотность $f(x)$ и х. ф. $\alpha(t)$. Показать, что

$$\left[\frac{\partial}{\partial t_2} \varphi(t_1, t_2) \right]_{t_2=0} = \frac{n-1}{n} i \left[\alpha^{n-1}(t_1/n) \int_{-\infty}^{\infty} x^2 e^{it_1 x/n} f(x) dx - \right. \\ \left. - \alpha^{n-2}(t_1/n) \left\{ \int_{-\infty}^{\infty} x e^{it_1 x/n} f(x) dx \right\}^2 \right].$$

Показать далее, что если среднее и дисперсия независимы, то

$$-\alpha \frac{d^2 \alpha}{dt^2} + \left(\frac{d\alpha}{dt} \right)^2 = \sigma^2 \alpha^2,$$

где σ^2 — дисперсия распределения генеральной совокупности, и что, следовательно, в этом случае генеральная совокупность необходимо нормальна.

[Лукач (1942). Гири (1936) установил этот результат для распределений с конечными семинвариантами всех порядков.]

11.20 Из генеральной совокупности с равномерным распределением $dF = dx/\theta$, $0 \leq x \leq \theta$, независимо извлекаются две выборки объемов m и n каждая. Наибольший элемент r -й выборки обозначим x_r и положим x_m — наибольшие элементы этих выборок. Показать, что для распределения x_m

$$dF(x_m) = \{m x_m^{m-1} / \theta^m\} dx_m$$

и что, следовательно, для распределения отношения x_m/x_n

$$dG(u) = \begin{cases} \frac{mn}{m+n} u^{m-1} du, & 0 \leq u \leq 1, \\ \frac{mn}{m+n} \frac{du}{u^{n+1}}, & 1 \leq u \leq \infty. \end{cases}$$

(Мэрти, 1955.)

11.21 Из генеральной совокупности с равномерным распределением $dF = dx$, $0 \leq x \leq 1$, независимо извлекаются k случайных выборок объема n каждая. Наибольший элемент r -й выборки обозначим x_r и положим $v = x_1 x_2 \dots x_k$. Пользуясь индукцией или каким-либо другим методом, показать, что для распределения v

$$dG(v) = \{n^k / (k-1)!\} v^{n-1} \{\log(1/v)\}^{k-1} dv.$$

(Райдер, 1955.)

11.22 Обобщая результат примера 11.21, показать, что если величины x и y имеют двумерное нормальное распределение с коэффициентом корреляции ρ , то статистика

$$z = \frac{x - \mu_x}{\sigma_x} / \frac{y - \mu_y}{\sigma_y}$$

имеет функцию плотности

$$g(z) = \frac{(1 - \rho^2)^{1/2} dz}{\pi (1 - 2\rho z + z^2)}, \quad -\infty \leq z \leq \infty.$$

(Филлер (Fieller), 1932, *Biometrika* 24, 428.)

11.23 Показать, что если в примере 11.21 вместо нормально распределенных величин рассматривать две независимые случайные величины, имеющие распределение, о котором шла речь в упражнении 2.12, то их отношение по-прежнему будет иметь распределение Коши (11.85).

(Стек (Steck), 1958, *Ann. Math., Statist.* 29, 604.)

11.24 Совместное распределение случайных величин x_1, x_2 ($x_2 \geq 0$, $M(x_2) = \mu$) имеет ф. п. $f(x_1, x_2)$ и х. ф. $\varphi(t_1, t_2)$. Обобщая (11.74), показать, что

ф. п. величины $u = x_1/x_2$ равна $h(u) = \int_0^\infty f(uv, v) v dv$. Показать далее, что

х. ф. распределения с плотностью

$$g(x_1, x_2) = \frac{x_2}{\mu} f(x_1, x_2)$$

равна

$$\psi(t_1, t_2) = \frac{1}{i\mu} \frac{\partial \varphi(t_1, t_2)}{\partial t_2}$$

и что х. ф. величины $w = x_1 - ux_2$ есть $[\psi(t_1, t_2)]_{t_2 = -ut_1}$. Убедиться в том, что ф. п. величины w при $w=0$ принимает значение $h(u)/\mu$, и, используя формулу обращения, доказать (11.78).

(Даниэлс, 1954.)

11.25 Пусть x — нормированная нормально распределенная случайная величина. Показать, что х. ф. распределения величины $y = 1/x^2$ равна

$$\varphi_y(t) = \exp \{ -(-2it)^{1/2} \}$$

и что, следовательно, х. ф. распределения среднего $\bar{y}$ в выборке объема n из генеральной совокупности с распределением y есть

$$\varphi_{\bar{y}}(t) = \varphi_y(nt) = \varphi_{ny}(t).$$

Таким образом, распределение среднего из n наблюдений совпадает с распределением одного наблюдения, умноженного на n , и распределение y становится все более рассеянным при возрастании n .

АППРОКСИМАЦИЯ ВЫБОРОЧНЫХ РАСПРЕДЕЛЕНИЙ — (1)

12.1 В предыдущей главе мы рассмотрели различные методы получения выборочных распределений в точной форме в предположении, что распределение генеральной совокупности известно полностью. Однако эти методы становятся неприменимыми, если распределение генеральной совокупности не известно точно. Формальное же их применение наталкивается в практике на большие трудности. Например, приходится иметь дело с нетабулированными интегралами. В таких случаях мы сталкиваемся с проблемой отыскания подходящего приближения для выборочного распределения, например, путем нахождения некоторых моментов выборки не очень высокого порядка и построения с их помощью, скажем, одной из кривых системы Пирсона.

Процедуры подобного рода в действительности уже рассматривались в главе 10. Там мы видели, что приближенное выражение для частотного распределения можно получить, отправляясь от первых двух выборочных моментов, которые могут быть выражены через низшие моменты распределения генеральной совокупности. В случае, когда выборочное распределение асимптотически нормально, эта процедура на самом деле решает поставленную проблему, поскольку первые два момента целиком определяют нормальное распределение. Методы настоящей главы являются развитием этой идеи. Мы обсудим точные методы, позволяющие находить моменты выборочных распределений по моментам генеральной совокупности. Наши результаты важны не только тем, что они имеют большой самостоятельный интерес, но и тем, что они при больших n дают точный метод оценки степени аппроксимации выражений, рассмотренных в главе 10. В частности, мы будем в состоянии решить ряд вопросов, лишь слегка затронутых в главе 10, например вопрос о скорости, с которой распределение некоторых функций от моментов выборки (как, скажем, $\sqrt{b_1}$) стремится к нормальному.

12.2 Статистики, с распределениями которых нам обычно приходится иметь дело, можно разбить на три группы. Наибольшую

группу составляют статистики, представляющие собой симметрические функции от результатов наблюдений, т. е. такие функции от результатов наблюдений $x_1, x_2, \dots, x_n$, которые не изменяются при перестановках их. Часто эти статистики оказываются, кроме того, алгебраическими. В этот класс статистик входят арифметическое среднее, моменты, такие величины как b_1 и b_2 , получающиеся из моментов, и смешанные моменты — в многомерном случае. Вторую группу составляют порядковые статистики: медиана, квантили, широты, крайние значения и т. д. В третью группу отнесем такие статистики, как, например, мода, не вошедшие в первые две группы.

Методы, развиваемые нами в этой и следующей главах, применимы только к первой из названных групп статистик. Отметим три различных типа моментов, на изучении которых мы будем останавливаться: (а) моменты истинного или теоретического распределения (например, μ_1, μ_2), (б) моменты выборки (например,

$$m'_r = \frac{1}{n} \sum_{j=1}^n x_j^r, \quad m_r = \frac{1}{n} \sum_{j=1}^n (x_j - m'_1)^r); \quad (\text{в}) \text{ моменты выборочных}$$

распределений, т. е. моменты распределений статистик (например, Mm_r, Dm'_r).

Соответственно мы будем рассматривать теоретические или истинные семиинварианты, семиинварианты выборки и семиинварианты выборочных распределений.

12.3 В главе 10 мы получили следующие точные формулы:

$$\left. \begin{aligned} Mm'_r &= \mu'_r, \\ Dm'_r &= M(m'_r - \mu'_r)^2 = \frac{1}{n} (\mu'_{2r} - \mu'^2_r), \end{aligned} \right\} \quad (12.1)$$

и отмечали, что получение моментов выборочных распределений относительно среднего не столь уже просто. Позднее мы предпочтем другой вывод этих формул, а сейчас только посмотрим, что получится, если подход главы 10 мы попытаемся применить к настоящему случаю. Предположим, к примеру, что нас интересует выборочное распределение дисперсии. Вышеприведенные уравнения дают нам первые два момента ее выборочного распределения относительно произвольной точки. Для первого момента получим

$$\begin{aligned} Mm_2 &= M \left\{ \sum x^2/n - \left(\sum x/n \right)^2 \right\} = \\ &= M \left\{ \frac{n-1}{n^2} \sum x^2 - \frac{1}{n^2} \sum x_i x_j \right\} = \frac{n-1}{n} Mx^2 - \frac{n-1}{n} Mx_i x_j, \quad i \neq j. \end{aligned} \quad (12.2)$$

Поскольку x_i и x_j независимы, то $\mathbf{M}x_i x_j$ равно $(\mu'_1)^2$. Уравнение (12.2) тогда дает

$$\mathbf{M}m_2 = \frac{n-1}{n} (\mu'_2 - \mu_1'^2) = \frac{n-1}{n} \mu_2. \quad (12.3)$$

Приближенное же выражение для $\mathbf{M}m_2$, полученное в главе 10, имеет вид

$$\mathbf{M}m_2 = \mu_2. \quad (12.4)$$

12.4 Довольно ясно, как использовать этот метод с целью получения моментов любого порядка для других статистик, которые могут быть выражены как симметрические функции (рациональные или целые) от результатов наблюдений. Например, для нахождения четвертого момента дисперсии выборки нам пришлось бы сначала разложить $\{\sum x^2/n - (\sum x/n)^2\}^4$ в сумму произведений вида $\sum x_i^a x_j^b \dots x_k^y$, затем произвести усреднение, что даст $n(n-1) \dots (n-t+1) \mu'_a \mu'_b \dots \mu'_y$, где t — число различных индексов $i, j, \dots, k$ в сумме. Таким образом, мы получим требуемый момент в терминах исходных моментов, т. е. моментов теоретического распределения.

12.5 Как видим, этот метод достаточно прост. Однако он сопровождается утомительными алгебраическими преобразованиями и приводит к громоздким выражениям, за исключением разве лишь очень простых случаев. В целях упрощения дальнейших выкладок мы используем некоторые понятия, связанные с симметрическими функциями.

Предположим, что мы имеем множество $x_1, \dots, x_n$. При записи выражений типа $\sum x_i^2 x_j x_l^3 x_k$ мы будем всегда предполагать, что все участвующие индексы различны и что суммирование производится по всем индексам с различными индексами. Поэтому в данном примере суммируется всего $n(n-1)(n-2)(n-3)$ членов. Нужно заметить, что в некоторых выражениях, таких как, например, $\sum x_i x_j$, каждая пара встречается дважды (к примеру, $x_1 x_2$ и $x_2 x_1$), тогда как в $\sum x_i^2 x_j x_1^2 x_2$ и $x_2^2 x_1$ встречаются только по одному разу.

Назовем *расширенной* симметрическую функцию, определенную соотношением

$$[p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}] = \sum x_i^{p_1} x_j^{p_1} \dots x_q^{p_2} x_r^{p_2} \dots x_n^{p_s} x_v^{p_s}, \quad (12.5)$$

где π_i — число степеней p_i . Например,

$$\sum x_i^2 x_j x_l^3 x_k = [1^2 2 3], \quad \sum x_i^2 x_j^2 x_k^2 = [2^3].$$

Более часто используется не расширенная симметрическая функция, а *мономиальная*, определяемая формулой

$$(p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}) = [p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}] / \{\pi_1! \pi_2! \dots \pi_s!\} \quad (12.6)$$

Особенно важны два ее частных случая: унитарная функция

$$a_r = (1^r) = \sum x_i x_j \dots x_l / r! \quad (12.7)$$

и сумма степеней

$$s_r = (r) = \sum_i x_i^r = [r]. \quad (12.8)$$

Имеются функции, выражающие введенные функции одну через другую. Наиболее часто применяются таблицы, дающие суммы степеней через расширенные функции и наоборот (Дэвид и Кендалл, 1949). Из (12.5) мы сразу же получаем следующий важный результат:

$$M[p_1^{\pi_1} p_2^{\pi_2} \dots p_s^{\pi_s}] = n(n-1) \dots \dots (n-\rho+1) (\mu'_{p_1})^{\pi_1} (\mu'_{p_2})^{\pi_2} \dots (\mu'_{p_s})^{\pi_s}, \quad (12.9)$$

где $\rho = \sum_{i=1}^s \pi_i$ и $p = \sum_{i=1}^s p_i \pi_i$ — вес симметрической функции.

Для $p \leq 6$ в конце книги в приложении (таблица 10) приводятся таблицы, аналогичные таблицам Дэвид и Кендалла.

Пример 12.1

Дисперсию выборки в терминах сумм степеней можно записать как

$$m_2 = \frac{(2)}{n} - \frac{(1)^2}{n^2}.$$

Непосредственно или с помощью таблицы 10 устанавливаем, что

$$(2) = [2]; (1)^2 = [2] + [1^2].$$

Следовательно,

$$m_2 = \frac{[2]}{n} - \frac{[2] + [1^2]}{n^2} = \frac{n-1}{n^2} [2] - \frac{1}{n^2} [1^2], \quad (12.10)$$

откуда, беря математическое ожидание и учитывая (12.9), находим

$$Mm_2 = \frac{n-1}{n^2} n\mu'_2 - \frac{1}{n^2} n(n-1)\mu_1'^2 = \frac{n-1}{n} \mu_2. \quad (12.11)$$

Мы видим, что введенные нами функции существенно сокращают выкладки. Статистика m_2 не зависит от выбора начала

отсчета и, следовательно, ее выборочные моменты не зависят от μ'_1 . Без ограничения общности поэтому можно μ'_1 считать равным нулю, а это в свою очередь позволяет считать равной нулю каждую расширенную симметрическую функцию, содержащую единицу. Поэтому из (12.10) сразу получаем

$$Mm_2 = \frac{n-1}{n} \mu_2.$$

Аналогично можно найти

$$m_2^2 = \frac{(2)^2}{n^2} - \frac{2(2)(1)^2}{n^3} + \frac{(1)^4}{n^4}. \quad (12.12)$$

Согласно таблицам

$$(2)^2 = [4] + [2^2],$$

$$(2)(1)^2 = [4] + 2[31] + [2^2] + [21^2],$$

$$(1)^4 = [4] + 4[31] + 3[2^2] + 6[21^2] + [1^4].$$

Пренебрегая теми расширенными симметрическими функциями, которые содержат единицы, и пользуясь предыдущими выражениями, получим

$$\begin{aligned} m_2^2 &= \frac{[4] + [2^2]}{n^2} - 2 \frac{[4] + [2^2]}{n^3} + \frac{[4] + 3[2^2]}{n^4} = \\ &= \frac{(n-1)^2}{n^4} [4] + \frac{n^2 - 2n + 3}{n^4} [2^2], \end{aligned}$$

откуда

$$Mm_2^2 = \frac{(n-1)^2}{n^3} \mu_4 + \frac{(n-1)(n^2 - 2n + 3)}{n^3} \mu_2^2. \quad (12.13)$$

Используя (12.11), найдем также

$$Dm_2 = Mm_2^2 - (Mm_2)^2 = \frac{(n-1)^2}{n^3} \mu_4 - \frac{(n-1)(n-3)}{n^3} \mu_2^2, \quad (12.14)$$

что для больших n равно примерно

$$Dm_2 = \frac{\mu_4 - \mu_2^2}{n}.$$

Этот результат согласуется с тем, который мы получили в (10.9) для $r=2$.

Если в (12.14) положить $\kappa_4 = \mu_4 - 3\mu_2^2$, $\kappa_2 = \mu_2$, то можно записать

$$Dm_2 = \left(\frac{n-1}{n} \right)^2 \left\{ \frac{\kappa_4}{n} + \frac{2\kappa_2^2}{n-1} \right\}. \quad (12.15)$$

***k*-статистики**

12.6 Алгебраическая сложность этого метода и большой объем вычислительной работы, требуемый при его использовании, особенно до того, как были созданы упомянутые выше таблицы симметрических функций, привели к поискам более простых методов. В 1928 году сэр Рональд Фишер революционизировал этот предмет в следующих двух направлениях: предложил новые симметрические функции, так называемые *k*-статистики, характеризующие распределение, и показал, как их выборочные семинварианты могут быть получены комбинаторными методами.

Рассмотрим семейство статистик $k_1, k_2, \dots, k_p, \dots$, являющихся симметрическими функциями от результатов наблюдений, обладающих тем свойством, что среднее значение статистики k_p является *p*-м семинвариантом:

$$Mk_p = \kappa_p. \quad (12.16)$$

Предостережем читателя от возможных недоразумений. Величины μ_p и m_p по своему характеру имеют одно и то же происхождение, а именно они являются соответственно теоретическими моментами и моментами выборки, и Mm_p не равно μ_p , точно так же k_p и κ_p различны по своему характеру, но зато $Mk_p = \kappa_p$.

12.7 Заметим прежде всего, что k_p однозначно определяются формулой (12.16). В самом деле, если имеются две функции k_p и k'_p , удовлетворяющие (12.16), то их разность $k_p - k'_p$ будет иметь нулевое среднее значение. Но эта разность сама является симметрической функцией и может быть поэтому выражена как сумма членов $\sum x^p$, $\sum x_j x_k^{p-1}$ и т. д., и, следовательно, ее среднее значение представляет собой сумму, состоящую из произведений моментов. Обращение этой суммы в нуль будет означать, что между моментами существуют некоторые соотношения, что возможно лишь для отдельных частных совокупностей. Следовательно, $k_p = k'_p$.

Во-вторых, отметим, что статистики k_p подобно центральным моментам не зависят от начала отсчета, за исключением k_1 , являющейся средним значением. Действительно, по формуле Тейлора

$$\begin{aligned} k_p(x_1 + h, x_2 + h, \dots, x_n + h) &= \\ &= k_p(x_1, x_2, \dots, x_n) + \frac{h}{1!} Dk_p(x_1, x_2, \dots, x_n) + \\ &\quad + \frac{h^2}{2!} D^2k_p(x_1, x_2, \dots, x_n) + \dots, \end{aligned} \quad (12.17)$$

где

$$D = \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_n}.$$

Беря среднее значение и вспоминая, что $k_p (p > 1)$ сами не зависят от начала отсчета, мы получим

$$\kappa_p = \kappa_p + \frac{h}{1!} M(Dk_p) + \dots \quad (12.18)$$

В силу произвольности h из (12.18) следует, что $M(Dk_p)$ и остальные слагаемые обращаются в нуль, откуда согласно сделанному выше замечанию $Dk_p = 0$, $D^2 k_p = 0$ и т. д. Следовательно,

$$k_p(x_1 + h, x_2 + h, x_n + h) = k_p(x_1, x_2, \dots, x_n).$$

Исключение составляет лишь статистика k_1 , среднее значение которой $\kappa_1 = \mu'_1$, т. е.

$$k_1 = \frac{1}{n} \sum x. \quad (12.19)$$

12.8 Сейчас мы найдем точные выражения для k -статистик в терминах результатов наблюдений $x_1, \dots, x_n$. По определению k_p относительно этих наблюдений имеет степень p (точно так же как κ_p относительно моментов имеет порядок p , т. е. в любом члене κ_p сумма порядков моментов равна p). Поэтому k_p -статистика согласно (12.16) должна иметь следующий вид:

$$k_p = \sum \sum \left(x_{i_1}^{p_1} \dots x_{i_{\pi_1}}^{p_1} x_{i_{\pi_1+1}}^{p_2} \dots \right. \\ \left. \dots x_{i_{\pi_1+\pi_2}}^{p_2} \dots x_{i_{\pi_1+\dots+\pi_s}}^{p_s} \right) A(p_1^{\pi_1} \dots p_s^{\pi_s}), \quad (12.20)$$

где внутреннее суммирование производится по всевозможным несовпадающим индексам $i_1, \dots, i_{\pi_1+\dots+\pi_s}$, каждый из которых может принимать любое из n значений $1, 2, \dots, n$. Внешнее суммирование производится по всевозможным разбиениям числа p :

$$p = p_1 \pi_1 + \dots + p_s \pi_s. \quad (12.21)$$

Каждое такое разбиение числа p будем обозначать $(p_1^{\pi_1} \dots p_s^{\pi_s})$. $A(p_1^{\pi_1} \dots p_s^{\pi_s})$ — число, зависящее от этого разбиения. Определим далее число ρ , положив

$$\rho = \pi_1 + \dots + \pi_s. \quad (12.22)$$

Всюду мы предполагаем, конечно, что $n \geq p$. Беря среднее значение от обеих частей равенства (12.20) и учитывая независи-

мость иксов, мы получим

$$\kappa_p = \sum \left\{ \left(\mu_{p_1}^{\pi_1} \mu_{p_2}^{\pi_2} \dots \mu_{p_s}^{\pi_s} \right) AB \right\}, \quad (12.23)$$

где B есть число способов, которыми p индексов можно извлечь из совокупности, состоящей из n индексов, включая и перестановки, т. е. B равно $n(n-1) \dots (n-p+1) = n^{[p]}$.

Из уравнения (3.39) имеем

$$\kappa_p = p! \sum \sum \left(\frac{\mu_{p_1}}{p_1!} \right)^{\pi_1} \dots \left(\frac{\mu_{p_s}}{p_s!} \right)^{\pi_s} \frac{(-1)^{p-1} (p-1)!}{\pi_1! \dots \pi_s!}, \quad (12.24)$$

где суммирование производится по всевозможным разбиениям, отвечающим (12.21) и (12.22). Отождествляя соответствующие члены в (12.23) и (12.24), мы найдем значения коэффициентов $A(p_1^{\pi_1} \dots p_s^{\pi_s})$, после подстановки которых в (12.20) получим окончательно

$$k_p = \frac{p! \sum (-1)^{p-1} (p-1)!}{n^{[p]}} \sum \frac{x_{i_1}^{p_1} \dots x_{i_{\pi_1}}^{p_1} \dots x_{i_p}^{p_s}}{(p_1!)^{\pi_1} \dots (p_s!)^{\pi_s} \pi_1! \dots \pi_s!}, \quad (12.25)$$

где второе суммирование производится по всем несовпадающим индексам $i_1, \dots, i_p$.

Приведем иную, более простую, запись этого выражения, которая чрезвычайно облегчает исследование выборочных свойств k -статистик. Отвлекаясь от множителей, зависящих от p и n , типичное слагаемое в (12.25) можно представить как

$$p! \left(\frac{x_1^{p_1}}{p_1!} \frac{x_2^{p_2}}{p_1!} \dots \frac{x_{\pi_1}^{p_1}}{p_1!} \dots \frac{x_p^{p_s}}{p_s!} \right) \frac{1}{\pi_1! \dots \pi_s!}.$$

Каждому разбиению чисел p и p (по формулам (12.21) и (12.22)) отвечает несколько членов указанного вида. Следовательно, мы можем записать

$$k_p = \sum \frac{(-1)^{p-1} (p-1)!}{n^{[p]}} \sum (x_{\gamma_1} x_{\gamma_2} \dots x_{\gamma_p}), \quad (12.26)$$

где второе суммирование производится по всевозможным индексам $\gamma_1, \dots, \gamma_p$, которые изменяются независимо один от другого. Поэтому в этой сумме наверняка содержится (с точностью до константы) и выделенное выше типичное слагаемое. В действительности формулы (12.25) и (12.26) в точности совпадают. В самом деле, пусть из n индексов $\gamma_1, \dots, \gamma_p$ мы выбираем p_1 индексов равными i_1 , p_1 индексов равными $i_2, \dots, p_1$ индексов равными i_{π_1} , p_2 индексов равными i_{π_1+1} и т. д. Тогда нетрудно

понять, что в $\sum (x_{\nu_1} \dots x_{\nu_p})$ будет существовать ровно

$$\frac{p!}{(p_1!)^{\pi_1} \dots (p_s!)^{\pi_s} \pi_1! \dots \pi_s!} \quad (12.27)$$

слагаемых вида $x_{i_1}^{p_1} x_{i_2}^{p_2} \dots x_{i_p}^{p_p}$, что и доказывает совпадение формул (12.25) и (12.26).

Представление k -статистик в терминах симметрических произведений и сумм

12.9 Воспользовавшись выражением семиинвариантов через моменты, мы можем представить k -статистики в терминах расширенных симметрических функций. Например,

$$\kappa_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3,$$

откуда, сравнивая (12.24) и (12.25), получаем

$$k_3 = \frac{[3]}{n} - \frac{3[21]}{n(n-1)} + \frac{2[1^3]}{n(n-1)(n-2)}.$$

Заменив расширенные симметрические функции по формулам

$$[3] = (3), \quad [21] = -(3) + (2)(1), \quad [1^3] = 2(3) - 3(2)(1) + (1)^3,$$

найдем

$$k_3 = \frac{1}{n(n-1)(n-2)} \{n^2(3) - 3n(2)(1) + 2(1)^3\}.$$

Полагая $s_r = r$, можно записать $k_3 = \frac{1}{n(n-1)(n-2)} \times \times (n^2s_3 - 3ns_2s_1 + 2s_1^3)$.

12.10 Первые восемь k -статистик в терминах степенных сумм s_r выражаются следующим образом:

$$\left. \begin{aligned} k_1 &= \frac{1}{n} s_1, \\ k_2 &= \frac{1}{n^{[2]}} (ns_2 - s_1^2), \\ k_3 &= \frac{1}{n^{[3]}} (n^2s_3 - 3ns_2s_1 + 2s_1^3), \\ k_4 &= \frac{1}{n^{[4]}} \{ (n^3 + n^2)s_4 - 4(n^2 + n)s_3s_1 - 3(n^2 - n)s_2^2 + \\ &\quad + 12ns_2s_1^2 - 6s_1^4 \}, \\ k_5 &= \frac{1}{n^{[5]}} \{ (n^4 + 5n^3)s_5 - 5(n^3 + 5n^2)s_4s_1 - 10(n^3 - n^2)s_3s_2 + \\ &\quad + 20(n^2 + 2n)s_3s_1^2 + 30(n^2 - n)s_2^2s_1 - 60ns_2s_1^3 + 24s_1^5 \}, \end{aligned} \right\} \quad (12.28)$$

$$\begin{aligned}
 k_6 &= \frac{1}{n^{[6]}} \{ (n^5 + 16n^4 + 11n^3 - 4n^2) s_6 - 6(n^4 + 16n^3 + \\
 &\quad + 11n^2 - 4n) s_5 s_1 - 15n(n-1)^2(n+4) s_4 s_2 - \\
 &\quad - 10(n^4 - 2n^3 + 5n^2 - 4n) s_3^2 + 30(n^3 + 9n^2 + 2n) s_4 s_1^2 + \\
 &\quad + 120(n^3 - n) s_3 s_2 s_1 + 30(n^3 - 3n^2 + 2n) s_3^2 - \\
 &\quad - 120(n^2 + 3n) s_3 s_1^3 - 270(n^2 - n) s_2^2 s_1^2 + 360n s_2 s_4^2 - 120s_1^6 \}, \\
 k_7 &= \frac{1}{n^{[7]}} \{ (n^6 + 42n^5 + 119n^4 - 42n^3) s_7 - 7(n^5 + 42n^4 + \\
 &\quad + 119n^3 - 42n^2) s_6 s_1 - 21(n^5 + 12n^4 - 31n^3 + 18n^2) s_5 s_2 - \\
 &\quad - 35(n^5 + 5n^3 - 6n^2) s_4 s_3 + 42(n^4 + 27n^3 + 44n^2 - 12n) s_5 s_1^2 + \\
 &\quad + 210(n^4 + 6n^3 - 13n^2 + 6n) s_4 s_2 s_1 + \\
 &\quad + 140(n^4 + 5n^2 - 6n) s_3^2 s_1 + 210(n^4 - 3n^3 + 2n^2) s_3 s_2^2 - \\
 &\quad - 210(n^3 + 13n^2 + 6n) s_4 s_1^3 - 1260(n^3 + n^2 - 2n) s_3 s_2 s_1^2 - \\
 &\quad - 630(n^3 - 3n^2 + 2n) s_3^2 s_1 + 840(n^2 + 4n) s_3 s_1^4 + \\
 &\quad + 2520(n^2 - n) s_2^2 s_1^3 - 2520n s_2 s_4^2 + 720s_1^7 \}, \\
 k_8 &= \frac{1}{n^{[8]}} \{ (n^7 + 99n^6 + 757n^5 + 141n^4 - 398n^3 + 120n^2) s_8 - \\
 &\quad - 8(n^6 + 99n^5 + 757n^4 + 141n^3 - 398n^2 + 120n) s_7 s_1 - \\
 &\quad - 28(n^6 + 37n^5 - 39n^4 - 157n^3 + 278n^2 - 120n) s_6 s_2 - \\
 &\quad - 56(n^6 + 9n^5 - 23n^4 + 111n^3 - 218n^2 + 120n) s_5 s_3 - \\
 &\quad - 35(n^6 + n^5 + 33n^4 - 121n^3 + 206n^2 - 120n) s_4^2 + \\
 &\quad + 56(n^5 + 68n^4 + 359n^3 - 8n^2 - 60n) s_6 s_1^2 + 336(n^5 + \\
 &\quad + 23n^4 - 31n^3 - 23n^2 + 30n) s_5 s_2 s_1 + 560(n^5 + 5n^4 + \\
 &\quad + 5n^3 + 5n^2 - 6n) s_4 s_3 s_1 + 420(n^5 + 2n^4 - 25n^3 + \\
 &\quad + 46n^2 - 24n) s_4 s_2^2 + 560(n^5 - 4n^4 + 11n^3 - 20n^2 + \\
 &\quad + 12n) s_3^2 s_2 - 336(n^4 + 38n^3 + 99n^2 - 18n) s_5 s_1^3 - \\
 &\quad - 2520(n^4 + 10n^3 - 17n^2 + 6n) s_4 s_2 s_1^2 - 1680(n^4 + 2n^3 + \\
 &\quad + 7n^2 - 10n) s_3^2 s_1^2 - 5040(n^4 - 2n^3 - n^2 + 2n) s_3 s_2^2 s_1 - \\
 &\quad - 630(n^4 - 6n^3 + 11n^2 - 6n) s_2^4 + 1680(n^3 + 17n^2 + \\
 &\quad + 12n) s_4 s_1^4 + 13440(n^3 + 2n^2 - 3n) s_3 s_2 s_1^3 + \\
 &\quad + 10080(n^3 - 3n^2 + 2n) s_2^3 s_1^2 - 6720(n^2 + 5n) s_3 s_1^5 - \\
 &\quad - 25200(n^2 - n) s_2^2 s_1^4 + 20160n s_2 s_4^2 - 5040s_1^8 \}.
 \end{aligned} \tag{12.28}$$

В частности,

$$\left. \begin{aligned} k_1 &= m'_1, \\ k_2 &= \frac{n}{n-1} m_2, \\ k_3 &= \frac{n^2}{(n-1)(n-2)} m_3, \\ k_4 &= \frac{n^2}{(n-1)(n-2)(n-3)} \{(n+1)m_4 - 3(n-1)m_2^2\}, \end{aligned} \right\} \quad (12.29)$$

что дает выражение k -статистик в терминах выборочных моментов.

Значения k_9 и k_{10} были даны Зя Уд-дином (1954).

12.11 В силу хорошо известной теоремы каждая рациональная целая симметрическая функция от $x_1, \dots, x_n$ может быть однозначно представлена как некоторая рациональная целая функция от сумм вида s_r . Поэтому такие функции могут быть выражены через k -статистики, поскольку в силу уравнений, аналогичных (12.28), суммы s_r можно выразить в терминах k -статистик. Тем самым изучение выборочных характеристик таких симметрических функций сводится к рассмотрению свойств k -статистик.

Читатель, который последующие алгебраические преобразования готов принять на веру, может при чтении оставшейся части этой и следующей главы опустить доказательства, усвоив лишь основные результаты. Мы сейчас переходим к изложению комбинаторных методов, позволяющих выразить семиинварианты k -статистик в терминах семиинвариантов исходного распределения. Будет показано, что эти результаты получаются существенно проще, чем эквивалентные им результаты, найденные в терминах моментов.

Семиинварианты k -статистик

12.12 Задача нахождения моментов или семиинвариантов k -статистик сводится к отысканию средних значений для различных степеней и произведений этих статистик. Каждому разбиению $(a_1^{a_1} a_2^{a_2} \dots a_s^{a_s})$ числа a соответствуют момент

$$\mu'(a_1^{a_1} \dots a_s^{a_s}) = \mathbf{M}(k_{a_1}^{a_1} \dots k_{a_s}^{a_s}) \quad (12.30)$$

и семинвариант $\kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s})$, связанный с моментами тождеством (см. (3.28))

$$\sum \left\{ \kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s}) \frac{t_{a_1}^{\alpha_1}}{\alpha_1!} \dots \frac{t_{a_s}^{\alpha_s}}{\alpha_s!} \right\} = \\ = \log \left\{ \sum \mu'(b_1^{\beta_1} \dots b_m^{\beta_m}) \frac{t_{b_1}^{\beta_1}}{\beta_1!} \dots \frac{t_{b_m}^{\beta_m}}{\beta_m!} \right\}. \quad (12.31)$$

К примеру, четвертый семинвариант статистики k_2 будет выражаться через четвертый момент k_2 и моменты низших порядков. В соответствии с (12.30) эти величины будут записываться как $\kappa(2^4)$ и $\mu'(2^4)$. Точно так же семинвариант $\kappa(32)$ соответствует моменту $\mu'(32) = Mk_3k_2$. Вообще же для каждого данного a формулы будут различны в зависимости от разбиения этого числа.

Заметим, что произведение $k_{a_1}^{\alpha_1} \dots k_{a_s}^{\alpha_s}$ однородно относительно иксов, суммарная степень которых равна a . Следовательно, среднее значение $\mu'(a_1^{\alpha_1} \dots a_s^{\alpha_s}) = Mk_{a_1}^{\alpha_1} \dots k_{a_s}^{\alpha_s}$ однородно относительно теоретических моментов μ' , суммарный порядок которых равен также a . Но поскольку κ сами являются однородными функциями от μ , то $\kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s})$ является однородной функцией от κ . Отсюда следует первое правило для семинвариантов k -статистик.

Правило 1. $\kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s})$ представляет собой сумму слагаемых, каждое из которых с точностью до константы является произведением $\kappa_{a_1}^{\alpha_1} \dots \kappa_{a_s}^{\alpha_s}$ исходных семинвариантов таких, что $a_1\alpha_1 + \dots + a_s\alpha_s = a$.

Например, в случае $\kappa(2^4)$ $a=8$ и поэтому $\kappa(2^4)$ состоит из суммы членов (с некоторыми коэффициентами) вида κ_8 , $\kappa_6\kappa_2$, $\kappa_5\kappa_3$, κ_4^2 , $\kappa_4\kappa_2^2$, $\kappa_3^2\kappa_2$ и κ_2^4 . Точно так же $\kappa(32)$ будет содержать только члены с κ_5 , $\kappa_3\kappa_2$. Заметим, что в согласии со следующим правилом в этих произведениях нигде не участвует семинвариант κ_1 .

Правило 2. В представлении $\kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s})$ не содержится членов κ_1 . Исключение составляет лишь $\kappa(1) = \kappa_1$.

Это правило выводится точно так же, как это делается для частного случая в примере 12.1. Статистики k не зависят от выбора начала отсчета, и, следовательно, их выборочное распределение не зависит от κ_1 . Исключение составляет лишь одна статистика k_1 , для которой по определению $\kappa(1) = \kappa_1$.

12.13 Мы сейчас сформулируем и проиллюстрируем правила, по которым $\kappa(a_1^{a_1} \dots a_s^{a_s})$ может быть выражен через семиинварианты исходного распределения. Доказательство справедливости этих правил довольно трудно и будет отложено до следующей главы.

Для нахождения членов $\kappa_{b_1}^{\beta_1} \dots \kappa_{b_m}^{\beta_m}$ в $\kappa(a_1^{a_1} \dots a_s^{a_s})$ рассмотрим таблицу с двумя входами

$$\begin{array}{c|c}
 & \begin{array}{c} b_1 \\ b_1 \\ \vdots \\ \vdots \\ b_2 \\ \vdots \\ \vdots \end{array} \\
 \hline
 \begin{array}{c} a_1 a_1 \dots a_2 \dots \end{array} & a
 \end{array} \quad (12.32)$$

где имеется β_i строк, отвечающих b_i и α_j столбцов, отвечающих a_j .

Рассмотрим различные способы, которыми эта таблица может быть заполнена такими числами, что их суммы по строкам и столбцам равны соответственно числам b и a , стоящим справа и снизу в таблице напротив этих строк и столбцов. Например, если мы интересуемся коэффициентом, с которым $\kappa_2 \kappa_2^2$ входит в $\kappa(4^2 2)$, то мы должны рассмотреть следующие таблицы:

$$\begin{array}{c|c|c|c}
 \begin{array}{ccc|c} 2 & 2 & 2 & 6 \\ 1 & 1 & . & 2 \\ 1 & 1 & . & 2 \\ \hline 4 & 4 & 2 & 10 \end{array} &
 \begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 1 & 1 & . & 2 \\ 1 & . & 1 & 2 \\ \hline 4 & 4 & 2 & 10 \end{array} &
 \begin{array}{ccc|c} 3 & 3 & . & 6 \\ 1 & . & 1 & 2 \\ . & 1 & 1 & 2 \\ \hline 4 & 4 & 2 & 10 \end{array} &
 \end{array} \quad (12.33)$$

Тогда правило, по которому отыскивается коэффициент у $\kappa_{b_1}^{\beta_1} \dots \kappa_{b_m}^{\beta_m}$, состоит в следующем.

П р а в и л о 3. Каждая таблица, числа которой образуют два или более непересекающихся блоков, должна не приниматься во внимание. Так, в предыдущем примере одна из возможных таблиц

$$\begin{array}{c|c|c|c}
 \begin{array}{ccc|c} 4 & 2 & . & 6 \\ . & 2 & . & 2 \\ . & . & 2 & 2 \\ \hline 4 & 4 & 2 & 10 \end{array} & & &
 \end{array}$$

должна быть отброшена, поскольку блок 2×2 в левом верхнем углу не имеет с блоком 1×1 , стоящим в правом нижнем углу, общих строк и столбцов.

П р а в и л о 4. Коэффициент при произведении $\kappa_{b_1}^{\beta_1} \dots \kappa_{b_m}^{\beta_m}$, участвующем в представлении $\kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s})$, складывается из коэффициентов, зависящих от способа заполнения таблиц типа (12.32) (таблицы, о которых шла речь в правиле 3, при подсчете коэффициента во внимание не принимаются). С каждой такой таблицей связан свой численный коэффициент, который, будучи умноженным на некоторую функцию, зависящую от n , и дает тот вклад, который эта таблица вносит в искомый коэффициент.

П р а в и л о 5. Связанный с каждой таблицей численный коэффициент, о котором говорилось выше, получается следующим образом. Пусть в рассматриваемой таблице в i -м столбце находятся элементы $C_1^{(i)}, \dots, C_{k_i}^{(i)}$, $C_1^{(i)} + \dots + C_{k_i}^{(i)} = a^{(i)}$, где $a^{(i)}$ — одно из чисел $a_1, a_1, \dots, a_2, \dots, a_s$. Тогда искомый численный коэффициент равен

$$\prod_{i=1}^{a_1 \alpha_1 + \dots + a_s \alpha_s} \frac{a^{(i)}!}{C_1^{(i)}! \dots C_{k_i}^{(i)}!} \cdot \frac{1}{\beta_1! \dots \beta_m!},$$

т. е. равен числу различных способов (поделенному на $\beta_1! \dots \beta_m!$), которыми можно заполнить таблицу с суммарной строкой $(a_1, a_1, \dots, a_2, \dots, a_s)$.

П р а в и л о 6. Функция от n , называемая шаблонной функцией, определяется только расположением нулей в таблице и не зависит от чисел, заполняющих таблицу, и чисел, составляющих суммарные строки $(a_1, a_1, \dots)$ и столбцы $(b_1, b_1, \dots)$. Эта функция получается в результате рассмотрения разбиений строк на разные группы:

(I) Если такое разбиение состоит из одной группы, то с ним связывается число n , если из двух, то $n(n-1), \dots$, с q группами связывается число $n(n-1) \dots (n-q+1)$.

(II) Если имеется какое-то разбиение, то мы подсчитываем число групп, в пересечении с которыми фиксированный столбец имеет ненулевые входы. Если таких групп ρ , то с ними связывается множитель

$$\frac{(-1)^{\rho-1} (\rho-1)!}{n(n-1) \dots (n-\rho+1)}.$$

Эти подсчеты проводятся для всех столбцов и соответствующие множители перемножаются.

(III) Для каждого разбиения коэффициенты, получаемые в (I) и (II), перемножаются.

(IV) Суммируются все произведения из (III) по различным разбиениям. Суммарное их выражение называется шаблонной функцией.

Правило 7. Для каждой таблицы, содержащей строку, состоящую лишь из одного ненулевого элемента, шаблонная функция равна нулю. Поэтому такие таблицы при подсчете коэффициентов можно не принимать в расчет.

Правило 8. Каждая таблица, содержащая столбец, в котором только один элемент отличен от нуля, имеет шаблонную функцию, в n раз меньшую, чем шаблонная функция для таблицы, получающейся из первоначальной путем вычеркивания этого столбца.

Правило 9. Каждая таблица, ненулевые входы которой образуют две подтаблицы, связанные только посредством общего столбца, имеет нулевую шаблонную функцию и поэтому может не приниматься в расчет.

Пример 12.2

В качестве иллюстрации этих правил (которые вовсе не так трудны, как это может показаться) найдем коэффициент при $x_6 x_2^2$ в $\kappa(4^2 2)$. Если читатель выпишет около тридцати таблиц вида (12.32) с суммарной строкой $(4, 4, 2)$ и суммарным столбцом $(6, 2, 2)$, то он сможет обнаружить, что учитывать надо только таблицы (12.33) и те, которые из них получаются с помощью перестановки строк и столбцов с сохранением суммарной строки $(4, 4, 2)$ и суммарного столбца $(6, 2, 2)$. Вот эти таблицы:

$$\begin{array}{c}
 \begin{array}{ccc|ccc} 2 & 2 & 2 & 6 & 2 & 3 & 1 \end{array} \\
 \begin{array}{ccc|ccc} 1 & 1 & . & 2 & 1 & 1 & . \end{array} \\
 \begin{array}{ccc|ccc} 1 & 1 & . & 2 & 1 & . & 1 \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{ccc|ccc} 6 & 3 & 2 & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & 1 & . \end{array} \\
 \begin{array}{ccc|ccc} 2 & . & 1 & 1 \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{ccc|ccc} 6 & 2 & 3 & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & . & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & 1 & . \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{ccc|ccc} 6 & 2 & 3 & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & . & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & 1 & . \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}$$

(a)

(b)

(c)

(d)

$$\begin{array}{c}
 \begin{array}{ccc|ccc} 3 & 2 & 1 & 6 & 3 & 3 & . \end{array} \\
 \begin{array}{ccc|ccc} . & 1 & 1 & 2 & 1 & . & 1 \end{array} \\
 \begin{array}{ccc|ccc} 1 & 1 & . & 2 & . & 1 & 1 \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{ccc|ccc} 6 & 3 & 3 & . \end{array} \\
 \begin{array}{ccc|ccc} 2 & . & 1 & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & . & 1 \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}
 \quad
 \begin{array}{c}
 \begin{array}{ccc|ccc} 6 & 3 & 3 & . \end{array} \\
 \begin{array}{ccc|ccc} 2 & . & 1 & 1 \end{array} \\
 \begin{array}{ccc|ccc} 2 & 1 & . & 1 \end{array} \\
 \hline
 \begin{array}{ccc|ccc} 4 & 4 & 2 & 10 & 4 & 4 & 2 \end{array}
 \end{array}$$

(e)

(f)

(g)

(12.34)

В действительности не было надобности выписывать такие таблицы, как (с), (d) и (e), которые получаются из (b) перестановкой столбцов и строк. Мы их выписали только для ясности. При рассмотрении таблиц (b) и (f) надо иметь в виду следующую тонкость. В таблице (b) сумма по двум столбцам равна четырем, а по двум строкам равна двум, и в результате их перестановок получатся четыре различные таблицы. Из таблицы же (f), хотя все ее столбцы и строки различны, в результате перестановок получаются всего лишь две различные таблицы.

Каждая из приведенных таблиц будет давать свой коэффициент. Рассмотрим сначала таблицу (a). Числовой коэффициент здесь согласно правилу 5 равен

$$\left(\frac{4!}{2!1!1!1!}\right)\left(\frac{4!}{2!1!1!1!}\right)\frac{1}{2!}=72.$$

Согласно правилу 8 шаблонная функция равна $1/n$ от аналогичной функции, отвечающей таблице

$$\begin{array}{cc} \times & \times \\ \times & \times \\ \times & \times \end{array}$$

Здесь мы имеем следующие пять разбиений:

$$\begin{array}{cccc} \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times \end{array}$$

Первое разбиение, состоящее из одной группы, является просто исходной таблицей. Каждое из следующих трех разбиений состоит из двух групп, получаемых, если первую, вторую и третью строки брать в качестве отдельных групп. Последнее разбиение состоит из трех групп, каждая из которых состоит из одной строки.

Тогда согласно правилу 6 соответствующие этим разбиениям слагаемые, составляющие шаблонную функцию, равны

$$n\left(\frac{1}{n}\right)\left(\frac{1}{n}\right)=\frac{1}{n},$$

$$3n(n-1)\left\{\frac{-1}{n(n-1)}\right\}\left\{\frac{-1}{n(n-1)}\right\}=\frac{3}{n(n-1)},$$

$$n(n-1)(n-2)\left\{\frac{(-1)^2 2!}{n(n-1)(n-2)}\right\}\left\{\frac{(-1)^2 2!}{n(n-1)(n-2)}\right\}=\frac{4}{n(n-1)(n-2)}.$$

Сумма их равна $\frac{n}{(n-1)(n-2)}$, откуда следует, что вклад, даваемый в искомый коэффициент таблицей (12.34) (а), равен $\frac{72}{(n-1)(n-2)}$.

Таблицы (b) — (e) дают один и тот же вклад и могут поэтому рассматриваться вместе. Для каждой из них численный множитель равен

$$\left(\frac{4!}{2!1!1!1!}\right)\left(\frac{4!}{3!1!}\right)\left(\frac{2!}{1!1!}\right) \cdot \frac{1}{2!} = 48,$$

и поскольку этих таблиц четыре, получим, что суммарный численный множитель равен 192.

Согласно правилу 6 шаблонная функция зависит от конфигурации

$$\begin{array}{ccc} \times & \times & \times \\ \times & \times & . \\ \times & . & \times \end{array}$$

где знак $\times$ означает ненулевой вход, а $.$ — нулевой. Здесь существуют пять разбиений: одно из одной группы, три из двух групп и одно из трех групп. Вклад, даваемый первым разбиением, равен

$$n \frac{1}{n} \cdot \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2},$$

поскольку в этом разбиении каждый столбец имеет ненулевой вход. Вклад от трех разбиений, получаемых при рассмотрении первой, второй и третьей строк в качестве отдельных групп, равен

$$n(n-1) \left[\frac{-1}{n^3(n-1)^3} + \frac{1}{n^3(n-1)^2} + \frac{1}{n^3(n-1)^2} \right] = \frac{2n-3}{n^2(n-1)^2}.$$

Вклад от разбиения, состоящего из трех групп, равен

$$n(n-1)(n-2) \left[\frac{2!}{n(n-1)(n-2)} \cdot \frac{-1}{n(n-1)} \cdot \frac{-1}{n(n-1)} \right] = \frac{2}{n^2(n-1)^2}.$$

Шаблонная функция есть сумма всех этих трех вкладов и, следовательно, равна $\frac{1}{(n-1)^2}$.

Так же можно найти, что вклады от таблиц (12.34) (f) и (g) в сумме дают $\frac{32}{(n-1)^2}$.

Итак, складывая вместе все полученные выражения, мы найдем, что коэффициент при $\kappa_6 \kappa_2^2$ в $\kappa(4^2 2)$ равен

$$\frac{72}{(n-1)(n-2)} + \frac{192}{(n-1)^2} + \frac{32}{(n-1)^2} = \frac{8(37n-65)}{(n-1)^2(n-2)},$$

что совпадает с результатом, даваемым ниже, в формуле (12.66).

12.14 Правило 10. Выражение для $\kappa(a_1^{a_1} \dots)$, где среди величин $a_i^{a_i}$ содержатся выражения вида 1^{a_i} , можно получить, зная представление $\kappa(a_1^{a_1} \dots)$, где величина 1^{a_i} опущена. Достаточно для формулировки этого правила рассмотреть случай $a_i = 1$. В этом случае для получения

$$\kappa(a_1^{a_1} \dots a_{i-1}^{a_{i-1}} 1 a_{i+1}^{a_{i+1}} \dots a_s^{a_s}) \text{ из } \kappa(a_1^{a_1} \dots a_{i-1}^{a_{i-1}} a_{i+1}^{a_{i+1}} \dots a_s^{a_s})$$

надо представление для последней величины умножить на $\frac{1}{n}$ и заменить всякий из членов в этом представлении на суммы членов, получающихся из исходного увеличением только одного индекса на единицу.

Например, можно показать, что

$$\kappa(2^2) = \frac{\kappa_4}{n} + \frac{2\kappa_2^2}{n-1},$$

откуда согласно сформулированному правилу

$$\kappa(2^2 1) = \frac{\kappa_5}{n^2} + \frac{4\kappa_3 \kappa_2}{n(n-1)},$$

$$\kappa(2^2 1^2) = \frac{\kappa_6}{n^3} + \frac{4\kappa_3^2}{n^2(n-1)} + \frac{4\kappa_4 \kappa_2}{n^2(n-1)} \text{ и т. д.}$$

12.15 Читателя нельзя упрекнуть за возможные сомнения в том, что развиваемая сложная комбинаторная процедура проще, чем прямой алгебраический подход, рассмотренный ранее в этой главе. Чтобы рассеять эти сомнения, мы рекомендуем читателю попробовать применить оба метода к каким-либо более сложным примерам. Легко, на самом деле, совершить ошибку при вычислениях, как пользуясь алгебраическими, так и комбинаторными методами. Именно поэтому все обычно используемые формулы были получены как тем, так и другим методом. Однако комбинаторные методы развиваются не только с целью сокращения алгебраических выкладок. В дальнейшем мы установим, что многие важные результаты можно сравнительно легко получить, пользуясь комбинаторными методами, в то время как алгебраические методы приводят к цели только после сложной и кропотливой работы.

Пример 12.3

Для сравнения найдем дисперсию статистики m_2 комбинаторными методами (эта дисперсия в примере 12.1 была найдена алгебраическим путем).

Из (12.29) имеем

$$k_2 = \frac{n}{n-1} m_2.$$

Следовательно, $Dm_2 = \left(\frac{n-1}{n}\right)^2 Dk_2 = \left(\frac{n-1}{n}\right)^2 \kappa(2^2)$, где $\kappa(2^2)$ состоит из суммы двух членов κ_4 и κ_2^2 с коэффициентами. Единственной таблицей, дающей коэффициент при κ_4 , является

$$\begin{array}{cc|c} 2 & 2 & 4 \\ \hline 2 & 2 & 4 \end{array}$$

Численный коэффициент здесь равен единице, а шаблонная функция равна $\frac{1}{n}$. Таблицы, служащие для подсчета коэффициента при κ_2^2 , имеют вид

$$\begin{array}{c|c} 2 & \\ \hline 2 & \\ \hline 2 & 2 & 4 \end{array}$$

Если на некотором месте в этой таблице будет стоять 2, то наверняка одна строка будет иметь нулевой элемент и, следовательно, такую таблицу не надо принимать в расчет. Поэтому единственной таблицей, дающей ненулевой вклад, будет

$$\begin{array}{cc|c} 1 & 1 & 2 \\ 1 & 1 & 2 \\ \hline 2 & 2 & 4 \end{array}$$

Численный коэффициент равен $\left(\frac{2!}{1!1!}\right)^2 \frac{1}{2!} = 2$. Шаблонная функция также легко может быть найдена, она равна $\frac{1}{n-1}$. Следовательно,

$$\kappa(2^2) = \frac{\kappa_4}{n} + \frac{2\kappa_2^2}{n-1}.$$

(см. также далее (12.35)). Поэтому

$$\begin{aligned} Dm_2 &= \left(\frac{n-1}{n}\right)^2 \kappa(2^2) = \left(\frac{n-1}{n}\right)^2 \left\{ \frac{1}{n} (\mu_4 - 3\mu_2^2) + \frac{2}{n-1} \mu_2^2 \right\} = \\ &= \frac{(n-1)^2}{n^3} \mu_4 + \frac{(3-n)(n-1)}{n^3} \mu_2^2, \end{aligned}$$

что согласуется с (12.14).

Пример 12.4

Чтобы найти третий момент статистики k^2 , мы должны рассмотреть $\kappa(2^3)$. С точностью до множителей $\kappa(2^3)$ равен сумме слагаемых κ_6 , $\kappa_4\kappa_2$, κ_3^2 и κ_2^3 .

Коэффициент при первом из них равен $1/n^2$. Для подсчета коэффициента при втором рассмотрим таблицу

1	1	2	4
1	1	.	2
			6
2	2	2	

Все остальные таблицы не надо принимать во внимание, за исключением двух таблиц, получающихся из выписанной подстановкой третьего столбца на место первого и второго. Подсчет показывает, что для них суммарный численный коэффициент равен

$$3 \cdot \left(\frac{2!}{1!1!1!} \right)^2 = 12.$$

Шаблонная функция равна $\frac{1}{n}$ от шаблонной функции таблицы

$$\begin{array}{cc} \times & \times \\ \times & \times \end{array}$$

т. е. равна $\frac{1}{n(n-1)}$. Итак, коэффициент при $\kappa_4\kappa_2$ равен $12/\{n(n-1)\}$.

Для подсчета коэффициента при κ_3^2 надо рассмотреть лишь таблицу

1	1	1	3
1	1	1	3
			6
2	2	2	

с численным коэффициентом $\left(\frac{2!}{1!1!1!} \right)^3 \frac{1}{2!} = 4$ и шаблонной функцией $(n-2)/n(n-1)^2$.

Для нахождения коэффициента при κ_2^3 мы должны рассмотреть таблицу

1	1	.	2
.	.	1	2
1	1	1	2
			6
2	2	2	

Подсчет показывает, что численный коэффициент равен 8, а шаблонная функция $\frac{1}{(n-1)^2}$.

Итак,

$$\kappa(2^3) = \frac{\kappa_6}{n^2} + \frac{12\kappa_4\kappa_2}{n(n-1)} + \frac{4(n-2)}{n(n-1)^2}\kappa_3^2 + \frac{8}{(n-1)^2}\kappa_2^3.$$

Нетрудно видеть, что $\kappa(2^3)$ равен третьему моменту $\mu(2^3)$, подсчитанному относительно среднего значения выборочного распределения. Тем самым мы видим, что если исходное распределение нормально, то третий момент сводится к $\frac{8\kappa_2^3}{(n-1)^2}$, т. е. имеет порядок n^{-2} , что показывает скорость стремления выборочного распределения к симметричной форме.

Пример 12.5

Полезность представления различных формул в терминах семиинвариантов исходного распределения и сила комбинаторных методов как нельзя лучше видны на примере нормального распределения. Единственным отличным от нуля семиинвариантом является κ_2 (считаем $\kappa_1=0$).

В качестве иллюстрации установим, что в нормальном случае $\kappa(pq)=0$, за исключением только случая $p=q$. Единственным членом, участвующим в представлении в $\kappa(pq)$, является $\frac{1}{2}(p+q)\kappa_2$, и поэтому очевидно, что если $p+q$ нечетно, то $\kappa(pq)=0$. В случае четных $p+q$ рассмотрим таблицу

	2
	2
	.
	.
	.
	2
pq	$p+q$

Если в ней имеется строка, содержащая 2, то такая таблица в расчет не принимается. Поэтому учитываться может только таблица, целиком состоящая из единиц, что возможно только в случае равенства p и q . Это и доказывает требуемый результат.

Пример 12.6

Если $\kappa(a_1^{\pi_1} \dots a_s^{\pi_s})$ содержит r частей, то каждый член в нем будет порядка $n^{-(r-1)}$. Например, каждый член в $\kappa(3^2 2^2)$ имеет порядок n^{-3} .

Согласно правилу 7 каждая таблица с ненулевой шаблонной функцией имеет по крайней мере два ненулевых входа в каждой строке. По правилу 8 мы можем ограничиться рассмотрением таблиц, имеющих по крайней мере два ненулевых входа в каждом столбце. Можно показать, что шаблонная функция каждой такой таблицы, состоящей из r столбцов, имеет порядок $n^{-(r-1)}$. Каждая таблица имеет разбиение, состоящее из одной группы; согласно правилу 6 соответствующий вклад равен $n \left(\frac{1}{n}\right)^r = n^{-(r-1)}$. Надо теперь только показать, что таблицы, обладающие более чем одним разбиением, дают вклад меньшего порядка. Рассмотрим какое-нибудь разбиение, состоящее из двух групп (разбиений); назовем его S_2 . Если соответствующая шаблонная функция не равна нулю, то согласно правилу 9 по меньшей мере два столбца должны иметь ненулевой вход в каждом разбиении, и согласно правилу 6 вклад S_2 по порядку равен $n^2 \left(\frac{1}{n^2}\right)^2 \left(\frac{1}{n}\right)^{r-2}$ или меньше, т. е. $o(n^{-(r-1)})$. Каждое разбиение, состоящее из трех групп, S_3 можно рассматривать как подразбиение S_2 , и согласно правилу 3 отсюда будет следовать, что вклад от S_3 порядка $n^3 \left(\frac{1}{n^2}\right)^3 \left(\frac{1}{n}\right)^{r-3}$, т. е. $o(n^{-(r-1)})$. Аналогичным способом доказывается, что и разбиения, состоящие из большего числа групп, дают вклад порядка $o(n^{-(r-1)})$. Этим заканчивается доказательство.

Пример 12.7

Используя свойства семиинвариантов выборочных распределений, можно доказать характеристическое свойство нормального распределения, установленное аналитически и геометрически в примерах 11.3 и 11.7, состоящее в том, что среднее и дисперсия выборки независимы.

Вспомним определение совместной х. ф. двух случайных величин. Они независимы в том и только том случае, когда их совместная х. ф. распадается (как в 4.16) в произведение двух одномерных х. ф. Рассматривая эти три х. ф. как производящие функции для моментов и беря логарифмы, в силу (3.74) получим

$$\sum_{r, s=1}^{\infty} \kappa_{rs} \frac{(it_1)^r}{r!} \frac{(it_2)^s}{s!} = \sum_{r=1}^{\infty} \kappa_{r0} \frac{(it_1)^r}{r!} + \sum_{s=1}^{\infty} \kappa_{0s} \frac{(it_2)^s}{s!},$$

откуда следует, что $\kappa_{rs}=0$, если $rs \neq 0$. В предположении, что семинварианты всех порядков существуют, это условие является необходимым и достаточным для независимости.

Рассмотрим, наконец, совместное распределение выборочных среднего и дисперсии, т. е. k_1 и k_2 . В силу того, что все исходные семинварианты порядка выше второго обращаются в нуль, мы получим согласно правилу 10

$$\kappa(2^s 1^r) = 0, \quad r, s \neq 0.$$

Следовательно, k_1 и k_2 независимы.

Справедлив и обратный результат, установленный Гири (1936): если k_1 и k_2 независимы, то исходное распределение нормально. В самом деле, согласно правилу 10

$$\kappa(21^r) = \frac{1}{n^r} \kappa_{r+2}, \quad r > 0.$$

Если k_1 и k_2 независимы, то эти семинварианты равны нулю для всех $r > 0$, что и означает нормальность исходного распределения.

Следует отметить, что здесь весьма замечательны два обстоятельства:

(1) нам не понадобилось использовать более общее требование $\kappa(2^s 1^r) = 0$ при $s > 1$;

(2) нормальность исходного распределения следует из независимости k_1 и k_2 при одном каком-либо значении n .

Позднее рядом авторов (см., например, Лукач (1942)) была доказана теорема Гири при менее стеснительных условиях, чем существование всех семинвариантов.

12.16 Следующие формулы дают семинварианты статистик до десятого порядка и некоторых статистик 12-го порядка. Эти формулы справедливы для всех n , не меньших чем порядок рассматриваемых k -статистик.

k -статистики второго порядка

$$\kappa(2^2) = \frac{\kappa_4}{n} + \frac{2\kappa_2^2}{n-1}, \quad (12.35)$$

$$\kappa(2^3) = \frac{\kappa_6}{n^2} + \frac{12\kappa_4\kappa_2}{n(n-1)} + \frac{4(n-2)}{n(n-1)^2} \kappa_3^2 + \frac{8}{(n-1)^2} \kappa_2^3, \quad (12.36)$$

$$\begin{aligned} \kappa(2^4) = & \frac{\kappa_8}{n^3} + \frac{24}{n^2(n-1)} \kappa_6\kappa_2 + \frac{32(n-2)}{n^2(n-1)^2} \kappa_5\kappa_3 + \\ & + \frac{8(4n^2-9n+6)}{n^2(n-1)^3} \kappa_4^2 + \frac{144}{n(n-1)} \kappa_4\kappa_2^2 + \frac{96(n-2)}{n(n-1)^3} \kappa_3^2\kappa_2 + \frac{48}{(n-1)^3} \kappa_2^4, \end{aligned} \quad (12.37)$$

$$\begin{aligned}
\kappa(2^5) = & \frac{\kappa_{10}}{n^4} + \frac{40\kappa_8\kappa_2}{n^3(n-1)} + \frac{80(n-2)}{n^3(n-1)^2} \kappa_7\kappa_3 + \frac{40(5n^2-12n+9)}{n^3(n-1)^3} \kappa_6\kappa_4 + \\
& + \frac{16(n-2)(6n^2-12n+7)}{n^3(n-1)^4} \kappa_5^2 + \frac{480}{n^2(n-1)^2} \kappa_6\kappa_2^2 + \\
& + \frac{1280(n-2)}{n^2(n-1)^3} \kappa_5\kappa_3\kappa_2 + \frac{320(4n^2-9n+6)}{n^2(n-1)^4} \kappa_4^2\kappa_2 + \\
& + \frac{480(2n^2-7n+6)}{n^2(n-1)^4} \kappa_4\kappa_3^2 + \frac{1920}{n(n-1)^3} \kappa_4\kappa_2^3 + \\
& + \frac{1920(n-2)}{n(n-1)^4} \kappa_3^2\kappa_2^2 + \frac{384}{(n-1)^4} \kappa_2^5,
\end{aligned} \quad (12.38)$$

$$\begin{aligned}
\kappa(2^6) = & \frac{1}{n^5} \kappa_{12} + \frac{60}{n^4(n-1)} \kappa_{10}\kappa_2 + \frac{160(n-2)}{n^4(n-1)^2} \kappa_9\kappa_3 + \\
& + \frac{240(2n^2-5n+4)}{n^4(n-1)^3} \kappa_8\kappa_4 + \frac{96(n-2)(7n^2-14n+9)}{n^4(n-1)^4} \kappa_7\kappa_5 + \\
& + \frac{4(113n^4-520n^3+950n^2-800n+265)}{n^4(n-1)^5} \kappa_6^2 + \\
& + \frac{1200}{n^3(n-1)^2} \kappa_8\kappa_2^2 + \frac{4800(n-2)}{n^3(n-1)^3} \kappa_7\kappa_3\kappa_2 + \\
& + \frac{2400(5n^2-12n+9)}{n^3(n-1)^4} \kappa_6\kappa_4\kappa_2 + \frac{160(n-2)(31n-53)}{n^3(n-1)^4} \kappa_6\kappa_3^2 + \\
& + \frac{960(n-2)(6n^2-12n+7)}{n^3(n-1)^5} \kappa_5^2\kappa_2 + \\
& + \frac{1920(n-2)(9n^2-23n+16)}{n^3(n-1)^5} \kappa_5\kappa_4\kappa_3 + \frac{480(11n^3-41n^2+59n-31)}{n^3(n-1)^5} \kappa_4^3 + \\
& + \frac{9600}{n^2(n-1)} \kappa_6\kappa_2^3 + \frac{38400(n-2)}{n^2(n-1)^4} \kappa_5\kappa_3\kappa_2^2 + \\
& + \frac{9600(4n^2-9n+6)}{n^2(n-1)^5} \kappa_4^2\kappa_3^2 + \frac{28800(2n^2-7n+6)}{n^2(n-1)^5} \kappa_4\kappa_3^2\kappa_2 + \\
& + \frac{960(n-2)(5n-12)}{n^2(n-1)^5} \kappa_4^3 + \frac{28800}{n(n-1)^4} \kappa_4\kappa_2^4 + \\
& + \frac{38400(n-2)}{n(n-1)^5} \kappa_3^2\kappa_2^3 + \frac{3840}{(n-1)^5} \kappa_2^6.
\end{aligned} \quad (12.39)$$

k -статистики третьего порядка

$$\kappa(3^2) = \frac{1}{n} \kappa_6 + \frac{9}{n-1} \kappa_4\kappa_2 + \frac{9}{n-1} \kappa_3^2 + \frac{6n}{(n-1)(n-2)} \kappa_2^3, \quad (12.40)$$

$$\begin{aligned}
\kappa(3^3) = & \frac{1}{n^2} \kappa_9 + \frac{27}{n(n-1)} \kappa_7\kappa_2 + \frac{27(3n-4)}{n(n-1)^2} \kappa_6\kappa_3 + \\
& + \frac{27(4n-7)}{n(n-1)^2} \kappa_5\kappa_4 + \frac{54(4n-7)}{(n-1)^2(n-2)} \kappa_5\kappa_2^2 +
\end{aligned}$$

$$\begin{aligned}
& + \frac{162(5n-12)}{(n-1)^2(n-2)} \kappa_4 \kappa_3 \kappa_2 + \frac{36(7n^2-30n+34)}{(n-1)^2(n-2)^2} \kappa_3^3 + \\
& + \frac{108n(5n-12)}{(n-1)^2(n-2)^2} \kappa_3 \kappa_2^3,
\end{aligned} \tag{12.41}$$

$$\begin{aligned}
\kappa(3^4) = & \frac{1}{n^3} \kappa_{12} + \frac{54}{n^2(n-1)} \kappa_{10} \kappa_2 + \frac{108(2n-3)}{n^2(n-1)^2} \kappa_9 \kappa_3 + \\
& + \frac{27(17n^2-49n+35)}{n^2(n-1)^3} \kappa_8 \kappa_4 + \frac{108(7n^2-20n+16)}{n^2(n-1)^3} \kappa_7 \kappa_5 + \\
& + \frac{27(17n^2-47n+39)}{n^2(n-1)^3} \kappa_6^2 + \frac{27(37n-70)}{n(n-1)^2(n-2)} \kappa_8 \kappa_2^2 + \\
& + \frac{324(19n^2-67n+54)}{n(n-1)^3(n-2)} \kappa_7 \kappa_3 \kappa_2 + \frac{162(65n^2-245n+234)}{n(n-1)^3(n-2)} \kappa_6 \kappa_4 \kappa_2 + \\
& + \frac{108(82n^3-481n^2+958n-640)}{n(n-1)^3(n-2)^2} \kappa_6 \kappa_3^2 + \\
& + \frac{108(59n^2-220n+224)}{n(n-1)^3(n-2)} \kappa_5^2 \kappa_2 + \\
& + \frac{324(75n^3-473n^2+1016n-756)}{n(n-1)^3(n-2)^2} \kappa_5 \kappa_4 \kappa_3 + \\
& + \frac{27(173n^4-1503n^3+4962n^2-7380n+4200)}{n(n-1)^3(n-2)^3} \kappa_4^3 + \\
& + \frac{108(71n^2-263n+234)}{(n-1)^3(n-2)^2} \kappa_6 \kappa_2^3 + \frac{648(79n^2-343n+378)}{(n-1)^3(n-2)^2} \kappa_5 \kappa_3 \kappa_2^2 + \\
& + \frac{486(63n^2-290n+352)}{(n-1)^3(n-2)^2} \kappa_4^2 \kappa_2^2 + \\
& + \frac{972(99n^3-688n^2+1612n-1280)}{(n-1)^3(n-2)^3} \kappa_4 \kappa_3^2 \kappa_2 + \\
& + \frac{162(87n^3-594n^2+1420n-1176)}{(n-1)^3(n-2)^3} \kappa_3^4 + \\
& + \frac{972n(23n^2-103n+118)}{(n-1)^3(n-2)^3} \kappa_4 \kappa_2^4 + \\
& + \frac{648n(103n^2-510n+640)}{(n-1)^3(n-2)^3} \kappa_3^2 \kappa_2^3 + \frac{648n^2(5n-12)}{(n-1)^3(n-2)^3} \kappa_2^6.
\end{aligned} \tag{12.42}$$

k -статистики четвертого порядка

$$\begin{aligned}
\kappa(4^2) = & \frac{1}{n} \kappa_8 + \frac{16}{n-1} \kappa_6 \kappa_2 + \frac{48}{n-1} \kappa_5 \kappa_3 + \frac{34}{n-1} \kappa_4^2 + \\
& + \frac{72n}{(n-1)(n-2)} \kappa_4 \kappa_2^2 + \frac{144n}{(n-1)(n-2)} \kappa_3^2 \kappa_2 + \\
& + \frac{24n(n+1)}{(n-1)(n-2)(n-3)} \kappa_2^4,
\end{aligned} \tag{12.43}$$

$$\begin{aligned}
\kappa(4^3) = & \frac{1}{n^2} \kappa_{12} + \frac{48}{n(n-1)} \kappa_{10} \kappa_2 + \frac{16(13n-17)}{n(n-1)^2} \kappa_9 \kappa_3 + \\
& + \frac{12(41n-65)}{n(n-1)^2} \kappa_8 \kappa_4 + \frac{48(16n-29)}{n(n-1)^2} \kappa_7 \kappa_5 + \\
& + \frac{12(37n-70)}{n(n-1)^2} \kappa_6^2 + \frac{72(11n-19)}{(n-1)^2(n-2)} \kappa_8 \kappa_2^2 + \\
& + \frac{288(19n-41)}{(n-1)^2(n-2)} \kappa_7 \kappa_3 \kappa_2 + \frac{48(203n-523)}{(n-1)^2(n-2)} \kappa_6 \kappa_4 \kappa_2 + \\
& + \frac{144(56n^2-257n+302)}{(n-1)^2(n-2)^2} \kappa_6 \kappa_3^2 + \frac{1440(4n-11)}{(n-1)^2(n-2)} \kappa_5^2 \kappa_2 + \\
& + \frac{1152(22n^2-106n+133)}{(n-1)^2(n-2)^2} \kappa_5 \kappa_4 \kappa_3 + \frac{8(709n^2-3430n+4456)}{(n-1)^2(n-2)^2} \kappa_4^3 + \\
& + \frac{288(19n^3-98n^2+125n+2)}{(n-1)^2(n-2)^2(n-3)} \kappa_6 \kappa_2^3 + \\
& + \frac{1728(24n^3-140n^2+200n+4)}{(n-1)^2(n-2)^2(n-3)} \kappa_5 \kappa_3 \kappa_2^2 + \\
& + \frac{432(61n^3-371n^2+552n+12)}{(n-1)^2(n-2)^2(n-3)} \kappa_4^2 \kappa_2^2 + \\
& + \frac{864(103n^3-629n^2+948n+24)}{(n-1)^2(n-2)^2(n-3)} \kappa_4 \kappa_3^2 \kappa_2 + \\
& + \frac{288(41n^4-384n^3+1209n^2-1282n-36)}{(n-1)^2(n-2)^2(n-3)^2} \kappa_3^4 + \\
& + \frac{288n(53n^2-179n-52)}{(n-1)^2(n-2)^2(n-3)} \kappa_4 \kappa_2^4 + \\
& + \frac{1728n(29n^3-196n^2+317n+62)}{(n-1)^2(n-2)^2(n-3)^2} \kappa_3^2 \kappa_2^3 + \\
& + \frac{1728n(n+1)(n^2-5n+2)}{(n-1)^2(n-2)^2(n-3)^2} \kappa_2^6.
\end{aligned} \tag{12.44}$$

k -статистики пятого порядка

$$\begin{aligned}
\kappa(5^2) = & \frac{1}{n} \kappa_{10} + \frac{25}{n-1} \kappa_8 \kappa_2 + \frac{100}{n-1} \kappa_7 \kappa_3 + \\
& + \frac{200}{n-1} \kappa_6 \kappa_4 + \frac{125}{n-1} \kappa_5^2 + \frac{200n}{(n-1)(n-2)} \kappa_6 \kappa_2^2 + \\
& + \frac{1200n}{(n-1)(n-2)} \kappa_5 \kappa_3 \kappa_2 + \frac{850}{(n-1)(n-2)} \kappa_4^2 \kappa_2 + \\
& + \frac{1500n}{(n-1)(n-2)} \kappa_4 \kappa_3^2 + \frac{600(n+1)}{(n-1)(n-2)(n-3)} \kappa_4 \kappa_2^3 + \\
& + \frac{1800n(n+1)}{(n-1)(n-2)(n-3)} \kappa_3^2 \kappa_2^2 + \frac{120n^2(n+5)}{(n-1)(n-2)(n-3)(n-4)} \kappa_2^5.
\end{aligned} \tag{12.45}$$

k-статистики шестого порядка

$$\begin{aligned}
\kappa(6^2) = & \frac{1}{n} \kappa_{12} + \frac{1}{n-1} (36\kappa_{10}\kappa_2 + 180\kappa_9\kappa_3 + 465\kappa_8\kappa_4 + \\
& + 780\kappa_7\kappa_5 + 461\kappa_6^2) + \frac{n}{(n-1)(n-2)} (450\kappa_8\kappa_2^2 + \\
& + 3600\kappa_7\kappa_3\kappa_2 + 7200\kappa_6\kappa_4\kappa_2 + 6300\kappa_6\kappa_3^2 + \\
& + 4500\kappa_5^2\kappa_2 + 21\,600\kappa_5\kappa_4\kappa_3 + 4950\kappa_4^3) + \\
& + \frac{n(n+1)}{(n-1)(n-2)(n-3)} (2400\kappa_6\kappa_2^3 + 21\,600\kappa_5\kappa_3\kappa_2^2 + \\
& + 15\,300\kappa_4^2\kappa_2^2 + 54\,000\kappa_4\kappa_3^2\kappa_2 + 8100\kappa_3^4) + \\
& + \frac{n^2(n+5)}{(n-1)(n-2)(n-3)(n-4)} (5400\kappa_4\kappa_2^4 + 21\,600\kappa_3^2\kappa_2^3) + \\
& + \frac{n(n+1)(n^2+15n-4)}{(n-1)(n-2)(n-3)(n-4)(n-5)} 720\kappa_2^6.
\end{aligned} \tag{12.46}$$

Формулы для смешанных семиинвариантов:

$$\kappa(32) = \frac{1}{n} \kappa_5 + \frac{6}{n-1} \kappa_3\kappa_2, \tag{12.47}$$

$$\kappa(42) = \frac{1}{n} \kappa_6 + \frac{8}{n-1} \kappa_4\kappa_2 + \frac{6}{n-1} \kappa_3^2, \tag{12.48}$$

$$\kappa(52) = \frac{1}{n} \kappa_7 + \frac{10}{n-1} \kappa_5\kappa_2 + \frac{20}{n-1} \kappa_4\kappa_3, \tag{12.49}$$

$$\kappa(62) = \frac{1}{n} \kappa_8 + \frac{12}{n-1} \kappa_6\kappa_2 + \frac{30}{n-1} \kappa_5\kappa_3 + \frac{20}{n-1} \kappa_4^2, \tag{12.50}$$

$$\kappa(72) = \frac{1}{n} \kappa_9 + \frac{14}{n-1} \kappa_4\kappa_2 + \frac{42}{n-1} \kappa_6\kappa_3 + \frac{70}{n-1} \kappa_5\kappa_4, \tag{12.51}$$

$$\begin{aligned}
\kappa(82) = & \frac{1}{n} \kappa_{10} + \frac{16}{n-1} \kappa_8\kappa_2 + \frac{56}{n-1} \kappa_7\kappa_3 + \frac{112}{n-1} \kappa_6\kappa_4 + \frac{70}{n-1} \kappa_3^2, \\
& \tag{12.52}
\end{aligned}$$

$$\begin{aligned}
\kappa(43) = & \frac{1}{n} \kappa_7 + \frac{12}{n-1} \kappa_5\kappa_2 + \frac{30}{n-1} \kappa_4\kappa_3 + \frac{36}{(n-1)(n-2)} \kappa_3\kappa_2^2, \\
& \tag{12.53}
\end{aligned}$$

$$\begin{aligned}
\kappa(53) = & \frac{1}{n} \kappa_8 + \frac{1}{n-1} (15\kappa_6\kappa_2 + 45\kappa_5\kappa_3 + 30\kappa_4^2) + \\
& + \frac{n}{(n-1)(n-2)} (60\kappa_4\kappa_2^2 + 90\kappa_3^2\kappa_2),
\end{aligned} \tag{12.54}$$

$$\begin{aligned}
\kappa(63) = & \frac{1}{n} \kappa_9 + \frac{1}{n-1} (18\kappa_7\kappa_2 + 63\kappa_6\kappa_3 + 105\kappa_4\kappa_4) + \\
& + \frac{n}{(n-1)(n-2)} (90\kappa_5\kappa_2^2 + 360\kappa_4\kappa_3\kappa_2 + 90\kappa_3^3),
\end{aligned} \tag{12.55}$$

$$\begin{aligned} \kappa(73) = & \frac{1}{n} \kappa_{10} + \frac{1}{n-1} (21\kappa_8\kappa_2 + 84\kappa_7\kappa_3 + 168\kappa_6\kappa_4 + 105\kappa_5^2) + \\ & + \frac{n}{(n-1)(n-2)} (126\kappa_6\kappa_2^2 + 630\kappa_5\kappa_3\kappa_2 + 420\kappa_4^2\kappa_2 + 630\kappa_4\kappa_3^2), \end{aligned} \quad (12.56)$$

$$\begin{aligned} \kappa(54) = & \frac{1}{n} \kappa_9 + \frac{1}{n-1} (20\kappa_7\kappa_2 + 70\kappa_6\kappa_3 + 120\kappa_5\kappa_4) + \\ & + \frac{n}{(n-1)(n-2)} (120\kappa_5\kappa_2^2 + 600\kappa_4\kappa_3\kappa_2 + 180\kappa_3^3) + \\ & + \frac{n(n+1)}{(n-1)(n-2)(n-3)} 240\kappa_3\kappa_2^3, \end{aligned} \quad (12.57)$$

$$\begin{aligned} \kappa(64) = & \frac{1}{n} \kappa_{10} + \frac{1}{n-1} (24\kappa_8\kappa_2 + 96\kappa_7\kappa_3 + 194\kappa_6\kappa_4 + 120\kappa_5^2) + \\ & + \frac{n}{(n-1)(n-2)} (180\kappa_6\kappa_2^2 + 1080\kappa_5\kappa_3\kappa_2 + 720\kappa_4^2\kappa_2 + 1260\kappa_4\kappa_3^2) + \\ & + \frac{n(n+1)}{(n-1)(n-2)(n-3)} (480\kappa_4\kappa_2^3 + 1080\kappa_3^2\kappa_2^2), \end{aligned} \quad (12.58)$$

$$\kappa(32^2) = \frac{1}{n^2} \kappa_7 + \frac{16}{n(n-1)} \kappa_5\kappa_2 + \frac{12(2n-3)}{n(n-1)^2} \kappa_4\kappa_3 + \frac{48}{(n-1)^2} \kappa_2^2\kappa_3, \quad (12.59)$$

$$\begin{aligned} \kappa(42^2) = & \frac{1}{n^2} \kappa_8 + \frac{20}{n(n-1)} \kappa_6\kappa_2 + \frac{8(5n-7)}{n(n-1)^2} \kappa_5\kappa_3 + \frac{4(7n-10)}{n(n-1)^2} \kappa_4^2 + \\ & + \frac{80}{(n-1)^2} \kappa_4\kappa_2^2 + \frac{120}{(n-1)^2} \kappa_3^2\kappa_2, \end{aligned} \quad (12.60)$$

$$\begin{aligned} \kappa(52^2) = & \frac{1}{n^2} \kappa_9 + \frac{24}{n(n-1)} \kappa_7\kappa_2 + \frac{20(3n-4)}{n(n-1)^2} \kappa_6\kappa_3 + \\ & + \frac{20(5n-7)}{n(n-1)^2} \kappa_5\kappa_4 + \frac{120}{(n-1)^2} \kappa_5\kappa_2^2 + \frac{480}{(n-1)^2} \kappa_4\kappa_3\kappa_2 + \\ & + \frac{120}{(n-1)^2} \kappa_3^3, \end{aligned} \quad (12.61)$$

$$\begin{aligned} \kappa(62^2) = & \frac{1}{n^2} \kappa_{10} + \frac{28}{n(n-1)} \kappa_8\kappa_2 + \frac{12(7n-9)}{n(n-1)^2} \kappa_7\kappa_3 + \\ & + \frac{4(41n-56)}{n(n-1)^2} \kappa_6\kappa_4 + \frac{20(5n-7)}{n(n-1)^2} \kappa_5^2 + \frac{168}{(n-1)^2} \kappa_6\kappa_2^2 + \\ & + \frac{840}{(n-1)^2} \kappa_5\kappa_3\kappa_2 + \frac{560}{(n-1)^2} \kappa_4^2\kappa_2 + \frac{840}{(n-1)^2} \kappa_4\kappa_3^2, \end{aligned} \quad (12.62)$$

$$\begin{aligned} \kappa(3^22) = & \frac{1}{n^2} \kappa_8 + \frac{21}{n(n-1)} \kappa_6\kappa_2 + \frac{6(8n-11)}{n(n-1)^2} \kappa_5\kappa_3 + \frac{9(3n-5)}{n(n-1)^2} \kappa_4^2 + \\ & + \frac{18(6n-11)}{(n-1)^2(n-2)} \kappa_4\kappa_2^2 + \frac{18(9n-20)}{(n-1)^2(n-2)} \kappa_3^2\kappa_2 + \frac{36n}{(n-1)^2(n-2)} \kappa_4^3, \end{aligned} \quad (12.63)$$

$$\begin{aligned}
\kappa(432) = & \frac{1}{n^2} \kappa_9 + \frac{26}{n(n-1)} \kappa_7 \kappa_2 + \frac{24(3n-4)}{n(n-1)^2} \kappa_6 \kappa_3 + \\
& + \frac{10(11n-17)}{n(n-1)^2} \kappa_5 \kappa_4 + \frac{36(5n-9)}{(n-1)^2(n-2)} \kappa_5 \kappa_2^2 + \\
& + \frac{12(61n-128)}{(n-1)^2(n-2)} \kappa_4 \kappa_3 \kappa_2 + \frac{36(5n-12)}{(n-1)^2(n-2)} \kappa_3^3 + \\
& + \frac{360n}{(n-1)^2(n-2)} \kappa_3 \kappa_2^3,
\end{aligned} \tag{12.64}$$

$$\begin{aligned}
\kappa(532) = & \frac{1}{n^2} \kappa_{10} + \frac{31}{n(n-1)} \kappa_8 \kappa_2 + \frac{101n-131}{n(n-1)^2} \kappa_7 \kappa_3 + \\
& + \frac{5(37n-55)}{n(n-1)^2} \kappa_6 \kappa_5 + \frac{5(23n-35)}{n(n-1)^2} \kappa_5^2 + \\
& + \frac{30(9n-16)}{(n-1)^2(n-2)} \kappa_6 \kappa_2^2 + \frac{30(45n-92)}{(n-1)^2(n-2)} \kappa_5 \kappa_3 \kappa_2 + \\
& + \frac{60(15n-31)}{(n-1)^2(n-2)} \kappa_4^2 \kappa_2 + \frac{30(45n-103)}{(n-1)^2(n-2)} \kappa_4 \kappa_3^2 + \\
& + \frac{720n}{(n-1)^2(n-2)} \kappa_4 \kappa_2^3 + \frac{1620n}{(n-1)^2(n-2)} \kappa_3^2 \kappa_2^2,
\end{aligned} \tag{12.65}$$

$$\begin{aligned}
\kappa(4^22) = & \frac{1}{n^2} \kappa_{10} + \frac{32}{n(n-1)} \kappa_8 \kappa_2 + \frac{8(13n-37)}{n(n-1)^2} \kappa_7 \kappa_3 + \\
& + \frac{4(49n-73)}{n(n-1)^2} \kappa_6 \kappa_1 + \frac{4(29n-46)}{n(n-1)^2} \kappa_5^2 + \frac{8(37n-65)}{(n-1)^2(n-2)} \kappa_6 \kappa_2^2 + \\
& + \frac{1536}{(n-1)^2} \kappa_5 \kappa_3 \kappa_2 + \frac{144(7n-15)}{(n-1)^2(n-2)} \kappa_4^2 \kappa_2 + \\
& + \frac{72(21n-50)}{(n-1)^2(n-2)} \kappa_4 \kappa_3^2 + \frac{96(10n^2-27n-1)}{(n-1)^2(n-2)(n-3)} \kappa_4 \kappa_2^3 + \\
& + \frac{144(17n^2-53n-2)}{(n-1)^2(n-2)(n-3)} \kappa_3^2 \kappa_2^2 + \frac{192n(n+1)}{(n-1)^2(n-2)(n-3)} \kappa_2^5,
\end{aligned} \tag{12.66}$$

$$\begin{aligned}
\kappa(43^2) = & \frac{1}{n^2} \kappa_{10} + \frac{33}{n(n-1)} \kappa_8 \kappa_2 + \frac{6(19n-25)}{n(n-1)^2} \kappa_7 \kappa_3 + \\
& + \frac{3(65n-107)}{n(n-1)^2} \kappa_6 \kappa_4 + \frac{6(19n-34)}{n(n-1)^2} \kappa_5^2 + \\
& + \frac{18(19n-33)}{(n-1)^2(n-2)} \kappa_6 \kappa_2^2 + \frac{72(23n-52)}{(n-1)^2(n-2)} \kappa_5 \kappa_3 \kappa_2 + \\
& + \frac{54(19n-48)}{(n-1)^2(n-2)} \kappa_4^2 \kappa_2 + \frac{54(33n^2-148n+172)}{(n-1)^2(n-2)^2} \kappa_4 \kappa_3^2 + \\
& + \frac{72n(17n-40)}{(n-1)^2(n-2)^2} \kappa_4 \kappa_2^3 + \frac{108n(27n-70)}{(n-1)^2(n-2)^2} \kappa_3^2 \kappa_2^2 + \\
& + \frac{216n^2}{(n-1)^2(n-2)^2} \kappa_2^5,
\end{aligned} \tag{12.67}$$

$$\begin{aligned} \kappa(3^2 3) = & \frac{1}{n^3} \kappa_9 + \frac{30}{n^2(n-1)} \kappa_7 \kappa_2 + \frac{2(31n-53)}{n^2(n-1)^2} \kappa_6 \kappa_3 + \\ & + \frac{12(9n^2-23n+16)}{n^2(n-1)^3} \kappa_5 \kappa_4 + \frac{240}{n(n-1)^2} \kappa_5 \kappa_2^2 + \\ & + \frac{360(2n-3)}{n(n-1)^3} \kappa_4 \kappa_3 \kappa_2 + \frac{24(5n-12)}{n(n-1)^3} \kappa_3^3 + \frac{480}{(n-1)^3} \kappa_3 \kappa_2^3, \end{aligned} \quad (12.68)$$

$$\begin{aligned} \kappa(42^3) = & \frac{1}{n^3} \kappa_{10} + \frac{36}{n^2(n-1)} \kappa_8 \kappa_2 + \frac{4(23n-37)}{n^2(n-1)^2} \kappa_7 \kappa_3 + \\ & + \frac{4(47n^2-120n+81)}{n^2(n-1)^3} \kappa_6 \kappa_4 + \frac{12(9n^2-24n+17)}{n^2(n-1)^3} \kappa_5^2 + \\ & + \frac{360}{n(n-1)^2} \kappa_6 \kappa_2^2 + \frac{288(5n-7)}{n(n-1)^3} \kappa_5 \kappa_3 \kappa_2 + \\ & + \frac{144(7n-10)}{n(n-1)^3} \kappa_4^2 \kappa_2 + \frac{24(49n-95)}{n(n-1)^3} \kappa_4 \kappa_3^2 + \\ & + \frac{960}{(n-1)^3} \kappa_4 \kappa_2^3 + \frac{2160}{(n-1)^3} \kappa_3^2 \kappa_2^2, \end{aligned} \quad (12.69)$$

$$\begin{aligned} \kappa(3^2 2^2) = & \frac{1}{n^3} \kappa_{10} + \frac{37}{n^2(n-1)} \kappa_8 \kappa_2 + \frac{6(17n-27)}{n^2(n-1)^2} \kappa_7 \kappa_3 + \\ & + \frac{3(61n^2-166n+117)}{n^2(n-1)^3} \kappa_6 \kappa_4 + \frac{2(59n^2-154n+113)}{n^2(n-1)^3} \kappa_5^2 + \\ & + \frac{6(67n-131)}{n(n-1)^2(n-2)} \kappa_6 \kappa_2^2 + \frac{24(71n^2-246n+202)}{n(n-1)^3(n-2)} \kappa_5 \kappa_3 \kappa_2 + \\ & + \frac{36(29n^2-103n+93)}{n(n-1)^3(n-2)} \kappa_4^2 \kappa_2 + \\ & + \frac{36(38n^2-155n+160)}{n(n-1)^3(n-2)} \kappa_4 \kappa_3^2 + \frac{72(14n-23)}{(n-1)^3(n-2)} \kappa_4 \kappa_2^3 + \\ & + \frac{144(19n-44)}{(n-1)^3(n-2)} \kappa_3^2 \kappa_2^2 + \frac{288n}{(n-1)^3(n-2)} \kappa_2^5. \end{aligned} \quad (12.70)$$

12.17 Приведем дополнительно формулы для случая нормального распределения.

Общие формулы:

$$\kappa(2^r) = \frac{2^{r-1}(r-1)!}{(n-1)^{r-1}} \kappa_2^r, \quad (12.71)$$

$$\kappa(p^q 2^r) = \frac{2^r \left(r + \frac{1}{2} pq - 1\right)!}{\left(\frac{1}{2} pq - 1\right)! (n-1)^r} \kappa_2^r \kappa(p^q), \quad (12.72)$$

и специальные формулы 12-й степени и выше (для степеней, меньших десятых, соответствующие формулы можно получить

из (12.34) — (12.70), полагая $\kappa_r = 0$, $r > 2$):

$$\kappa(3^2 2^4) = \frac{34\,560n}{(n-1)^5 (n-2)} \kappa_2^7, \quad (12.73)$$

$$\kappa(3^4 2) = \frac{7776n^2 (5n-12)}{(n-1)^4 (n-2)^3} \kappa_2^7, \quad (12.74)$$

$$\kappa(3^4 2^2) = \frac{108\,864n^2 (5n-12)}{(n-1)^5 (n-2)^3} \kappa_2^8, \quad (12.75)$$

$$\kappa(3^4 2^3) = \frac{1\,741\,824n^2 (5n-12)}{(n-1)^6 (n-2)^3} \kappa_2^9, \quad (12.76)$$

$$\kappa(3^6) = \frac{466\,560n^3 (22n^2 - 111n + 142)}{(n-1)^5 (n-2)^5} \kappa_2^9, \quad (12.77)$$

$$\kappa(3^6 2) = \frac{18}{n-1} \kappa_2 \kappa(3^6), \quad (12.78)$$

$$\kappa(3^6 2^2) = \frac{360}{(n-1)^2} \kappa_2^2 \kappa(3^6), \quad (12.79)$$

$$\kappa(4^2 2^2) = \frac{1920n (n+1)}{(n-1)^3 (n-2) (n-3)} \kappa_2^6, \quad (12.80)$$

$$\kappa(4^2 2^3) = \frac{23\,040n (n+1)}{(n-1)^4 (n-2) (n-3)} \kappa_2^7, \quad (12.81)$$

$$\kappa(4^2 2^4) = \frac{322\,560n (n+1)}{(n-1)^5 (n-2) (n-3)} \kappa_2^8, \quad (12.82)$$

$$\kappa(4^3 2) = \frac{20\,736n (n+1) (n^2 - 5n + 2)}{(n-1)^3 (n-2)^2 (n-3)^2} \kappa_2^7, \quad (12.83)$$

$$\kappa(4^3 2^2) = \frac{290\,304n (n+1) (n^2 - 5n + 2)}{(n-1)^4 (n-2)^2 (n-3)^2} \kappa_2^8, \quad (12.84)$$

$$\kappa(4^3 2^3) = \frac{4\,644\,864n (n+1) (n^2 - 5n + 2)}{(n-1)^5 (n-2)^2 (n-3)^2} \kappa_2^9, \quad (12.85)$$

$$\begin{aligned} \kappa(4^4) = & \frac{6912n (n+1)}{(n-1)^3 (n-2)^3 (n-3)^3} \{53n^4 - 428n^3 + \\ & + 1025n^2 - 474n + 180\} \kappa_2^8, \end{aligned} \quad (12.86)$$

$$\kappa(4^4 2) = \frac{16}{n-1} \kappa_2 \kappa(4^4), \quad (12.87)$$

$$\kappa(4^4 2^2) = \frac{288}{(n-1)^2} \kappa_2^2 \kappa(4^4), \quad (12.88)$$

$$\kappa(4^5) = \frac{484 \cdot 12^5}{n^4} \kappa_2^{10} \text{ (приближенно)}. \quad (12.89)$$

Нужно заметить, что в формулах (12.73) — (12.89) даны выражения только для четных степеней. Для нечетных степеней соответствующие выражения обращаются в нуль, поскольку в нормальном случае $\kappa_r = 0$, $r > 2$. Действительно, из (12.72) сле-

дует, что $\kappa(p^2 2^r) = 0$, если pq нечетно или если $q=1$ и $p>2$. Методы доказательства формул содержатся в упражнениях 12.9 и 12.10. Сюй и Лоли (1939) дали точные формулы для $\kappa(4^5)$ и $\kappa(4^6)$.

12.18 Чтобы показать, как приведенные выше формулы можно использовать для аппроксимации выборочных распределений, рассмотрим распределение статистики $\sqrt{b_1}$ для случая нормальной совокупности. Мы имеем

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}} = \frac{n-2}{\sqrt{n(n-1)}} \frac{k_3}{k_2^{3/2}}.$$

Для нормального распределения дисперсия статистики k_3 , $\kappa(3^2)$ согласно (12.40) равна

$$\frac{6n\kappa_2^3}{(n-1)(n-2)}.$$

Рассмотрим поэтому статистику

$$x = \sqrt{\frac{(n-1)(n-2)}{6n}} k_3 k_2^{-3/2} = \quad (12.90)$$

$$= \frac{(n-1)}{\sqrt{6(n-2)}} \sqrt{b_1}, \quad (12.91)$$

которая с точностью до величин порядка n^{-1} имеет единичную дисперсию. Очевидно, эту статистику можно записать в эквивалентной форме:

$$x = \sqrt{\frac{(n-1)(n-2)}{6n}} \frac{k_3}{\kappa_2^{3/2}} \left\{ 1 + \frac{k_2 - \kappa_2}{\kappa_2} \right\}^{-3/2}. \quad (12.92)$$

Поскольку распределение симметрично относительно нуля, среднее значение x равно нулю. Тогда, разлагая (12.92) в ряд, получим

$$\begin{aligned} x^2 = & \frac{(n-1)(n-2)}{6n} \cdot \frac{1}{\kappa_2^3} \left\{ k_3^2 - \frac{3}{\kappa_2} k_3^2 (k_2 - \kappa_2) + \frac{6}{\kappa_2^2} (k_2 - \kappa_2)^2 - \right. \\ & - \frac{10}{\kappa_2^3} k_3^2 (k_2 - \kappa_2)^3 + \frac{15}{\kappa_2^4} k_3^2 (k_2 - \kappa_2)^4 - \frac{21}{\kappa_2^5} k_3^2 (k_2 - \kappa_2)^5 + \\ & \left. + \frac{28}{\kappa_2^6} k_3^2 (k_2 - \kappa_2)^6 + \dots \right\}. \quad (12.93) \end{aligned}$$

Дисперсию можно получить, беря среднее значение от обеих частей. Поскольку κ_2 есть среднее значение k_2 , то

$$\begin{aligned} D_x = & \frac{(n-1)(n-2)}{6n} \frac{1}{\kappa_2^3} \left\{ \mu(3^2) - \frac{3}{\kappa_2} \mu(3^2 2) + \frac{6}{\kappa_2^2} \mu(3^2 2^2) - \right. \\ & - \frac{10}{\kappa_2^3} \mu(3^2 2^3) + \frac{15}{\kappa_2^4} \mu(3^2 2^4) - \frac{21}{\kappa_2^5} \mu(3^2 2^5) + \frac{28}{\kappa_2^6} \mu(3^2 2^6) + \dots \left. \right\}. \quad (12.94) \end{aligned}$$

Выразим теперь величины μ в терминах произведений κ , используя для этого соотношения (12.31) и отождествляя коэффициенты. В случае нормального распределения $\kappa(32^r) = 0$, и мы будем интересоваться только аппроксимацией с точностью до порядка n^{-4} , так что семиинварианты κ пятого и более высоких порядков во внимание можно не принимать. Тогда найдем

$$\begin{aligned} Dx = & \frac{(n-1)(n-2)}{6n} \frac{1}{\kappa_2^3} \left[\kappa(3^2) - \frac{3}{\kappa_2} \kappa(3^2 2) + \frac{6}{\kappa_2^2} \{ \kappa(3^2 2^2) + \right. \\ & + \kappa(3^2) \kappa(2^2) \} - \frac{10}{\kappa_2^3} \{ \kappa(3^2 2^3) + 3\kappa(3^2 2) \kappa(2^2) + \kappa(3^2) \kappa(2^3) \} + \\ & + \frac{15}{\kappa_2^4} \{ 6\kappa(3^2 2^2) \kappa(2^2) + 4\kappa(3^2 2) \kappa(2^3) + \kappa(3^2) \kappa(2^4) + \\ & + 3\kappa(3^2) \kappa^2(2^2) \} - \frac{21}{\kappa_2^5} \{ 15\kappa(3^2 2) \kappa^2(2^2) + 10\kappa(3^2) \kappa(2^3) \kappa(2^2) \} + \\ & \left. + \frac{28}{\kappa_2^6} 15\kappa(3^2) \kappa^3(2^2) \right]. \quad (12.95) \end{aligned}$$

Используя соответствующие значения из (12.35) — (12.89), после ряда чисто алгебраических преобразований получим

$$\begin{aligned} Dx = & 1 - \frac{6}{n-1} + \frac{28}{(n-1)^2} - \frac{120}{(n-1)^3} + \dots = \\ = & 1 - \frac{6}{n} + \frac{22}{n^2} - \frac{70}{n^3} + \dots \quad (12.96) \end{aligned}$$

Аналогичным путем можно показать, что

$$\mu_4(x) = 3 - \frac{1056}{n^2} + \frac{24132}{n^3} - \dots \quad (12.97)$$

Момент $\mu_3(x) = 0$ в силу симметричности распределения.

На самом деле в рассмотренном случае известны точные результаты (см. упражнение 12.9), которые можно использовать для проверки проделанных выше алгебраических преобразований.

Как следует из (12.96) и (12.97), при $n \rightarrow \infty$ второй момент для x стремится к 1, а четвертый — к 3, что соответствует сходимости к нормальному распределению. Но эта сходимость не очень быстрая. Когда $n=100$, дисперсия равна примерно 0,942. Поэтому допущение, что x имеет единичную дисперсию, будет давать ошибку около 6%. Существует два способа, улучшающих аппроксимацию распределения для x . Первый состоит в том, что от x мы переходим к новой переменной ξ , выбранной так, чтобы ξ была нормально распределена с точностью до величин порядка n^{-2} . Второй способ состоит в подгонке кривой

Пирсона к распределению статистики x с использованием моментов, задаваемых формулами (12.96) и (12.97).

Подходящая кривая относится к VII типу:

$$dF \propto \left(1 + \frac{x^2}{a^2}\right)^{-m} dx. \quad (12.98)$$

Первый способ был использован Фишером (1928), который получил следующее преобразование:

$$\xi = x \left(1 + \frac{3}{n} + \frac{91}{4n^2}\right) - \frac{3}{2n} \left(1 - \frac{111}{2n}\right) (x^3 - 3x) - \frac{33}{8n^2} (x^5 - 10x^3 + 15x). \quad (12.99)$$

Второй способ был использован Э. С. Пирсоном (1930), который для распределения (12.98) табулировал критические значения, соответствующие уровням значимости 1 и 5%, т. е. вычислил для различных n те значения переменной x , для которых в интервале (среднее значение $\pm x$) содержится 99 и 95% всей частоты.

12.19 С помощью указанных методов могут быть найдены аппроксимации выборочных распределений любых статистик, являющихся симметрическими функциями.

Читателю, интересующемуся этими вопросами, можно порекомендовать статьи Ф. Н. Дэвид (1949а, b), где он может ознакомиться с моментами выборочных распределений таких статистик, как коэффициент вариации, логарифм отношения дисперсий (z -статистика Фишера), отношение дисперсий. Трудности, связанные с разложением знаменателя в статистиках, являющихся отношениями двух статистик, можно преодолевать, используя следующие соображения.

Рассмотрим статистику t , представленную в виде a/b , где знаменатель всегда положителен. Тогда

$$\begin{aligned} F(t_0) &= P\{t \leq t_0\} = P\{a/b \leq t_0\} = \\ &= P\{a \leq bt_0\} = P\{a - bt_0 \leq 0\}. \end{aligned} \quad (12.100)$$

Теперь мы можем рассматривать статистику $a - bt_0$, которая уже не страдает недостатками, присущими отношению. Для примера, предположим, что рассматриваемая нами статистика t является t -статистикой Стьюдента:

$$t = \frac{(\bar{x} - \mu) \sqrt{n-1}}{s},$$

где μ — истинное среднее, а $s^2 = \sum (x - \bar{x})^2 / n$. Распределение этой статистики симметрично и поэтому

$$P\{|t| \leq |t_0|\} = P\{t^2 \leq t_0^2\} = P\{(\bar{x} - \mu)^2 (n-1) - s^2 t_0^2 \leq 0\}. \quad (12.101)$$

Рассмотрим теперь статистику

$$u = (n-1)(\bar{x} - \mu)^2 - t_0^2 s^2. \quad (12.102)$$

При заданных μ и t_0^2 мы можем для u найти столько моментов или семинвариантов, сколько этого захотим.

В частном случае нормального распределения мы знаем, что $\bar{x}$ и s независимы. Поэтому семинварианты статистики u легко найти прямым вычислением. Для характеристической функции случайной величины u , полагая $\theta = it$, получим с помощью (11.39) и (11.40)

$$\begin{aligned} \varphi(t) &\propto \int \int \exp\left\{[(n-1)(\bar{x} - \mu)^2 - t_0^2 s^2] \theta\right\} \exp\left\{-\frac{1}{2} n (\bar{x} - \mu)^2\right\} \times \\ &\quad \times d\bar{x} \exp\left\{-\frac{1}{2} n s^2\right\} s^{n-2} ds = \\ &= \left\{1 - \frac{2(n-1)}{n} \theta\right\}^{-1/2} \left\{1 + \frac{2t_0^2 \theta}{n}\right\}^{-\frac{1}{2}(n-1)}, \end{aligned} \quad (12.103)$$

откуда

$$\kappa_1 = \frac{n-1}{n} (1 - t_0^2), \quad (12.104)$$

$$\kappa_2 = \frac{2(n-1)}{n^2} (n-1 + t_0^4). \quad (12.105)$$

Когда $n \rightarrow \infty$, φ стремится к пределу $(1 - 2\theta)^{-1/2} \exp(-\theta t_0^2)$, первый сомножитель которого совпадает с х. ф. распределения III типа, т. е. с х. ф. χ^2 -распределения с одной степенью свободы. Рассмотрим теперь вероятность того, что $u \leq 0$. Для $n=20$, $t_0=2$ семинварианты $\kappa_1 = -2,85$, $\kappa_2 = 3,42$.

Если подсчитать для нормального распределения со средним $\kappa_1 = -2,85$ и дисперсией $\kappa_2 = 3,42$ вероятность того, что $u \leq 0$, то получим, что она равна 0,941. Для истинного же распределения эта вероятность равна 0,940.

Хотя мы сейчас рассматривали наш метод только в применении к нормальному распределению, уместно отметить, что он полезен при проверке эффектов отклонения распределений от нормального с помощью некоторых статистик. Укажем в этой связи на статью Дэвид и Джонсона (1951), где подробно изучается статистика, являющаяся отношением дисперсий.

Выбор из конечных совокупностей

12.20 Если результаты наблюдений получены в результате выбора без возвращения из конечной совокупности, то нахождение средних значений сильно усложняется по той причине, что последовательные извлечения не являются независимыми. Па-

раллельно с 12.3 рассмотрим некоторые средние значения выборочных распределений в случае конечной совокупности. Уравнения (12.2) остаются справедливыми:

$$Mm_2 = \frac{n-1}{n} Mx^2 - \frac{n-1}{n} Mx_i x_j, \quad i \neq j. \quad (12.106)$$

Если использовать M для обозначения истинных моментов конечной генеральной совокупности объема N , то по определению будем иметь

$$\begin{aligned} Mx^2 &= M'_2, \\ Mx_i x_j &= \sum_{i,j=1}^N x_i x_j / \{N(N-1)\} = \quad (i \neq j) \\ &= \frac{\left(\sum_1^N x_i\right)^2 - \sum_1^N x_i^2}{N(N-1)} = \frac{N^2 M'_2 - N M'_2}{N(N-1)}. \end{aligned}$$

Подставляя эти моменты в (12.106), получим

$$Mm_2 = \frac{n-1}{n} \left\{ \frac{N}{N-1} [M'_2 - (M'_1)^2] \right\} = \frac{n-1}{n} \frac{N}{N-1} M_2. \quad (12.107)$$

С помощью того же самого подхода мы можем найти все моменты статистик, являющихся симметрическими функциями. Но вычисления с каждым шагом будут все более усложняться и трудно поддаваться контролю. Более простой путь состоит в следующем.

12.21 Заметим сначала, что среднее значение каждой расширенной симметрической функции от наблюдаемых значений очень просто связано с соответствующей функцией от истинных значений. Снабжая эти две функции индексами n и N соответственно, мы получим

$$M[p_1^{\pi_1} \dots p_s^{\pi_s}]_n / n^{|\rho|} = [p_1^{\pi_1} \dots p_s^{\pi_s}]_N / N^{|\rho|}, \quad (12.108)$$

где $\rho = \sum_i \pi_i$. Отсюда следует, что если K_p есть p -я k -статистика исходного распределения, то

$$M_N k_p = K_p, \quad (12.109)$$

где M_N означает усреднение по множеству из N значений, образующих исходную совокупность. Это привлекательное свойство k -статистик очень удобно в работе. К примеру, рассмотрим $k_2 = nm_2 / (n-1)$. Тогда

$$M_N \frac{nm_2}{n-1} = K_2 = \frac{NM_2}{N-1},$$

откуда сразу же вытекает (12.107).

Представим теперь, что совокупность из N элементов получена как выборка из бесконечной совокупности, семинварианты которой равны κ_p . Если символ математического ожидания M относится к этой бесконечной совокупности, то

$$Mk_p = MK_p = \kappa_p.$$

Если мы имеем дело с некоторой симметрической функцией f , среднее значение которой может быть выражено в терминах κ линейным образом:

$$M(f) = \sum a_j \kappa_j, \quad (12.110)$$

то справедливо и аналогичное соотношение

$$M_N(f) = \sum a_j K_j. \quad (12.111)$$

В самом деле, если это не так, то $M_N(f)$ может быть выражено как некоторая другая функция от K , усреднение M которой будет таким же, как усреднение от $\sum a_j K_j$, и из этого тогда бы следовала связь между различными K . Это обстоятельство известно как принцип Ирвина — Кендалла, по имени авторов, которые его ввели в 1944 году.

Пример 12.8. Моменты среднего значения в случае конечных совокупностей

Среднее значение выборки m_1 равно k_1 . Поэтому

$$M_N(k_1) = K_1. \quad (12.112)$$

Из хорошо известного результата относительно дисперсии выборочного среднего значения следует, что

$$M(k_1^2) = \frac{\kappa_2}{n} + \kappa_1^2, \quad M(K_1^2) = \frac{\kappa_2}{N} + \kappa_1^2.$$

Следовательно,

$$M(k_1 - K_1)^2 = Mk_1^2 - MK_1^2 = \kappa_2 \left(\frac{1}{n} - \frac{1}{N} \right).$$

Поэтому

$$M_N(k_1 - K_1)^2 = K_2 \left(\frac{1}{n} - \frac{1}{N} \right), \quad (12.113)$$

что и дает дисперсию среднего значения. Если бы мы пожелали представить это выражение в терминах исходных моментов, то получили бы

$$M_N(k_1 - K_1)^2 = \frac{N-n}{(N-1)n} M_2. \quad (12.114)$$

Для третьего момента имеем

$$\begin{aligned} M(k_1 - \kappa_1)^3 &= M(k_1 - K_1 + K_1 - \kappa_1)^3 = \\ &= M(k_1 - K_1)^3 + 3M(k_1 - K_1)^2(K_1 - \kappa_1) + 3M(k_1 - K_1)(K_1 - \kappa_1)^2 + \\ &\quad + M(K_1 - \kappa_1)^3. \end{aligned} \quad (12.115)$$

Третье слагаемое справа равно нулю, поскольку среднее значение $k_1 - K_1$ для конечных совокупностей обращается в нуль. Далее,

$$M(k_1 - \kappa_1)^3 = \frac{\kappa_3}{n^2}, \quad M(K_1 - \kappa_1)^3 = \frac{\kappa_3}{N^2}.$$

Следовательно, из (12.115) получаем

$$M(k_1 - K_1)^3 = \kappa_3 \left(\frac{1}{n^2} - \frac{1}{N^2} \right) - 3M(K_1 - \kappa_1)(k_1 - K_1)^2.$$

Положим

$$\alpha_r = \frac{1}{n^r} - \frac{1}{N^r}, \quad (12.116)$$

тогда

$$\begin{aligned} M(k_1 - K_1)^3 &= \alpha_2 \kappa_3 - 3M(K_1 - \kappa_1) M_N(k_1 - K_1)^2 = \\ &= \alpha_2 \kappa_3 - 3M(K_1 - \kappa_1) \alpha_1 K_2 = \\ &= \alpha_2 \kappa_3 - 3\alpha_1 M(K_1 - \kappa_1)(K_2 - \kappa_2) = \\ &= \alpha_2 \kappa_3 - \frac{3\alpha_1 \kappa_3}{N} = \left(\alpha_2 - \frac{3\alpha_1}{N} \right) \kappa_3. \end{aligned}$$

Таким образом,

$$M_N(k_1 - K_1)^3 = \left(\alpha_2 - \frac{3\alpha_1}{N} \right) K_3. \quad (12.117)$$

Подставляя сюда значения α_2 , α_1 и

$$K_3 = \frac{N^2}{(N-1)(N-2)} M_3,$$

получим

$$M_N(m_1 - M_1)^3 = \frac{(N-n)(N-2n)}{n^2(N-1)(N-2)} M_3. \quad (12.118)$$

Подобным же образом можно найти, что

$$\begin{aligned} M_N(k_1 - K_1)^4 &= K_4 \left\{ \alpha_3 - 4\alpha_2/N + G\alpha_1/N^2 - \right. \\ &\quad \left. - \frac{3(N-1)}{N(N+1)} (\alpha_2 - 2\alpha_1/N) \right\} + 3K_2^2 \frac{N-1}{N+1} (\alpha_2 - 2\alpha_1/N), \end{aligned} \quad (12.119)$$

что эквивалентно соотношению

$$\begin{aligned} M_N(m_1 - M_1)^4 &= \frac{N-n}{n^3(N-1)(N-2)(N-3)} \{ (N^2 - 6nN + N + 6n^2) M_4 + \\ &\quad + 3N(n-1)(N-n-1) M_2^2 \}. \end{aligned} \quad (12.120)$$

Пример 12.9

Мы имеем

$$M(k_r - \kappa_r)(k_1 - \kappa_1) = M(k_r - K_r)(k_1 - K_1) + M(K_r - \kappa_r)(K_1 - \kappa_1)$$

и, следовательно,

$$M(k_r - K_r)(k_1 - K_1) = \frac{1}{n} \kappa_{r+1} - \frac{1}{N} \kappa_{r+1} = \alpha_1 \kappa_{r+1},$$

откуда

$$M_N(k_r - K_r)(k_1 - K_1) = \alpha_1 K_{r+1}. \quad (12.121)$$

Поэтому если конечная совокупность симметрична, то каждая k -статистика четного порядка не коррелирована с k_1 , так как в противном случае $K_{r+1} = 0$.

12.22 Продолжение наших исследований наталкивается на некоторые трудности, связанные с обозначениями. Мы уже использовали такие символы, как κ , k и K . В следующей главе мы будем заниматься многомерным случаем и нам придется рассматривать аналогичные величины, но уже снабженные несколькими индексами, например κ_2 , k_{rs} , K_{rs} . Сейчас нам требуется ввести обозначение для статистик, чьи средние значения равны произведению семиинвариантов κ , например $\kappa_r \kappa_s$. Вопреки обозначениям Тьюки (1950), Уишарта (1952), Кендалла (1952), мы будем обозначать эти величины буквами l с индексами. Они определяются из соотношений

$$Ml_{rs \dots u} = \kappa_r \kappa_s \dots \kappa_u. \quad (12.122)$$

Для случая конечных совокупностей будем пользоваться обозначениями $L_{rs \dots u}$, так что

$$M_N l_{rs \dots u} = L_{rs \dots u}. \quad (12.123)$$

Использование этих функций основано на том факте, что мы можем их раз и навсегда выразить через некоторые расширенные симметрические функции, представить нелинейные функции от них как некоторые линейные и затем к ним применить рассуждения 12.21. Необходимые таблицы до 12-го порядка были даны Абдель Ати (1954). Уишарт (1952) указал (вплоть до членов восьмого порядка) представление этих многомерных k -статистик через произведение простых k -статистик. Так, из таблиц Уишарта следует, что

$$k_2 k_1 = l_2 l_1 = l_3/n + l_{21} = k_3/n + l_{21}, \quad (12.124)$$

откуда

$$M(k_2 k_1) = \kappa_3/n + \kappa_2 \kappa_1, \quad (12.125)$$

$$M_N(k_2 k_1) = K_3/n + K_2 K_1. \quad (12.126)$$

Подобным образом получаем

$$k_1^4 = k_4/n^3 + 4l_{31}/n^2 + 3l_{22}/n^2 + 6l_{211}/n + l_{1111}, \quad (12.127)$$

откуда следует, что

$$M(k_1^4) = \kappa_4/n^3 + 4\kappa_3\kappa_1/n^2 + 3\kappa_2^2/n^2 + 6\kappa_2\kappa_1^2/n + \kappa_1^4. \quad (12.128)$$

Аналогичное представление имеет место и для $M_N(k_1^4)$.

Пример 12.10

Одно из наиболее важных практических применений этих результатов состоит в том, что с их помощью мы можем получать несмещенные оценки произведений семинвариантов. В примерах 12.1 и 12.3 мы нашли, что

$$Dm_2 = \frac{(n-1)^2}{n^2} \left\{ \frac{\kappa_4}{n} + \frac{2\kappa_2^2}{n-1} \right\}. \quad (12.129)$$

Часто случается так, что сами значения κ_4 и κ_2 , входящие в (12.129), нам неизвестны и приходится вместо них пользоваться их оценками. Спрашивается, какие оценки для κ_4 и κ_2 надо взять, чтобы полученная оценка для Dm_2 была бы несмещенной? Как легко видеть, такой несмещенной оценкой является величина

$$\frac{(n-1)^2}{n^2} \left\{ \frac{k_4}{n} + \frac{2l_{22}}{n-1} \right\}. \quad (12.130)$$

По таблицам Уишарта

$$l_{22} = \frac{n-1}{n+1} \left(k_2^2 - \frac{k_4}{n} \right),$$

откуда, подставляя l_{22} в (12.130), получим следующую несмещенную оценку:

$$\frac{(n-1)^2}{n^2(n+1)} \left\{ \frac{(n-1)k_4}{n} + 2k_2^2 \right\}. \quad (12.131)$$

Пример 12.11

Найти несмещенную оценку для $\mu_3(\mu'_1)^2$. Эта величина равна $\kappa_3\kappa_1^2$ и ее оценкой является l_{311} . Пользуясь таблицей 11 приложения, мы можем эту оценку выразить в терминах расширенных симметрических функций:

$$l_{311} = \frac{2}{n^{[5]}} [1^5] - \frac{3}{n^{[4]}} [21^3] + \frac{[31^2]}{n^{[3]}}.$$

С помощью таблицы 10 приложения можно выразить расширенные симметрические функции в терминах степенных сумм (r) .

В результате получим

$$l_{311} = \frac{2}{n^{[5]}} \{24(5) - 30(4)(1) - 20(3)(2) + 20(3)(1)^2 + \\ + 15(2)^2(1) - 10(2)(1)^3 + (1)^5\} - \frac{3}{n^{[4]}} \{-6(5) + 6(4)(1) + \\ + 5(3)(2) - 3(3)(1)^2 - 3(2)^2(1) + (2)(1)^3\} + \\ + \frac{1}{n^{[3]}} \{2(5) - 2(4)(1) - (3)(2) + (3)(1)^2\}.$$

Если положить $(r) = s_r$, то l_{311} сведется к выражению

$$\frac{1}{n^{[5]}} \{2n(n+2)s_5 - 2(n^2 + 2n + 6)s_4s_1 - (n^2 + 8n - 8)s_3s_2 + \\ + (n^2 + 2n + 16)s_3s_1^2 + (9n - 16)s_2^2s_1 - (3n + 8)s_2s_1^3 + 2s_1^5\}.$$

Некоторые американские авторы называют l -статистики «многократными»^{*}). Мы не будем придерживаться этого названия, поскольку считаем, что должен существовать предел лингвистическим увлечениям.

УПРАЖНЕНИЯ

12.1 Используя непосредственно (12.25), выразить k_3 и k_4 через расширенные симметрические функции и, исходя из этого, доказать (12.29).

12.2 Предполагая, что

$$k_3 = a_0s_3 + a_1s_2s_1 + a_2s_1^3,$$

найти коэффициенты a путем усреднения обеих частей и тем самым получить выражение для k_3 , содержащееся в (12.28).

12.3 Используя подход примера 12.1, получить выражение $\kappa(2^3)$ в терминах семинвариантов исходного распределения.

12.4 Показать, что шаблонные функции для таблиц

$$\begin{array}{cccccc} \times & \times & \times & \times & \times & \times \\ \times & . & \times & . & . & \times \\ \times & \times & . & \times & \times & . \end{array}$$

равны $\frac{1}{(n-1)^2}$ и $\frac{1}{n(n-1)^2}$ соответственно.

(Фишер, 1928.)

12.5 Показать, что шаблонная функция таблицы

$$\begin{array}{cccc} \times & \times & \times & . \\ \times & \times & \times & . \end{array}$$

с p столбцами равна $\frac{1}{n^{p-1}} \left\{ 1 - \frac{(-1)^{p-1}}{(n-1)^{p-1}} \right\}$.

(Фишер, 1928.)

12.6 Доказать формулы (12.37) и (12.43).

^{*}) Нам представляется, что такой перевод слова «polykaus» отвечает существу дела. (Прим. перев.)

12.7 Предположим, что некоторый шаблон содержит столбец с тремя входами (X). Рассмотрим шаблоны, получающиеся из исходного зачеркиванием этого столбца и (1) объединением трех строк, которым принадлежат три входа отброшенного столбца; (2) попарным объединением этих трех строк; (3) оставлением этих строк без изменения. Пусть этим шаблонам отвечают шаблонные функции A ; B_1 , B_2 , B_3 и C соответственно. Показать, что шаблонная функция исходного шаблона равна

$$\frac{n}{(n-1)(n-2)} A - \frac{1}{(n-1)(n-2)} (B_1 + B_2 + B_3) + \frac{2}{n(n-1)(n-2)} C.$$

Вывести отсюда, что шаблон

$$\begin{array}{ccc} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & . \end{array}$$

имеет шаблонную функцию, равную

$$\frac{n^3 - 8n^2 + 17n + 2}{(n-1)^2 (n-2)^2 (n-3)}.$$

(Фишер и Уишарт, 1931.)

12.8 Используя комбинаторный метод, доказать, что с точностью до членов порядка n^{-1}

$$Dm_4 = (\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2 - 8\mu_5\mu_3)/n.$$

12.9 Доказать, что в случае нормальных выборок k_2 не зависит от $k_p/k_2^{\frac{1}{2}p}$ (ср. упражнение 11.4). Используя это, показать, что

$$D\left(k_3/k_2^{\frac{3}{2}}\right) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)},$$

$$\mu_4\left(k_3/k_2^{\frac{3}{2}}\right) = \frac{108n^2(n-1)^2(n^2+27n-70)}{(n-2)^8(n+1)(n+3)(n+5)(n+7)(n+9)},$$

и проверить (12.96) и (12.97).

12.10 Пусть статистика y определена соотношением

$$y = \sqrt{\frac{(n-1)(n-2)(n-3)}{24n(n+1)}} \frac{k_4}{k_2^2}.$$

Показать с помощью методов 12.18, что моменты y в случае нормального исходного распределения имеют вид

$$\mu_1 = 0$$

$$\mu_2 = 1 - \frac{12}{n} + \frac{88}{n^2} - \frac{532}{n^3} \dots$$

$$\mu_3 = 6 \sqrt{\frac{6}{n}} \left(1 - \frac{65}{2n} + \frac{4811}{8n^2} - \frac{136605}{16n^3} \dots \right),$$

$$\mu_4 = 3 + \frac{468}{n} - \frac{32196}{n^2} + \frac{1118388}{n^3} \dots$$

(Э. С. Пирсон, 1930.)

12.11 Показать, что в случае конечных генеральных совокупностей

$$M_N(k_2 - K_2)^2 = \frac{(N-n)(Nn-n-N-1)}{n(n-1)N(N+1)} K_4 + \frac{2(N-n)}{(n-1)(N+1)} K_2^2,$$

что эквивалентно выражению

$$M_N \left(m_2 - \frac{n-1}{n} \frac{N}{N-1} M_2 \right)^2 = \frac{(n-1)N(N-n)}{n^3(N-1)^2(N-2)(N-3)} \times \\ \times \{ (nN - N - n - 1)(N-1) M_4 - (nN^2 - 3N^2 + 6N - 3N - 3) M_2^2 \}.$$

12.12 Аналогично установить, что

$$M_N(k_3 - K_3)(k_2 - K_2) = \\ = \frac{(N-n)\{(n-1)(N-1)-6\}}{n(n-1)N(N+5)} K_5 + \frac{6(N-n)}{(n-1)(N+5)} K_3 K_2.$$

12.13 Показать, что для случая симметрических конечных генеральных совокупностей k_{2r} не коррелирована с k -статистиками нечетного порядка.

12.14 Показать, что если $M(z)$ — производящая функция моментов для x , то производящая функция моментов $M(\xi)$ для $y=f(x)$ может быть представлена в виде

$$M(\xi) = \left[\exp \left\{ \xi f \left(\frac{d}{dz} \right) \right\} M(z) \right]_{z=0}.$$

Установить, что производящая функция моментов k -статистик

$$\sum \left\{ \mu(p_1^{\pi_1} \dots p_s^{\pi_s}) \frac{t_1^{\pi_1}}{\pi_1!} \dots \frac{t_s^{\pi_s}}{\pi_s!} \right\}$$

равна

$$\left[\exp \{ t_1 K_1 + t_2 K_2 + t_3 K_3 + \dots \} \exp \left\{ \kappa_1 s_1 + \frac{\kappa_2 s_2}{2!} + \dots \right\} \right]_{\kappa=0},$$

где K_r так же зависит от операторов $\frac{\partial}{\partial x}$, как и k_r от результатов наблюдений x , а $s_r = \sum (x^r)$.

Установить, что

$$K_p(s_p) = p!, \quad K_p(s_{p_1} s_{p_2} \dots) = 0,$$

где $(p_1 p_2 \dots)$ — некоторое разбиение числа p .

(Фишер, 1928.)

12.15 Показать, что

$$M \{ k_p(x_1 + h, x_2, \dots, x_n) \} = \kappa_p + \frac{1}{n} h^p$$

и что

$$S_p k_p = p!, \quad S_q k_p = 0, \quad q \neq p,$$

где S_p так же зависит от операторов $\frac{\partial}{\partial x}$, как s_p от наблюдений x .

(Кендалл, 1940а.)

12.16 Показать, что производящая функция моментов k -статистик задается формулой

$$\left[\exp \left\{ \frac{S_1 x_1}{1!} + \frac{S_2 x_2}{2!} + \dots \right\} \exp \{ k_1 t_1 + k_2 t_2 + \dots \} \right]_{x=0}.$$

Вывести отсюда результат предыдущего упражнения.

(Кендалл, 1942.)

12.17 Используя упражнение 12.16, показать, что в представлении k_p в терминах симметрических сумм s сумма коэффициентов равна $1/n$.

12.18 Показать, что для $(\mu_1')^4$ несмещенная оценка равна

$$\frac{1}{n^{[4]}} \{ s_1^4 - 6s_2s_1^2 + 3s_2^2 + 8s_3s_1 - 6s_4 \}.$$

12.19 Показать, что симметрические функции l_r ... могут быть представлены в виде

$$\sum' \frac{(-1)^{\sum (p-1)} \Pi (p-1)!}{\Pi \{ \pi_1! \pi_2! \dots \}} \cdot \frac{r! s! \dots}{\Pi \{ (p_1!)^{\pi_1} (p_2!)^{\pi_2} \dots \}} \frac{\{ \Pi (p_1^{\pi_1} p_2^{\pi_2} \dots) \}}{n^{(\sum p)}},$$

где $(p_1^{\pi_1} p_2^{\pi_2} \dots)$ есть разбиение r , $\sum p_i \pi_i = r$, $\sum \pi_i = \rho$; аналогичные разбиения отвечают s и т. д.; Π и Σ означают произведения и суммы по всем $r, s, \dots$; $\sum'$ означает суммирование по всем разбиениям.

Вывести отсюда, что

$$l_{222} = -\frac{[1^6]}{n^{[6]}} + \frac{3[21^4]}{n^{[5]}} - \frac{3[2^2 1^2]}{n^{[4]}} + \frac{[2^3]}{n^{[3]}}$$

(ср. с таблицей 11 приложения).

12.20 С помощью методов примера 12.7 показать, что если k_1 и k_p не коррелированы для всех $p > 1$, то исходное распределение нормально. Если же k_1 не зависит от какой-либо k_p ($p > 1$), то распределение также нормально. Указание — использовать упражнение 4.5.

АППРОКСИМАЦИЯ ВЫБОРОЧНЫХ РАСПРЕДЕЛЕНИЙ — (2)

13.1 Эта глава посвящена главным образом следующим двум темам. Первая — обобщение развитых в предыдущей главе методов для одномерного случая на двумерные и многомерные статистики. Вторая — изучение структуры k -статистик и семинвариантов; при этом особое внимание будет уделено доказательству комбинаторных правил для вывода выборочных свойств этих статистик.

Двумерные k -статистики

13.2 Для всякой пары чисел pp' имеется k -статистика $k_{pp'}$, среднее значение которой равно двумерному семинварианту $\kappa_{pp'}$. Она имеет вид

$$k_{pp'} = p! p'! \sum \frac{(-1)^p (p-1)!}{n^{[p]}} \sum \frac{(x_{i_1}^{p_1} y_{i_1}^{p'_1} x_{i_2}^{p_2} y_{i_2}^{p'_2} \dots x_{i_p}^{p_p} y_{i_p}^{p'_p})}{(p_1!)^{\pi_1} (p'_1)^{\pi'_1} \dots \pi_1! \pi_2! \dots} \quad (13.1)$$

В частности, аналогично (12.28),

$$\left. \begin{aligned} k_{11} &= \frac{1}{n^{[2]}} (ns_{11} - s_{10}s_{01}), \\ k_{21} &= \frac{1}{n^{[3]}} (n^2 s_{21} - 2ns_{10}s_{11} - ns_{20}s_{01} + 2s_{10}^2 s_{01}), \\ k_{31} &= \frac{1}{n^{[4]}} \{n^2 (n+1) s_{31} - n(n+1) s_{30}s_{01} - 3n(n-1) s_{11}s_{20} - \\ &\quad - 3n(n+1) s_{21}s_{10} + 6ns_{11}s_{10}^2 + 6ns_{20}s_{10}s_{01} - 6s_{01}s_{10}^2\}, \\ k_{22} &= \frac{n}{(n-1)(n-2)(n-3)} \left\{ (n+1) s_{22} - \frac{2(n+1)}{n} s_{21}s_{01} - \right. \\ &\quad - \frac{2(n+1)}{n} s_{12}s_{10} - \frac{n-1}{n} s_{20}s_{02} - \frac{2(n-1)}{n} s_{11}^2 + \\ &\quad \left. + \frac{8}{n} s_{11}s_{10}s_{01} + \frac{2}{n} s_{02}s_{10}^2 + \frac{2}{n} s_{20}s_{01}^2 - \frac{6}{n^2} s_{10}^2 s_{01}^2 \right\}. \end{aligned} \right\} \quad (13.2)$$

Для средних значений произведений таких статистик k можно принять, например, следующие обозначения:

$$M(k_{20}k_{11}) = \mu' \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix},$$

$$M(k_{20}k_{11}k_{02}) = \mu' \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix};$$

соответствующие обозначения принимаются и для семиинвариантов κ . Как и в одномерном случае, последние могут быть выражены в терминах семиинвариантов исходного двумерного распределения; коэффициенты в этих выражениях будут теперь зависеть от разбиений пар чисел. Сформулированные в предыдущей главе правила по-прежнему применимы (в частности, шаблонные функции, отвечающие определенным таблицам, те же самые), однако числовые коэффициенты, которые ставятся в соответствие заполненным таблицам, меняются, поскольку теперь мы должны рассматривать число способов размещения двух разных классов элементов, так как здесь мы имеем разбиения пары чисел. Чтобы сделать это более ясным, приведем пример.

Предположим, что нужно найти коэффициент при $\kappa_{33}\kappa_{11}^2$ в $\kappa \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}$. Так как общий порядок равен здесь 10 и порядки произведения суть 6, 2, 2, то нужно рассматривать таблицы типа

				6
				2
				2
4	4	2		10

Такие таблицы уже рассматривались в главе 12. Там были найдены соответствующие шаблонные функции. Чтобы вычислить коэффициенты, суммарный столбец нужно рассматривать как состоящий из двойных элементов: (3,3), (1,1) и (1,1), и аналогично для суммарной строки: (2,2), (2,2), (1,1). Например, таблицу

2	2	2		6
1	1	.		2
1	1	.		2
4	4	2		10

можно представить как

$$\begin{array}{ccc|c}
 (1,1) & (1,1) & (1,1) & (3,3) \\
 (0,1) & (1,0) & . & (1,1) \\
 (1,0) & (0,1) & . & (1,1) \\
 \hline
 (2,2) & (2,2) & (1,1) & (5,5)
 \end{array} \quad (13.3)$$

или как

$$\begin{array}{ccc|c}
 (2,0) & (0,2) & (1,1) & (3,3) \\
 (0,1) & (1,0) & . & (1,1) \\
 (0,1) & (1,0) & . & (1,1) \\
 \hline
 (2,2) & (2,2) & (1,1) & (5,5)
 \end{array} \quad (13.4)$$

Каждая из таблиц (13.3), (13.4) дает свой вклад в числовой коэффициент. Оказывается, что, кроме этих таблиц, а также таблиц, получающихся из них перестановкой первых двух столбцов, других подходящих таблиц нет.

Суммарный вклад таблицы (13.3) и таблицы, получающейся из нее перестановкой, равен

$$2 \left(\frac{2!}{1!1!} \right) \left(\frac{2!}{1!1!} \right) \left(\frac{2!}{1!1!} \right) \left(\frac{2!}{1!1!} \right) \frac{1}{2!} = 16.$$

Аналогично таблица (13.4) и таблица, получающаяся из нее перестановкой, дают суммарный вклад

$$2! \left(\frac{2!}{2!} \right) \left(\frac{2!}{1!1!} \right) \left(\frac{2!}{1!1!} \right) \left(\frac{2!}{2!} \right) \frac{1}{2!} = 4.$$

Общий вклад равен, следовательно, 20. Соответствующая шаблонная функция есть $\frac{1}{(n-1)(n-2)}$.

Таким же образом можно убедиться, что отвечающие разбиениям

$$\begin{array}{ccc|c}
 2 & 3 & 1 & 6 \\
 1 & 1 & . & 2 \\
 1 & . & 1 & 2 \\
 \hline
 4 & 4 & 2 & 10
 \end{array}
 \quad
 \begin{array}{ccc|c}
 3 & 3 & . & 6 \\
 1 & . & 1 & 2 \\
 . & 1 & 1 & 2 \\
 \hline
 4 & 4 & 2 & 10
 \end{array}$$

коэффициенты суть 48 и 8. Следовательно, искомый коэффициент при $\kappa_{33}\kappa_{11}^2$ равен

$$\frac{20}{(n-1)(n-2)} + \frac{48}{(n-1)^2} + \frac{8}{(n-1)^2} = \frac{4(19n-33)}{(n-1)^2(n-2)}.$$

Пример 13.1

Найти точное выражение для ковариации оценок дисперсий двух коррелированных случайных величин, т. е. для

$$\kappa \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Ясно, что искомое выражение есть линейная комбинация трех членов κ_{22} , $\kappa_{02}\kappa_{20}$ и κ_{11}^2 . Для первого члена имеем разбиение

$$\frac{\begin{array}{cc|c} (2,0) & (0,2) & (2,2) \\ \hline (2,0) & (0,2) & (2,2) \end{array}}{1},$$

которому отвечают шаблонная функция $\frac{1}{n}$ и числовой коэффициент, равный единице. Коэффициент при втором члене есть нуль, поскольку для единственной соответствующей ему таблицы

$$\frac{\begin{array}{cc|c} (2,0) & . & (2,0) \\ . & (0,2) & (0,2) \\ \hline (2,0) & (0,2) & (2,2) \end{array}}{0} = 0,$$

шаблонная функция равна нулю. Таблице для третьего члена

$$\frac{\begin{array}{cc|c} (1,0) & (0,1) & (1,1) \\ (1,0) & (0,1) & (1,1) \\ \hline (2,0) & (0,2) & (2,2) \end{array}}{2} = \frac{1}{n-1},$$

отвечают шаблонная функция $\frac{1}{n-1}$ и числовой коэффициент 2. Следовательно,

$$\kappa \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \frac{1}{n} \kappa_{22} + \frac{2}{n-1} \kappa_{11}^2.$$

Отметим аналогию между этим выражением и выражением для дисперсии k_{20} (ср. (12.35))

$$\kappa \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix} = \frac{\kappa_{40}}{n} + \frac{2\kappa_{20}^2}{n-1}.$$

В частности, если генеральная совокупность нормальна, то все κ порядка выше второго обращаются в нуль, так что коэффициент корреляции k_{20} и k_{02} равен $\frac{\kappa_{11}^2}{\kappa_{20}\kappa_{02}} = \rho^2$, где ρ — коэффициент корреляции исходного двумерного нормального распределения.

13.3 Формулы для двумерного случая можно вывести из формул для одномерного случая с помощью предложенного Кендал-

лом (1940с) формального процесса. Этим методом Кук (1951) получила двумерные формулы до четвертого порядка по каждой из переменных. За формулами и деталями следует обращаться к ее статье; в одной из более поздних статей ею также дан ряд приложений этих формул. Здесь мы выпишем некоторые из наиболее простых результатов:

$$\kappa \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{31} + \frac{2}{n-1} \kappa_{20} \kappa_{11}, \quad (13.5)$$

$$\kappa \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \frac{1}{n} \kappa_{22} + \frac{2}{n-1} \kappa_{11}^2, \quad (13.6)$$

$$\kappa \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{22} + \frac{1}{n-1} \kappa_{20} \kappa_{02} + \frac{1}{n-1} \kappa_{11}^2, \quad (13.7)$$

$$\kappa \begin{pmatrix} 3 & 3 \\ 0 & 0 \end{pmatrix} = \frac{1}{n} \kappa_{60} + \frac{9}{n-1} \kappa_{40} \kappa_{20} + \frac{9}{n-1} \kappa_{30}^2 + \frac{6n}{(n-1)(n-2)} \kappa_{20}^2 \kappa_{11} \quad (13.8)$$

(эта формула, по существу, относится к одномерному случаю),

$$\begin{aligned} \kappa \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{51} + \frac{6}{n-1} \kappa_{31} \kappa_{20} + \frac{3}{n-1} \kappa_{40} \kappa_{11} + \\ + \frac{9}{n-1} \kappa_{30} \kappa_{21} + \frac{6n}{(n-1)(n-2)} \kappa_{20}^2 \kappa_{11}, \end{aligned} \quad (13.9)$$

$$\begin{aligned} \kappa \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} = \frac{1}{n} \kappa_{42} + \frac{3}{n-1} \kappa_{22} \kappa_{20} + \frac{6}{n-1} \kappa_{31} \kappa_{11} + \\ + \frac{3}{n-1} \kappa_{30} \kappa_{12} + \frac{6}{n-1} \kappa_{21}^2 + \frac{6n}{(n-1)(n-2)} \kappa_{20} \kappa_{11}^2, \end{aligned} \quad (13.10)$$

$$\kappa \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \frac{1}{n} \kappa_{33} + \frac{9}{n-1} \kappa_{22} \kappa_{11} + \frac{9}{n-1} \kappa_{12} \kappa_{21} + \frac{6n}{(n-1)(n-2)} \kappa_{11}^3, \quad (13.11)$$

$$\begin{aligned} \kappa \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{42} + \frac{4}{n-1} \kappa_{22} \kappa_{20} + \frac{4}{n-1} \kappa_{31} \kappa_{11} + \frac{1}{n-1} \kappa_{40} \kappa_{02} + \\ + \frac{4}{n-1} \kappa_{30} \kappa_{12} + \frac{5}{n-1} \kappa_{21}^2 + \\ + \frac{4n}{(n-1)(n-2)} \kappa_{20} \kappa_{11}^2 + \frac{2n}{(n-1)(n-2)} \kappa_{20}^2 \kappa_{02}, \end{aligned} \quad (13.12)$$

$$\begin{aligned} \kappa \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{n} \kappa_{33} + \frac{5}{n-1} \kappa_{22} \kappa_{11} + \frac{2}{n-1} \kappa_{13} \kappa_{20} + \\ + \frac{2}{n-1} \kappa_{31} \kappa_{02} + \frac{1}{n-1} \kappa_{30} \kappa_{03} + \frac{8}{n-1} \kappa_{21} \kappa_{12} + \\ + \frac{4n}{(n-1)(n-2)} \kappa_{11} \kappa_{20} \kappa_{02} + \frac{2n}{(n-1)(n-2)} \kappa_{11}^3. \end{aligned} \quad (13.13)$$

Эти формулы по своему характеру, очевидно, очень схожи. Ниже мы увидим, как их можно записать в сжатой форме.

Пример 13.2

Рассмотренные методы можно обобщить на случай конечных двумерных генеральных совокупностей. Получающиеся при этом формулы нужны не столь часто, чтобы их стоило здесь выписывать, однако будет полезно проиллюстрировать способ их нахождения. Вычислим дисперсию величины k_{11} для выборки объема n из генеральной совокупности объема N . Как в 12.21, имеем

$$M(k_{11} - K_{11})^2 = M(k_{11}^2) - M(K_{11}^2). \quad (13.14)$$

В силу (13.7)

$$\begin{aligned} M(k_{11}^2) &= \mu' \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \kappa \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \kappa_{11}^2 = \\ &= \frac{1}{n} \kappa_{22} + \frac{1}{n-1} \kappa_{20} \kappa_{02} + \frac{1}{n-1} \kappa_{11}^2 + \kappa_{11}^2, \end{aligned} \quad (13.15)$$

$$M(K_{11}^2) = \frac{1}{N} \kappa_{22} + \frac{1}{N-1} \kappa_{20} \kappa_{02} + \frac{1}{N-1} \kappa_{11}^2 + \kappa_{11}^2. \quad (13.16)$$

Подставляя (13.15) и (13.16) в (13.14), получаем

$$\begin{aligned} Dk_{11} &= \left(\frac{1}{n} - \frac{1}{N} \right) \kappa_{22} + \left(\frac{1}{n-1} - \frac{1}{N-1} \right) \kappa_{20} \kappa_{02} + \\ &+ \left(\frac{1}{n-1} - \frac{1}{N-1} \right) \kappa_{11}^2, \end{aligned}$$

что сводится к равенству

$$\frac{Dk_{11}}{N-n} - \frac{1}{nN} \kappa_{22} = \frac{1}{(n-1)(N-1)} \kappa_{20} \kappa_{02} + \frac{1}{(n-1)(N-1)} \kappa_{11}^2. \quad (13.17)$$

Теперь нужно исключить стоящие в (13.17) справа нелинейные члены. Согласно (13.16)

$$M(K_{11}^2) - \frac{1}{N} \kappa_{22} = \frac{1}{N-1} \kappa_{20} \kappa_{02} + \frac{N}{N-1} \kappa_{11}^2. \quad (13.18)$$

Пользуясь формулой (13.6) в терминах K , аналогично находим

$$M(K_{20} K_{02}) - \frac{1}{N} \kappa_{22} = \kappa_{20} \kappa_{02} + \frac{2}{N-1} \kappa_{11}^2. \quad (13.19)$$

Из (13.17), (13.18) и (13.19) следует, что

$$\begin{vmatrix} \frac{Dk_{11}}{N-n} - \frac{1}{nN} \kappa_{22} & \frac{1}{(n-1)(N-1)} & \frac{1}{(n-1)(N-1)} \\ M(K_{11}^2) - \frac{1}{N} \kappa_{22} & \frac{1}{N-1} & \frac{N}{N-1} \\ M(K_{20} K_{02}) - \frac{1}{N} \kappa_{22} & 1 & \frac{2}{N-1} \end{vmatrix} = 0, \quad (13.20)$$

откуда

$$Dk_{11} = \frac{N-n}{(n-1)(N+1)(N-2)} \left\{ \frac{N-2}{Nn} (Nn - N - n - 1) K_{22} + \right. \\ \left. + (N-1) K_{20} K_{02} + (N-3) K_{11}^2 \right\}. \quad (13.21)$$

При выводе этого выражения мы, в соответствии с принципом Ирвина — Кендалла (см. 12.21), заменили математические ожидания, относящиеся к бесконечной генеральной совокупности, на математические ожидания, относящиеся к конечной генеральной совокупности объема N ; при этом члены вида $M(K_{11}^2)$ были заменены соответствующими постоянными значениями, в данном случае K_{11}^2 . Полученное выражение полезно сравнить с результатом упражнения 12.11.

Доказательство комбинаторных правил

13.4 Мы переходим теперь к доказательству комбинаторных правил, сформулированных и проиллюстрированных выше. Правила 1 и 2 уже были доказаны.

Предварительно определим оператор ∂_p , положив

$$\left. \begin{aligned} \partial_p \mu'_r &= r(r-1) \dots (r-p+1) \mu'_{r-p}, & r > p, \\ \partial_p \mu'_p &= p!, \\ \partial_p \mu'_r &= 0, & r < p, \end{aligned} \right\} \quad (13.22)$$

и

$$\partial_p (AB) = (\partial_p A) B + A (\partial_p B). \quad (13.23)$$

Свойство (13.23) означает дистрибутивность ∂ при действии на произведение.

В силу этого свойства

$$\partial_p (\mu'_r)^m = m (\mu'_r)^{m-1} \partial_p \mu'_r = \frac{\partial}{\partial \mu'_r} (\mu'_r)^m \partial_p \mu'_r.$$

Следовательно, если функция f — многочлен от моментов μ , или если f разлагается в ряд по таким многочленам, то

$$\partial_p f = \frac{\partial f}{\partial \mu'_1} \partial_p \mu'_1 + \frac{\partial f}{\partial \mu'_2} \partial_p \mu'_2 + \dots \quad (13.24)$$

Рассмотрим теперь выражение, определяющее семиинварианты в терминах моментов:

$$\exp \left(\kappa_1 t + \dots + \kappa_p \frac{t^p}{p!} + \dots \right) = 1 + \mu'_1 t + \dots + \frac{\mu'_p t^p}{p!} + \dots$$

Применяя к обеим его частям оператор ∂_p , получим

$$\begin{aligned} \exp\left(\kappa_1 t + \dots + \kappa_p \frac{t^p}{p!} + \dots\right) \left(\partial_p \kappa_1 t + \dots + \partial_p \kappa_p \frac{t^p}{p!} + \dots\right) = \\ = t^p + \mu'_1 t^{p-1} + \dots = t^p \left(1 + \mu'_1 t + \dots + \mu'_p \frac{t^p}{p!} + \dots\right), \end{aligned}$$

и, следовательно,

$$\partial_p \kappa_1 t + \dots + \partial_p \kappa_p \frac{t^p}{p!} + \dots = t^p.$$

Так как последнее соотношение тождественно относительно t , то

$$\left. \begin{aligned} \partial_p \kappa_p &= p!, \\ \partial_q \kappa_p &= 0, \quad q \neq p. \end{aligned} \right\} \quad (13.25)$$

Например,

$$\begin{aligned} \kappa_4 &= \mu'_4 - 4\mu'_3\mu'_1 - 3\mu'^2_2 + 12\mu'_2\mu'^2_1 - 6\mu'^4_1, \\ \partial_1 \kappa_4 &= 4\mu'_3 - 4\mu'_3 - 12\mu'_2\mu'_1 - 12\mu'_2\mu'_1 + 24\mu'^3_1 + 24\mu'_2\mu'_1 - 24\mu'^3_1 = 0, \\ \partial_2 \kappa_4 &= 12\mu'_2 - 24\mu'^2_1 - 12\mu'_2 + 24\mu'^2_1 = 0, \\ \partial_3 \kappa_4 &= 24\mu'_1 - 24\mu'_1 = 0, \\ \partial_4 \kappa_4 &= 4!. \end{aligned}$$

13.5 В соответствии с уже доказанным нами правилом 1 се-
мийнварианты $\kappa(a_1^{\alpha_1} \dots a_s^{\alpha_s})$ и, следовательно, моменты
 $\mu(a_1^{\alpha_1} \dots a_s^{\alpha_s})$ можно выразить в терминах семийнвариантов
исходного распределения

$$\mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots) = \sum \{A(\kappa_{b_1}^{\beta_1} \kappa_{b_2}^{\beta_2} \dots)\}, \quad (13.26)$$

где коэффициенты A пока неизвестны и наша задача их найти.
В целях экономии штрихи при моментах μ на протяжении всего
доказательства опускаются. Применим к обеим частям (13.26)
оператор $(\partial_{b_1}^{\beta_1} \partial_{b_2}^{\beta_2} \dots)$. Все стоящие справа члены, за исключением
 $(\kappa_{b_1}^{\beta_1} \kappa_{b_2}^{\beta_2} \dots)$, перейдут при этом в нуль, и мы получим

$$A(b_1!)^{\beta_1} (b_2!)^{\beta_2} \dots \beta_1! \beta_2! = (\partial_{b_1}^{\beta_1} \partial_{b_2}^{\beta_2} \dots) \mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots). \quad (13.27)$$

Рассмотрим теперь аналогичный оператору ∂_p оператор θ_p , ко-
торый, действуя на степень x^r (с любым индексом), понижает
показатель степени на p и производит умножение на $r(r-1) \dots$
 $\dots (r-p+1)$; предполагается также, что θ_p дистрибутивен*).

*) θ_p можно считать эквивалентным оператору

$$\left(\frac{\partial^p}{\partial x_1^p} + \frac{\partial^p}{\partial x_2^p} + \dots + \frac{\partial^p}{\partial x_n^p} \right).$$

Рассматривая $\mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$ как среднее значение произведения $(k_{a_1}^{\alpha_1} k_{a_2}^{\alpha_2} \dots)$, видим, что результат применения операторов ∂ к среднему значению, выраженному в терминах моментов μ , тот же самый, что и при взятии среднего значения от результата применения соответствующих операторов θ к произведению, выраженному в терминах иксов. Но в результате применения операторов θ получается константа, совпадающая, конечно, со своим средним значением. Следовательно,

$$A = \frac{(\theta_{b_1}^{\beta_1} \theta_{b_2}^{\beta_2} \dots)}{(b_1!)^{\beta_1} (b_2!)^{\beta_2} \dots \beta_1! \beta_2! \dots} (k_{a_1}^{\alpha_1} k_{a_2}^{\alpha_2} \dots). \quad (13.28)$$

Наши комбинаторные правила касаются вычислений, связанных с этими операторами.

13.6 Рассмотрим заполненную таблицу типа (12.32). Небольшое размышление показывает, что для каждого члена в (13.28), не переходящего в нуль при применении оператора, имеется такая таблица. Верно и обратное. Стоящие в таблице числа образуют степени иксов, входящих в произведение функций k ; если их сложить по строкам, то получатся порядки операторов, при сложении по столбцам получатся порядки соответствующих функций k . Заполненная таблица является, так сказать, схемой части операции (13.28); целиком операция представляется как сумма всех возможных заполненных таблиц.

Операция (13.28) дает коэффициенты в представлении $\mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$, но нас интересуют коэффициенты в представлении $\kappa(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$. Для их нахождения используется правило 3, которое мы теперь докажем. Согласно этому правилу коэффициент при $(\kappa_{b_1}^{\beta_1} \kappa_{b_2}^{\beta_2} \dots)$ в $\kappa(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$ задается всеми заполненными таблицами, за исключением тех, которые распадаются на отдельные блоки, образованные различными строками и столбцами.

Из уравнения (12.31), выражающего соотношение между многомерными моментами и семиинвариантами, следует, что $\kappa(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$ есть сумма $\mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$ и членов, образованных произведением двух, трех и т. д. многомерных моментов (с некоторыми коэффициентами). Рассмотрим член вида $\mu(a_1^{\alpha'_1} a_2^{\alpha'_2} \dots) \times \mu(a_1^{\alpha''_1} a_2^{\alpha''_2} \dots)$, где $\alpha'_1 + \alpha''_1 = \alpha_1$ и т. д. Коэффициент при нем в разложении правой части (12.31) равен

$$-\frac{1}{2} \frac{2!}{1!1!} \frac{t_{a_1}^{\alpha_1}}{\alpha'_1! \alpha''_1!} \frac{t_{a_2}^{\alpha_2}}{\alpha'_2! \alpha''_2!} \dots,$$

и, следовательно, коэффициент, с которым этот член входит в формулу для $\kappa(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$, есть

$$-\frac{\alpha_1!}{\alpha_1'! \alpha_1''!} \frac{\alpha_2!}{\alpha_2'! \alpha_2''!} \dots = -\binom{\alpha_1}{\alpha_1'} \binom{\alpha_2}{\alpha_2'} \dots \quad (13.29)$$

Далее, моменту $\mu(a_1^{\alpha_1'} a_2^{\alpha_2'} \dots)$ соответствует таблица типа (12.32) с суммарной строкой $(a_1^{\alpha_1'} a_2^{\alpha_2'} \dots)$ и суммарным столбцом, скажем $(b_1^{\beta_1'} b_2^{\beta_2'} \dots)$, и аналогичная таблица соответствует $\mu(a_1^{\alpha_1''} a_2^{\alpha_2''} \dots)$. Если $\beta_1' + \beta_1'' = \beta_1$ и т. д., то эти таблицы в выражениях через семиинварианты κ будут соответствовать тем членам, которые по умножении дадут член с $(\kappa_{b_1}^{\beta_1} \kappa_{b_2}^{\beta_2} \dots)$. Таким образом, произведение этих двух членов можно рассматривать как таблицу типа (12.32) с суммарной строкой $(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$ и суммарным столбцом $(b_1^{\beta_1} b_2^{\beta_2} \dots)$, распадающуюся на два отдельных блока. Поскольку имеется α_1 столбцов с суммами α_1 и т. д., то всего в выражении для $\mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$ будет $\binom{\alpha_1}{\alpha_1'} \binom{\alpha_2}{\alpha_2'} \dots$ произведений такого рода. Это число отличается от числа (13.29) знаком и имеет ту же абсолютную величину. Следовательно, если в таблице для μ отбрасывать двойные блоки, то тем самым будет учитываться произведение пар моментов, которые нужно вычесть из μ для получения κ .

Некоторые из этих отдельных блоков сами распадаются на два блока, и, вычитая их все из $\mu(a_1^{\alpha_1} a_2^{\alpha_2} \dots)$, мы вычитаем слишком много. Например, если имеются три отдельных блока L , M и N , то, рассматривая L и $(M+N)$ как два блока, мы вычли бы L , M , N . То же самое было бы вычтено, если рассматривать M , $(L+N)$ и N , $(L+M)$ как два блока. Таким образом, мы вычли бы излишне $2L$, $2M$, $2N$. Эти блоки нужно вернуть в таблицу для μ . Такое добавление, произведенное по всем тройным блокам, учитывает члены в разложении (12.31), получающиеся в результате произведения трех моментов.

Возвращая эти блоки, мы возвращаем слишком много в тех случаях, когда имеются четыре отдельных блока. Их следует опять отбросить, они соответствуют отрицательным членам в разложении (12.31), содержащим произведения четырех моментов. Рассуждая дальше таким же образом, убеждаемся в справедливости правила 3*).

*) Для обоснования этого правила Капланом (1952) были приведены также статистические соображения.

13.7 Докажем теперь правила 4, 5 и 6, которые являются основными для всего процесса. Рассмотрим опять таблицу типа (12.32). Для определенности пусть это будет

$$\begin{array}{ccc|c}
 2 & 3 & 1 & 6 \\
 1 & 1 & . & 2 \\
 1 & . & 1 & 2 \\
 \hline
 4 & 4 & 2 & 10
 \end{array} \quad (13.30)$$

Таблица (13.30) представляет некоторое число членов операции, каждый из которых состоит в действии θ_6 на $x^2.x^3.x$ (первая строка), а θ_2 на $x.x$ (вторая строка) и т. д. Если считать индексы иксов в каждой строке одинаковыми, то каждому индексу будет соответствовать некоторый член, поскольку k_p содержит члены с любым распределением степеней (дающих в сумме p) и индексов. Например, будут члены вида

$$\begin{array}{ccc}
 x_1^2 & x_1^3 & x_1 \\
 x_2 & x_2 & . \\
 x_3 & . & x_3
 \end{array}
 \quad
 \begin{array}{ccc}
 x_1^2 & x_1^3 & x_1 \\
 x_1 & x_1 & . \\
 x_2 & . & x_2
 \end{array}
 \quad
 \begin{array}{ccc}
 x_1^2 & x_1^3 & x_1 \\
 x_1 & x_1 & . \\
 x_1 & . & x_1
 \end{array}$$

Вообще для каждой заполненной таблицы имеются члены, в которых

(I) все иксы имеют один и тот же индекс (всего их n , один для каждого индекса),

(II) все иксы, за исключением одной строки, имеют один и тот же индекс (всего их $n(n-1)$),

(III) все иксы, за исключением двух строк, имеют один и тот же индекс, и в этих двух строках у иксов тоже одинаковый индекс (всего их $n(n-1)$) и т. д. Эти случаи соответствуют различным разбиениям на группы в правиле 6.

В случае (I) член, отвечающий любому столбцу, появляется из члена с x^p в k_p , и с точностью до числового множителя он в силу (12.25) равен n^{-1} . Следовательно, любой столбец, имеющий ненулевые элементы, дает множитель n^{-1} , и функция от n , появляющаяся в случае (I), есть произведение n на n^{-1} в степени, равной числу столбцов с ненулевыми элементами.

Подобно этому в случаях (II) и (III) функция от n для каждого разбиения и каждого столбца есть произведение $n(n-1)$ на n^{-1} или на $\frac{-1}{n(n-1)}$, в зависимости от того, имеет ли столбец в одной или в обеих группах разбиения ненулевые элементы, и т. д.

Тем самым объясняется, откуда взялась шаблонная функция, описанная в правиле 6. Однако чтобы доказать это правило полностью (и попутно также правила 4 и 5), нужно показать, что появляющиеся от каждого разбиения на группы числовые коэффициенты одни и те же. После того как это сделано, правило 6 доказано, так как вклады по n от разбиений можно сложить вместе (в результате чего получается шаблонная функция), и затем результат умножить на числовой коэффициент.

Действие оператора θ_1 можно рассматривать как выбрасывание всеми возможными способами x из выражения, на которое θ_1 действует, с заменой его на единицу. Аналогично $\frac{\theta_p}{p!}$ можно рассматривать как выбрасывание p иксов с одним и тем же индексом и замену их на единицу. При этом очевидно, что результат действия произведения операторов $\frac{\theta_{b_1}\theta_{b_2}\dots}{b_1!b_2!\dots}$ на произведение функций k той же степени есть число способов, которыми множества иксов могут быть выброшены из произведения функций k , причем каждое множество должно содержать b_1 иксов с первым индексом, b_2 — со вторым (второй индекс может совпадать с первым) и т. д.

Рассмотрим теперь операцию (13.28), предполагая, что функции k выражены в упрощенной форме (12.26). Напомним, что операторы θ дистрибутивны. Выделим из операции с суммой членов, включающей все возможные способы, которыми отдельные иксы могут быть выброшены из произведения функций k , те члены, которые отвечают суммарному столбцу и суммарной строке рассматриваемой таблицы. Рассмотрим множества, соответствующие определенной таблице, например (13.30). Общий вклад доставляется способами выбрасывания отдельных иксов так, что

(I) из элементов первой функции k_1 выбираются четыре в разбиении (2, 1, 1),

(II) из второй функции k_2 выбираются четыре в разбиении (3, 1),

(III) из функции k_3 выбираются два в разбиении (1, 1),

(IV) все это объединяется по всем возможным способам, при которых элементы в одной и той же строке имеют одинаковый индекс.

По рассмотрении возможностей становится ясно, что общее число способов осуществить эту процедуру равно числу способов размещения элементов суммарной строки в соответствии с правилом 5. И это верно независимо от того, имеют ли множества, образующие строки, одинаковый индекс или нет.

Отсюда немедленно вытекают правила 5 и 6 и, следовательно, правило 4.

13.8 Оставшиеся правила выводятся из рассмотренных выше.

Правило 7 следует из правила 2. В самом деле, шаблонная функция не зависит от чисел, заполняющих таблицу, и шаблон со строкой, содержащей один элемент, может, следовательно, быть основой для таблицы, в которой этот элемент равен единице. Отличие соответствующей шаблонной функции от нуля привело бы к появлению χ_1 , что согласно правилу 2 невозможно.

Правило 8 следует из правила 6. Если столбец имеет единственный ненулевой элемент, то при всех разбиениях на группы имеется в точности одна группа, в которой представлен этот ненулевой элемент, так что следствием наличия этого столбца является умножение всех вкладов на n^{-1} .

Правило 10 следует из правила 8. Добавление в аргумент χ единицы эквивалентно добавлению лишнего столбца, состоящего из одной единицы и нулей. Добавление такого столбца ведет к умножению всех шаблонных функций на n^{-1} , числовые коэффициенты оставляет неизменными и увеличивает индекс каждого χ соответственно строке, в которой находится единица.

13.9 Осталось доказать только правило 9. Отметим, что любая шаблонная функция может быть линейно выражена в терминах функций от шаблонов, получающихся в результате отбрасывания одного из столбцов. Например, рассмотрим правый столбец в шаблоне

$$\begin{array}{cccc} & \times & \times & \times \\ \times & . & \times & \times \\ \times & \times & . & . \\ \times & \times & . & . \end{array} \quad (13.31)$$

и тот вклад, который он дает в шаблонную функцию. Пятнадцать различных разбиений, возможных в случае четырех строк, можно разделить на два класса: в первый класс входят те разбиения, при которых два ряда четвертого столбца принадлежат одной и той же группе, во второй — разбиения, относящие эти ряды к разным группам. При разбиениях первого класса вклад от первых трех столбцов совпадает с вкладом, даваемым всеми разбиениями шаблона

$$\begin{array}{ccc} \times & \times & \times \\ \times & \times & . \\ \times & \times & . \end{array} \quad (13.32)$$

получающегося из первых трех столбцов объединением двух верхних строк. Вклад первых трех столбцов

$$\begin{array}{ccc} \cdot & \times & \times \\ \times & \cdot & \times \\ \times & \times & \cdot \\ \times & \times & \cdot \end{array} \quad (13.33)$$

при разбиениях второго класса состоит из всех вкладов, не определяемых разбиениями первого класса. Если A — вклад от первых трех столбцов при разбиениях первого класса и B — вклад при всех разбиениях, то общий вклад равен

$$\frac{1}{n} A - \frac{1}{n(n-1)} (B - A) = \frac{1}{n-1} A - \frac{1}{n(n-1)} B,$$

т. е. является линейной функцией от величин A и B . В общем случае доказательство следует той же схеме.

Если шаблон можно разделить на две группы, связанные только одним столбцом, то его можно редуцировать шаг за шагом, отбрасывая другие столбцы. В результате получится один столбец с шаблонной функцией, равной нулю. В самом деле, суммарная строка a соответствует функции k_a , среднее значение которой представляет эта одностолбцовая таблица. Но так как согласно определению среднее значение k_a есть κ_a , то не может появиться составных членов, которые имелись бы при двух или большем числе рядов.

13.10 Вышеизложенные рассуждения непосредственно обобщаются на двумерный случай. В дополнительном изучении нуждается только правило нахождения числовых коэффициентов, на которые умножаются шаблонные функции. Это изучение использует оператор

$$\partial_{pq} \mu_{rs} = r^{[p]} s^{[q]} \mu_{r-p, s-q} \quad (13.34)$$

и структуру двумерных k -статистик и приводит к выводу, что схема, набросанная в 13.2, остается применимой.

Без доказательства отметим также, что подобные методы применимы и к многомерным статистикам большей размерности. Для трехмерных статистик, например, нужно рассматривать трехчастные разбиения. Практически для этих более общих случаев требуются лишь простейшие результаты, и их обычно можно получить, используя одномерные и двумерные результаты и применяя рассуждения типа тех, которые развиваются в последующих пунктах.

Сводка результатов

13.11 Резюмируем результаты этой и предшествующей глав. Мы видели, что с помощью простых алгебраических методов можно найти сколь угодно много моментов и семиинвариантов любой статистики, симметрично зависящей от наблюдений. Было также отмечено, что работа значительно упрощается, если использовать k -статистики и семиинварианты.

Естественно спросить, за счет чего происходит это упрощение? Каковы те структурные свойства k -статистик и семиинвариантов, которые дают им эти преимущества, и не представляют ли они собой предела для возможных упрощений? Ответ на последний вопрос будет положительным, и поучительно посмотреть, почему это должно быть так.

13.12 Определение одномерных семиинвариантов в терминах моментов тождеством по t ,

$$\exp \{ \sum \kappa_p t^p / p! \} = \sum \mu_p t^p / p!, \quad (13.35)$$

не выявляет существа структуры отношений между семиинвариантами и моментами. Рассмотрим поэтому отношение между многомерными семиинвариантами и моментами, определенное формально тождеством (12.31). Предположим для определенности, что имеется четыре случайных величины, и найдем выражение для их смешанного момента μ'_{1111} через семиинварианты. Подсчитывая коэффициенты при t_1, t_2, t_3, t_4 , получим

$$\begin{aligned} \mu'_{1111} = & \kappa_{1111} + (\kappa_{1110}\kappa_{0001} + \kappa_{1101}\kappa_{0010} + \kappa_{1011}\kappa_{0100} + \kappa_{0111}\kappa_{1000}) + \\ & + (\kappa_{1100}\kappa_{0011} + \kappa_{1010}\kappa_{0101} + \kappa_{0110}\kappa_{1001}) + (\kappa_{1100}\kappa_{0010}\kappa_{0001} + \\ & + \kappa_{1010}\kappa_{0100}\kappa_{0001} + \kappa_{1001}\kappa_{0100}\kappa_{0010} + \kappa_{0110}\kappa_{1000}\kappa_{0001} + \\ & + \kappa_{0101}\kappa_{1000}\kappa_{0010} + \kappa_{0011}\kappa_{1000}\kappa_{0100}) + \kappa_{1000}\kappa_{0100}\kappa_{0010}\kappa_{0001}. \end{aligned} \quad (13.36)$$

В конечном счете для любых четырех случайных величин x_i, x_j, x_k, x_l можно написать

$$\mu'_{ijkl} = \kappa_{ijkl} + (\kappa_{ijfk}\kappa_l + \kappa_{ijlk}\kappa_k + \dots \text{ и т. д.}).$$

Эти формулы выявляют действительную простоту отношений между моментами и семиинвариантами. В выражении μ через κ каждая возможность в правой части встречается в точности один раз. Обращая соотношение (чтобы получить κ в терминах μ), имеем, аналогично (13.36),

$$\begin{aligned} \kappa_{1111} = & \mu'_{1111} - (\mu'_{1110}\mu'_{0001} + \mu'_{1101}\mu'_{0010} + \mu'_{1011}\mu'_{0100} + \mu'_{0111}\mu'_{1000}) - \\ & - (\mu'_{1100}\mu'_{0011} + \mu'_{1010}\mu'_{0101} + \mu'_{0110}\mu'_{1001}) + \\ & + 2(\mu'_{1100}\mu'_{0010}\mu'_{0001} + \dots) - 6\mu'_{1000}\mu'_{0100}\mu'_{0010}\mu'_{0001}. \end{aligned} \quad (13.37)$$

Формула такого типа отличается от (13.36) только множителями $(\rho-1)!(-1)^{\rho-1}$ перед членами, состоящими из ρ частей.

13.13 В частности, эти формулы верны, когда две или более из четырех случайных величин совпадают. Если, например, совпадают первая и вторая величины, то такие выражения, как κ_{1111} , записываются в виде κ_{211} и становятся трехмерными; если, кроме того, вторая и четвертая величины одинаковы, то этот семиинвариант переходит в κ_{22} и становится двумерным. В крайнем случае, когда все величины идентичны, он переходит в κ_4 . Таким образом, посредством объединения индексов из (13.36) и (13.37) можно получить все соотношения четвертого порядка для любого числа случайных величин. При этом некоторые первоначально различные члены переходят в одинаковые и могут быть объединены. Например, из (13.36) находим

$$\begin{aligned} \mu'_{22} = & \kappa_{22} + 2\kappa_{21}\kappa_{01} + 2\kappa_{12}\kappa_{10} + \kappa_{20}\kappa_{02} + 2\kappa_{11}^2 + \\ & + \kappa_{20}\kappa_{01}^2 + \kappa_{02}\kappa_{10}^2 + 4\kappa_{11}\kappa_{10}\kappa_{01} + \kappa_{10}^2\kappa_{01}^2. \end{aligned} \quad (13.38)$$

Еще раз объединяя индексы, получаем известную формулу для одномерного случая

$$\mu'_4 = \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4. \quad (13.39)$$

Теперь становится ясным происхождение числовых коэффициентов в (13.39). Формальные методы получения многомерных формул из одномерных (см., например, 3.28) являются просто сокращенными способами обращения процесса объединения индексов, что возможно в силу симметрии относительно индексов членов в (13.36).

13.14 Таким образом, полезность формальных определений типа (12.31) и (13.35), включающих в себя экспоненциальную функцию, проистекает из алгебраического свойства этой функции, проиллюстрированного в (13.36), а не из аналитических ее свойств, таких как совпадение со своей производной или простота связи с логарифмической функцией (для которой она является обратной). Числовые коэффициенты в соотношениях между семиинвариантами и моментами появляются в результате свертывания соотношений, структурно значительно более простых, чем они кажутся. Аналогичные рассуждения применимы к k -статистикам. Если отвлечься от множителей $(-1)^{\rho-1}(\rho-1)!/n^{\rho!}$, то эти статистики, как было отмечено в 12.8, являются просто суммами произведений переменных x , причем каждое произведение встречается один и только один раз.

13.15 Теперь мы можем записать формулы в свернутом виде. Например, (13.36) перейдет в

$$\mu_{ijkl} = \kappa_{ijkl} + \sum_4 \kappa_i \kappa_{jkl} + \sum_3 \kappa_i \kappa_{jk l} + \sum_6 \kappa_i \kappa_{j k l} + \kappa_i \kappa_{j k} \kappa_l, \quad (13.40)$$

где суммирования ведутся по всем способам группирования индексов, и число этих способов указывается над знаком суммы. (Обозначения здесь несколько изменены против обычных: μ_i , например, в этих обозначениях есть среднее значение i -й случайной величины, а не среднее значение i -й степени некоторой величины x .)

Подобным образом, обозначая символом $s_{ij\dots l}$ сумму произведений $x_i x_j \dots x_l$ по всей выборке, получим следующие выражения для k -статистик:

$$\left. \begin{aligned} k_i &= s_i/n, \\ k_{ij} &= (ns_{ij} - s_i s_j)/n^{[2]}, \\ k_{ijk} &= (n^2 s_{ijk} - n \sum s_i s_j s_k + 2s_i s_j s_k)/n^{[3]} \end{aligned} \right\} \quad (13.41)$$

и т. д.

Считая, например, i -ю и j -ю случайные величины идентичными, из второй формулы (13.41) получим (в более привычных обозначениях)

$$k_2 = (ns_2 - s_1^2)/(n(n-1)), \quad (13.42)$$

в то время как для двумерной статистики k_{11} имеем

$$k_{11} = (ns_{11} - s_{10}s_{01})/(n(n-1)). \quad (13.43)$$

Для нахождения k_{21} нужно в третьей формуле (13.41) объединить индексы i и j , что в результате дает

$$k_{21} = \{n^2 s_{21} - 2ns_{10}s_{11} - ns_{20}s_{01} + 2s_{10}^2 s_{01}\}/n^{[3]}$$

в соответствии с (13.2).

13.16 Эта система обозначений была введена Каплагом (1952) и названа им тензорными обозначениями. Она позволяет записать многомерные формулы в виде, занимающем лишь немного более места, чем его требуют одномерные формулы. Например, в четырехмерном случае для ковариации двух пар имеем

$$\begin{aligned} \kappa(ab, ij) &= M\{(k_{ab} - \kappa_{ab})(k_{ij} - \kappa_{ij})\} = \\ &= \frac{1}{n} \kappa_{abij} + \frac{1}{n-1} (\kappa_{ai} \kappa_{bj} + \kappa_{aj} \kappa_{bi}). \end{aligned} \quad (13.44)$$

Это выражение содержит в себе семь типов формул. Если все случайные величины совпадают (a, b, i, j относятся к одной и той же величине), то (13.44) сводится к

$$Dk_2 = \kappa(2^2) = \frac{1}{n} \kappa_4 + \frac{2}{n-1} \kappa_2^2. \quad (13.45)$$

Если a совпадает с i и b совпадает с j , то получим

$$Dk_{11} = \kappa \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{22} + \frac{1}{n-1} \kappa_{20} \kappa_{02} + \frac{1}{n-1} \kappa_{11}^2. \quad (13.46)$$

Если a совпадает с b , а i совпадает с j , то получим

$$\text{cov}(k_{20}, k_{02}) = \kappa \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \frac{1}{n} \kappa_{22} + \frac{2}{n-1} \kappa_{11}^2. \quad (13.47)$$

При совпадающих a, b, i и отличающемся j

$$\text{cov}(k_{20}, k_{11}) = \kappa \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{31} + \frac{2}{n-1} \kappa_{11} \kappa_{20}. \quad (13.48)$$

При совпадающих a, b и отличающихся от них и друг от друга i, j

$$\text{cov}(k_{200}, k_{011}) = \kappa \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{211} + \frac{2}{n-1} \kappa_{110} \kappa_{101}. \quad (13.49)$$

При совпадающих a, i и отличающихся от них и друг от друга b, j

$$\text{cov}(k_{110}, k_{101}) = \kappa \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{211} + \frac{1}{n-1} (\kappa_{200} \kappa_{011} + \kappa_{110} \kappa_{101}). \quad (13.50)$$

Наконец, если все a, b, i, j различны, то будем иметь

$$\text{cov}(k_{1100}, k_{0011}) = \kappa \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = \frac{1}{n} \kappa_{1111} + \frac{1}{n-1} (\kappa_{1010} \kappa_{0101} + \kappa_{1001} \kappa_{0110}). \quad (13.51)$$

13.17 Следующие формулы резюмируют дальнейшие многомерные результаты:

$$\kappa(ab, ijk) = \kappa_{abijk}/n + \sum^6 \kappa_{ai} \kappa_{bjk}/(n-1), \quad (13.52)$$

$$\kappa(ab, ijkl) = \kappa_{abijkl}/n + \left(\sum^8 \kappa_{ai} \kappa_{bjkl} + \sum^6 \kappa_{aj} \kappa_{bkl} \right) / (n-1), \quad (13.53)$$

$$\begin{aligned} \kappa(ab, ij, pq) &= \kappa_{abijpq}/n^2 + \sum^{12} \kappa_{ablp} \kappa_{jq} / \{n(n-1)\} + \\ &+ \sum^4 \kappa_{aip} \kappa_{bjq} (n-2) / \{n(n-1)^2\} + \sum^8 \kappa_{ai} \kappa_{bp} \kappa_{jq} / (n-1)^2, \end{aligned} \quad (13.54)$$

$$\begin{aligned} \kappa(abc, ijk) &= \kappa_{abcijk}/n + \left(\sum^9 \kappa_{ai} \kappa_{bcjk} + \sum^9 \kappa_{abi} \kappa_{cjk} \right) / (n-1) + \\ &+ \sum^6 \kappa_{ai} \kappa_{bj} \kappa_{ck} n / \{(n-1)(n-2)\}, \end{aligned} \quad (13.55)$$

$$\begin{aligned} \kappa(abc, ijkl) = & \kappa_{abcljkl}/n + \sum^{12} \kappa_{ai} \kappa_{bcjkl}/(n-1) + \\ & + \left(\sum^{12} \kappa_{abi} \kappa_{cjk} + \sum^{18} \kappa_{aij} \kappa_{bckl} \right) / (n-1) + \\ & + \sum^{36} \kappa_{ai} \kappa_{bj} \kappa_{ckl} n / \{(n-1)(n-2)\}, \quad (13.56) \end{aligned}$$

$$\begin{aligned} \kappa(ab, ij, pq, uv) = & \kappa_{abijpquv}/n + \sum^{24} \kappa_{ai} \kappa_{bjpquv} / \{n^2(n-1)\} + \\ & + \sum^{32} \kappa_{aip} \kappa_{bjquv} (n-2) / \{n^2(n-1)^2\} + \\ & + \sum^8 \kappa_{aipu} \kappa_{bjqv} (n^2 - 3n + 3) / \{n^2(n-1)^3\} + \\ & + 3 \sum^{24} \kappa_{abpu} \kappa_{ijqv} / \{n^2(n-1)\} + \\ & + \left(\sum^{96} \kappa_{ai} \kappa_{bp} \kappa_{jqv} + \sum^{48} \kappa_{ai} \kappa_{pu} \kappa_{bjqv} \right) / \{n(n-1)^2\} + \\ & + \sum^{96} \kappa_{ai} \kappa_{bpu} \kappa_{jqv} (n-2) / \{n(n-1)^3\} + \sum^{48} \kappa_{bi} \kappa_{jp} \kappa_{qu} \kappa_{va} / (n-1)^3, \quad (13.57) \end{aligned}$$

$$\begin{aligned} \kappa(abcd, ijkl) = & \kappa_{abcdijkl}/n + \sum^{16} \kappa_{ai} \kappa_{bcdjkl} / (n-1) + \\ & + \sum^{48} \kappa_{abi} \kappa_{cdjkl} / (n-1) + \sum^{72} \kappa_{ai} \kappa_{bj} \kappa_{cdkl} n / \{(n-1)(n-2)\} + \\ & + \left(\sum^{16} \kappa_{aijk} \kappa_{bcdl} + \sum^{18} \kappa_{abl} \kappa_{cdkl} \right) / (n-1) + \\ & + \sum^{141} \kappa_{ai} \kappa_{bcj} \kappa_{dkl} / \{(n-1)(n-2)\} + \\ & + \sum^{24} \kappa_{ai} \kappa_{bj} \kappa_{ck} \kappa_{dl} n (n+1) / \{(n-1)(n-2)(n-3)\}. \quad (13.58) \end{aligned}$$

Эти формулы включают в себя 41 формулу, полученную Кук (1951), а также и ряд других формул, однако, как заметил Каплан, при выводе из них формул для частных случаев необходима осторожность.

Пример 13.3

Для дисперсии статистики k_{11} в случае бесконечной генеральной совокупности имеем (см. пример 13.2)

$$Dk_{11} = \frac{1}{n} \kappa_{22} + \frac{1}{n-1} \kappa_{20} \kappa_{02} + \frac{1}{n-1} \kappa_{11}^2.$$

Если генеральная совокупность нормальна и компоненты имеют единичные дисперсии, то эта формула переходит в

$$Dk_{11} = \frac{1}{n-1} (1 + \rho^2). \quad (13.59)$$

Найдем третий и четвертый семиинварианты величины k_{11} .

Третий семиинвариант получается из (13.54), если индексы в каждой из троек a, i, p и b, j, q считать совпадающими. Имеем

$$\begin{aligned} \kappa \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} &= \frac{1}{n^2} \kappa_{33} + \frac{1}{n(n-1)} \{6\kappa_{22}\kappa_{11} + 3\kappa_{20}\kappa_{13} + 3\kappa_{31}\kappa_{02}\} + \\ &+ \frac{n-2}{n(n-1)^2} \{3\kappa_{21}\kappa_{12} + \kappa_{30}\kappa_{03}\} + \frac{1}{(n-1)^2} \{2\kappa_{11}^3 + 6\kappa_{11}\kappa_{20}\kappa_{02}\}. \end{aligned} \quad (13.60)$$

В нормальном случае остается только один член, и мы получаем

$$\kappa_3(k_{11}) = \frac{1}{(n-1)^2} (2\rho^3 + 6\rho). \quad (13.61)$$

Аналогично, полагая в (13.57) индексы в каждой из четверок a, i, p, u и b, j, q, v совпадающими, для нормального случая находим

$$\kappa_4(k_{11}) = \frac{1}{(n-1)^3} \{36\rho^2 + 6 + 6\rho^4\} \quad (13.62)$$

(все члены в (13.62) происходят из последнего члена в (13.57): только он дает ненулевые члены). Из формул (13.59) и (13.62) получается следующее выражение для эксцесса распределения k_{11} :

$$\beta_2 - 3 = \frac{6}{n-1} \left(1 + \frac{4\rho^2}{(1+\rho^2)^2} \right). \quad (13.63)$$

УПРАЖНЕНИЯ

13.1 Показать, что коэффициент корреляции между k_{30} и k_{03} для выборок любого объема n из двумерной нормальной генеральной совокупности равен ρ^3 , где ρ — коэффициент корреляции компонент исходного распределения.

13.2 Показать, что в условиях упражнения 13.1 коэффициент корреляции между k_{r0} и k_{0r} равен ρ^r для всех r .

13.3 Случайные величины x_1, x_2, x_3, x_4 имеют единичные дисперсии, совместное распределение этих величин нормально и коэффициент корреляции между x_i и x_j равен ρ_{ij} . Показать, что коэффициент корреляции между

$$(x_1 - Mx_1)(x_2 - Mx_2) \quad \text{и} \quad (x_3 - Mx_3)(x_4 - Mx_4)$$

дается формулой

$$\frac{\rho_{13}\rho_{24} + \rho_{14}\rho_{23}}{\sqrt{(1+\rho_{12}^2)(1+\rho_{34}^2)}}.$$

13.4 Вывести формулу

$$\kappa \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \frac{1}{n^2} \kappa_{33} + \frac{6}{n(n-1)} \kappa_{22} \kappa_{11} + \frac{3}{n(n-1)} \kappa_{20} \kappa_{13} + \frac{3}{n(n-1)} \kappa_{31} \kappa_{02} + \\ + \frac{3(n-2)}{n(n-1)^2} \kappa_{21} \kappa_{12} + \frac{n-2}{n(n-1)^2} \kappa_{30} \kappa_{03} + \frac{2}{(n-1)^2} \kappa_{11}^3 + \frac{6}{(n-1)^2} \kappa_{11} \kappa_{20} \kappa_{02}. \\ \text{(Кук, 1951.)}$$

13.5 Используя методы этой главы, доказать, что в случае больших выборок из двумерной нормальной совокупности дисперсию выборочного коэффициента корреляции можно вычислять по формуле

$$Dr = (1 - \rho^2)^2 / n.$$

13.6 Случайная величина x имеет семинварианты всех порядков, и для всех r $\kappa_r = O(n^{-r+1})$. Показать, что преобразованная случайная величина $y = f(x)$ имеет семинварианты всех порядков, тоже удовлетворяющих этому условию. (Это не мешает семинвариантам иметь меньший порядок, например, $o(n^{-r+1})$.)

(Джеймс (James), 1955, *Biometrika* 42, 529.)

13.7 Показать, что для двумерной нормальной генеральной совокупности ковариация статистик k_{tu} и k_{vw} равна нулю при $t+u \neq v+w$.

(Уишарт, 1929.)

13.8 Показать, что для двумерного нормального распределения

$$dF = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left(\frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2} \right) \right\} dx_1 dx_2$$

дисперсия k_{tu} равна

$$Dk_{tu} = \kappa \begin{pmatrix} t & t \\ u & u \end{pmatrix} = t! u! \sum_{j=1}^{t+u} \frac{(j-1)!}{j} \frac{\Delta^j 0^{t+u}}{n^{[j]}} \sigma_1^{2t} \sigma_2^{2u} F(-t, -u, 1, \rho^2),$$

где $\Delta^j 0^k$ есть j -я разность k -й степени нуля *) и F — гипергеометрическая функция.

(Уишарт, 1929.)

13.9 В обозначениях 13.15 показать, что

$$k_{ijkl} = \left\{ n^2(n+1) s_{ijkl} - n(n+1) \sum^4 s_{ijskl} - n(n-1) \sum^3 s_{ijskl} + \right. \\ \left. + 2n \sum^6 s_{ijsksl} - 6s_{ijsksl} \right\} / n^{[4]}.$$

Вывести отсюда формулы (13.2) для k_{31} и k_{22} .

(Каплан, 1952.)

13.10 Используя результат 13.9, показать, что в случае нормальной генеральной совокупности добавление новой части 2 к $\kappa(a^{a_1} \dots a_s^{a_s})$ ведет к умножению всех шаблонных функций на $1/(n-1)$; проверить, что при этом

*) Под этим здесь понимается, как обычно, $\Delta^j x^k|_{x=0}$. (Прим. перев.)

числовой коэффициент умножается на удвоенное число строк таблицы, которой он соответствует. Вывести отсюда, что результат добавления новой части 2 равносильно применению оператора

$$\frac{2\kappa_2^2}{n-1} \frac{d}{d\kappa_2}.$$

(Фишер и Уишарт, 1931.)

13.11 Использовать предыдущее упражнение для доказательства формул (12.71) и (12.72).

13.12 Обобщая результат упражнения 13.10, показать, что в случае многомерной нормальной генеральной совокупности результат добавления ковариации k_{pq} (p, q относятся к p -й и q -й случайным величинам) равносильно применению оператора

$$\sum_{rs} \frac{1}{n-1} (\kappa_{pr}\kappa_{qs} + \kappa_{ps}\kappa_{qr}) \frac{d}{d\kappa_{rs}},$$

где κ_{pq} — ковариация p -й и q -й величин.

(Фишер и Уишарт, 1931.)

13.13 Показать, что для выборок из нормальной генеральной совокупности

$$D\left(m_r/m_2^{\frac{1}{2}r}\right) \sim r! / n.$$

ПОРЯДКОВЫЕ СТАТИСТИКИ

14.1 В главах 10 и 11 был кратко изложен ряд свойств выборочной медианы, квантилей, а также более общих порядковых статистик от результатов наблюдений. Напомним, что r -я порядковая статистика выборки объема n есть r -е наименьшее выборочное значение, которое будем обозначать $x_{(r)}$. В отличие от статистик, являющихся симметрическими функциями от наблюдений, к порядковым статистикам неприменимы методы глав 12 и 13. Эти статистики обладают тем отличительным свойством, что если $x_1, \dots, x_n$ — результаты наблюдений, а $x_{(1)}, \dots, x_{(n)}$ — те же результаты, но расположенные в порядке возрастания их величин, то $x_{(1)}, \dots, x_{(n)}$ уже не будут независимыми (даже если первоначальные наблюдения были независимыми). Это замечание и объясняет специальный характер выборочных распределений порядковых статистик.

14.2 Несмотря на отсутствие симметрии и независимости, можно получить некоторые исключительно простые результаты. Мы видели в 10.10, что в выборке объема n из генеральной совокупности с функцией распределения $F(x)$, обладающей непрерывной плотностью $f(x)$, распределение G_r случайной величины $x_{(r)}$ выражается формулой

$$dG_r = \frac{\{F(x_{(r)})\}^{r-1} \{1 - F(x_{(r)})\}^{n-r} dF(x_{(r)})}{B(r, n-r+1)}. \quad (14.1)$$

Аналогичным образом можно получить и совместное распределение $x_{(r)}$ и $x_{(s)}$, $r < s$:

$$dG_{r,s} = \frac{\{F(x_{(r)})\}^{r-1} \{F(x_{(s)}) - F(x_{(r)})\}^{s-r-1} \{1 - F(x_{(s)})\}^{n-s} dF(x_{(r)}) dF(x_{(s)})}{B(r, s-r) B(s, n-s+1)}. \quad (14.2)$$

Обобщение на случай нескольких порядковых статистик не вызывает трудностей. Так, например, для $x_{(r)}, x_{(s)}, x_{(t)}, x_{(u)}$, $r < s <$

$t < u$, мы имеем (обозначая F_r вместо $F(x_{(r)})$ и т. д.)

$$dG_{r,s,t,u} = \frac{F_r^{r-1} (F_s - F_r)^{s-r-1} (F_t - F_s)^{t-s-1} (F_u - F_t)^{u-t-1} (1 - F_u)^{n-u} dF_r dF_s dF_t dF_u}{B(r, s-r) B(s, t-s) B(t, u-t) B(u, n-u+1)}. \quad (14.3)$$

14.3 Распределения типа (14.1) были изучены для различных семейств исходных распределений $F(x)$. Когда истинное распределение является равномерным на отрезке $[0, 1]$, то $F(x) = x$ и распределение для $x_{(r)}$ особенно просто: оно сводится к бета-распределению. К относительно простым результатам приводят и распределения, являющиеся экспоненциальными: $dF = \exp(-x/\sigma) d(x/\sigma)$. В других случаях (в частности, для нормального распределения) нахождение точного результата сводится, вообще говоря, к квадратурам некоторых интегралов.

Медиана и квартили в нормальном случае

14.4 Предположим, что рассматривается выборка нечетного объема $n = 2r + 1$. В этом случае выборочной медианой является $(r+1)$ -я порядковая статистика. Пусть

$$A_x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt. \quad (14.4)$$

Тогда, учитывая (14.1) и полагая $A = A_x$, распределение медианы можно выразить формулой

$$dG = n! A^r (1 - A)^r dA / (r!)^2. \quad (14.5)$$

В силу симметрии среднее значение выборочной медианы равно нулю и, следовательно,

$$Dm_e = \int x^2 dG = \frac{n!}{(r!)^2} \int_{-\infty}^{\infty} x^2 A^r (1 - A)^r \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx. \quad (14.6)$$

Разлагая $(1 - A)^r$ по формуле бинома, получим

$$Dm_e = \frac{n!}{(r!)^2} \int_{-\infty}^{\infty} \sum_{j=0}^r (-1)^j \binom{r}{j} A^{r+j} \frac{1}{\sqrt{2\pi}} x^2 e^{-\frac{1}{2}x^2} dx. \quad (14.7)$$

Каждое слагаемое можно проинтегрировать по частям, беря $x e^{-\frac{1}{2}x^2}$ в качестве одной из частей. Поэтому

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x A^{r+j} x e^{-\frac{1}{2}x^2} dx = \left[\frac{-x A^{r+j} e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right]_{-\infty}^{\infty} + \\ + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \left\{ A^{r+j} + (r+j) A^{r+j-1} \frac{1}{\sqrt{2\pi}} x e^{-\frac{1}{2}x^2} \right\} dx. \quad (14.8)$$

Первый член справа обращается в нуль. Интегрируя далее по частям третий член, мы получим для интеграла в левой части (14.8) выражение

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} A^{r+j} dx + \frac{(r+j)(r+j+1)}{4\pi} \int_{-\infty}^{\infty} A^{r+j-2} \exp\left(-\frac{3}{2}x^2\right) dx. \quad (14.9)$$

В (14.9) первый интеграл равен $1/(r+j+1)$. Подставляя это значение в (14.7), найдем, что

$$Dm_e = \frac{n!}{(r!)^2} \sum_{j=0}^r (-1)^j \binom{r}{j} \frac{1}{r+j+1} + \\ + \frac{n!}{(r!)^2 4\pi} \sum_{j=0}^r (-1)^j \binom{r}{j} (r+j)(r+j+1) \int_{-\infty}^{\infty} A^{r+j-2} \exp\left(-\frac{3}{2}x^2\right) dx. \quad (14.10)$$

Первая сумма в правой части (14.10) равна единице, поскольку

$$\sum_{j=0}^r (-1)^j \binom{r}{j} \frac{1}{r+j+1} = \int_0^1 \sum_{j=0}^r (-1)^j \binom{r}{j} t^{r+j} dt = \\ = \int_0^1 t^r (1-t)^r dt = B(r+1, r+1) = \frac{(r!)^2}{n!}. \quad (14.11)$$

Поэтому из (14.10) и (14.11) следует, что

$$Dm_e = 1 + \frac{n!}{4\pi (r!)^2} \sum_{j=0}^r (-1)^j \binom{r}{j} (r+j)(r+j+1) T_{r+j-2}, \quad (14.12)$$

где

$$T_{r+j-2} = \int_{-\infty}^{\infty} A^{r+j-2} \exp\left(-\frac{3}{2}x^2\right) dx. \quad (14.13)$$

14.5 Подобным образом мы могли бы получить любое количество моментов порядковых статистик. Результаты будут зависеть от интегралов типа

$$\int_{-\infty}^{\infty} A^p \exp\left(-\frac{1}{2}qx^2\right)dx, \quad (14.14)$$

частным случаем которых является (14.13). Эти интегралы известны как интегралы Хойо, по имени автора, который впервые (1931) детально их изучил и табулировал для некоторых p и q . Путем простого обобщения рассмотренного метода можно получить смешанные моменты двух порядковых статистик и, следовательно, найти моменты таких величин, как интерквартильная широта (величина, равная разности между верхним и нижним квартилями) и середина размаха (полусумма наибольшего и наименьшего значений). Этот подход чрезвычайно трудоемкий, и для обычных целей вполне достаточны приближенные методы, рассматриваемые ниже. Можно, однако, в качестве иллюстрации привести некоторые точные результаты Хойо.

14.6 Для нормального распределения стандартная ошибка выборочного среднего равна $\sigma/\sqrt{n}$. Стандартная ошибка выборочной медианы $c_n\sigma/\sqrt{n}$ всегда больше, чем $\sigma/\sqrt{n}$, так как $c_n \geq 1$ (равенство достигается только в тривиальном случае $n=2$). Определяя выборочную медиану при четном $n=2r$ формулой $m_e = \frac{1}{2}(x_{(r)} + x_{(r+1)})$, мы имеем следующие значения для c_n :

n	c_n	n	c_n
2	1,000	10	1,177
4	1,092	12	1,189
6	1,135	20	1,214
8	1,160	∞	1,253

Аналогично середина размаха $\frac{1}{2}(x_{(1)} + x_{(n)})$ имеет стандартную ошибку $d_n\sigma/\sqrt{n}$, где d_n стремится к бесконечности с ростом n , как это станет ясно из дальнейшего (пример 14.4). С увеличением объема выборки эффективность середины размаха (как оценки среднего значения нормальной совокупности) стремится

к нулю. Приведем значения d_n при некоторых малых n :

n	d_n	n	d_n
2	1,000	10	1,362
4	1,092	20	1,691
6	1,190	∞	∞

14.7 Интерквартильная ширина, деленная на удвоенную величину 0,67449 (расстояние между квартилями нормального распределения с единичной дисперсией), иногда употребляется как оценка истинного стандартного отклонения. Интересно сравнить стандартную ошибку этой величины (обозначим эту стандартную ошибку $e_n\sigma/\sqrt{2n}$) со стандартной ошибкой выборочного стандартного отклонения, равной $\sigma/\sqrt{2n}$. Ниже указаны значения e_n для некоторых n :

n	e_n	n	e_n
2	1,000	10	1,497
4	1,047	12	1,419
6	1,421	∞	1,649
8	1,313		

Нерегулярность в поведении этих значений до некоторой степени объясняется произвольностью в определении квартилей в случае, когда общее число наблюдений n не представимо в виде $4r+1$.

14.8 Хойо (1931, 1933) и К. Пирсон (1931) получили в этом направлении ряд других результатов. С более поздними работами вычислительного характера можно ознакомиться по статьям Хэстингса и др. (1947) и Годвина (1949). Несколько позже в настоящей главе будут даны ссылки на работы, посвященные крайним значениям.

Пирсоновское разложение

14.9 В главе 6 мы рассматривали разложение функции распределения $F(x)$ или ее плотности в ряд по степеням x . Сейчас мы рассмотрим данное К. Пирсоном обратное представление x по значениям F . Пусть X_r — истинное значение такое, что

$$F(X_r) = \frac{r}{n+1}. \quad (14.15)$$

Разложим $x_{(r)}$ в ряд Тейлора в окрестности точки X_r :

$$x_{(r)} = X_r + h_r X'_r + \frac{1}{2!} h_r^2 X''_r + \dots, \quad (14.16)$$

где

$$\left. \begin{aligned} h_r &= F(x_{(r)}) - F(X_r) = F_r - r/(n+1), \\ X'_r &= \frac{dX_r}{dF} = \frac{dx}{dF} \Big|_{x=X_r}, \\ X''_r &= \frac{d^2x}{dF^2} \Big|_{x=X_r}. \end{aligned} \right\} \quad (14.17)$$

С помощью (14.16) можно выразить степени x в виде ряда по степеням величины h , а с помощью (14.1) легко затем получить усреднение степеней h . Тогда в предположении, что наши ряды сходятся подходящим образом (или, точнее, дают хорошую аппроксимацию в некотором асимптотическом смысле), мы можем получить аппроксимацию для столь большого числа моментов порядковых статистик, сколько мы пожелаем.

Пример 14.1

Снова рассмотрим распределение медианы, т. е. распределение $(r+1)$ -го выборочного значения выборки нечетного объема $n=2r+1$ из нормальной совокупности.

Поскольку $(r+1)/(n+1) = \frac{1}{2}$ и $X_{r+1} = 0$, то

$$X'_r = \frac{dx}{dF} \Big|_{x=0} = 1 / \frac{dF}{dx} \Big|_{x=0} = \sqrt{2\pi},$$

$$X''_r = \frac{d}{dF} \left(1 / \frac{dF}{dx} \right) \Big|_{x=0} = \frac{dx}{dF} \cdot \frac{d}{dx} \left(1 / \frac{dF}{dx} \right) \Big|_{x=0} = \sqrt{2\pi} \cdot 0 = 0.$$

И аналогично $X'''_r = (2\pi)^{3/2}$. Поэтому из (14.16) и (14.4) при $h_{r+1} = A - 1/2$ получим

$$X_{me} = x_{(r+1)} = (2\pi)^{1/2} \left(A - \frac{1}{2} \right) + \frac{1}{6} (2\pi)^{3/2} \left(A - \frac{1}{2} \right)^3 + \dots$$

и

$$x_{me}^2 = 2\pi \left(A - \frac{1}{2} \right)^2 + \frac{1}{3} (2\pi)^2 \left(A - \frac{1}{2} \right)^4 + \dots$$

Подставляя разложение для x_{me}^2 в (14.6), найдем

$$Dm_e = \frac{n!}{(r!)^2} \int_{-\infty}^{\infty} A'(1-A)' \left\{ 2\pi \left(A - \frac{1}{2} \right)^2 + \frac{1}{3} (2\pi)^2 \left(A - \frac{1}{2} \right)^4 \right\} dA,$$

что сводится к выражению

$$Dm_e = \frac{\pi}{2(n+2)} + \frac{\pi^2}{4(n+2)(n+4)}. \quad (14.18)$$

Для обычных целей точность этой формулы вполне достаточна. При $n=11$ согласно (14.18) $Dm_e=0,137$, что совпадает в трех знаках с известным истинным значением 0,137227.

Пример 14.2

При равномерном распределении

$$dF=dx, \quad 0 \leq x \leq 1,$$

для медианы выборки нечетного объема мы получим

$$X'=1, \quad X''=X'''=\dots=0.$$

Дисперсия медианы поэтому совпадает со средним значением $(x - \frac{1}{2})^2$ для распределения

$$dG = \frac{n!}{(r!)^2} x^r (1-x)^r dx, \quad 0 \leq x \leq 1,$$

что приводит к следующему точному результату:

$$Dm_e = \frac{1}{4(n+2)}. \quad (14.19)$$

14.10 Дэвид и Джонсон (1954) исследовали этот круг вопросов более детально и привели разложения для семинвариантов и смешанных семинвариантов (до четвертого порядка включительно) порядковых статистик с точностью до третьей степени от $1/(n+2)$. Они выбрали разложение по степеням $(n+2)^{-1}$, а не n^{-1} по той причине, что в элементарных случаях (ср. с (14.18) и (14.19)) появляется именно $(n+2)$, а не n . Неясно, насколько эти разложения обладают требуемыми асимптотическими свойствами, но во всяком случае они себя хорошо оправдывают на практике.

Знание средних значений порядковых статистик $x_{(r)}$ в выборке объема n из нормальной совокупности часто бывает пугным при изучении свойств порядковых статистик, и поэтому такие средние значения были достаточно полно табулированы. См., например, *Biometrika Tables for Statisticians*, Part 1, где даны для нормальной совокупности средние значения всех порядковых статистик при $n=2(1)26(2)50$ (по меньшей мере с двумя десятичными знаками). Хартер (Harter, 1961, *Biom.* 48, 151) привел эти значения для $n=2(1)100(25)250(50)400$ с пятью десятичными знаками*). Тейкроу (1956) дал таблицу средних значений, а Сархан и Гринберг (1956) — таблицу дисперсий и ковариаций порядковых статистик в нормальном случае*) с десятью знаками для $n=2(1)20$.

14.11 При рассмотрении асимптотических распределений порядковых статистик мы можем получить два типа предельных

*) С некоторыми сокращениями эти таблицы воспроизведены в книге *Сборник статистических таблиц*. (Прим. ред.)

переходов. Первый тип получается, если считать r и n стремящимися к бесконечности таким образом, чтобы отношение r/n оставалось постоянным. Тем самым этот тип приводит к распределению квантилей выборки. Второй тип мы получим, если r остается фиксированным при $n \rightarrow \infty$. В этом случае мы приходим к результатам, по своему характеру совершенно иным, чем для квантилей. Это направление исследований часто называют теорией крайних (экстремальных) значений. Объясняется это название тем, что здесь наиболее интересны те случаи, когда r близко к 1 или n .

14.12 Мы уже получили в 10.10 и 10.11 ряд предельных результатов для квантилей и видели, что выборочное распределение квантиля $x_{(r)}$ асимптотически нормально. Его среднее значение равно x_r , причем $F(x_r) = r/n$, а дисперсия выражается отношением

$$p_r q_r / (n f_r^2), \quad (14.20)$$

где $p_r = r/n$, $q_r = 1 - p_r$ и f_r, n — ордината плотности распределения в точке X_r . Аналогично распределение двух квантилей $x_{(r)}$ и $x_{(s)}$ асимптотически нормально с дисперсиями (14.20) и ковариацией

$$p_r q_s / (n f_r, n f_s, n). \quad (14.21)$$

Этот результат обобщается и на случай более чем двух квантилей; при этом дисперсии и ковариации выражаются формулами типа (14.20) и (14.21), а предельное многомерное распределение будет нормальным.

Асимптотические распределения крайних значений

14.13 Остановимся теперь на асимптотической теории крайних значений, с которой до сих пор нам не приходилось сталкиваться*). Из (14.1), сделав преобразование $y = nF(x)$ ($dF = dy/n$), мы получим распределение для y :

$$dH_r(y) = \left(\frac{y}{n}\right)^{r-1} \left(1 - \frac{y}{n}\right)^{n-r} d(y/n) / B(r, n-r+1). \quad (14.22)$$

Тем самым отношение y/n имеет бета-распределение первого рода. Когда n стремится к бесконечности, то при каждом фиксированном r

$$\lim_{n \rightarrow \infty} dH_r(y) = y^{r-1} e^{-y} dy / \Gamma(r), \quad (14.23)$$

*) Весьма полное изложение этого вопроса можно найти в книге Гамбела (E. J. Gumbel) *Statistics of Extremes*, Columbia U. P., 1958. [Имеется русский перевод с другой транслитерацией фамилии автора: Э. Гумбель (1965), *Статистика экстремальных значений*, М., Изд-во «Мир», (Прим. ред.)]

что является элементом гамма-распределения с параметром r . В случае $r=1$ из (14.23) мы получаем, что трансформированное наименьшее значение имеет экспоненциальное распределение.

Если теперь мы преобразуем (14.1), положив $y=n(1-F(x))$, то получим результат, аналогичный (14.23), где вместо $r=1$ надо положить $n-r$. В случае $r=n$ также получается экспоненциальное распределение.

Эти результаты относились к переменной y . Теперь мы исследуем распределение и интересующей нас величины $x_{(r)}$. Если $F(x)$ известно точно, то преобразование, приведшее к результату (14.23), можно обратить и, таким образом, получить предельное распределение для $x_{(r)}$.

14.14 Для крайних значений ($r=1$ или n) распределение можно получить прямым путем, используя идеи Фишера и Типпетта (1928). Рассмотрим наибольшее значение и предположим, что $n=km$. Наибольшее значение $x_{(n)}$ тогда можно рассматривать как наибольшее из k наибольших значений в выборках объема m . Если $x_{(n)}$ имеет предельное распределение, то оно должно с точностью до сдвига и величины масштаба совпадать с распределением для наибольшего значения выборки объема m при $m \rightarrow \infty$ и фиксированном k . Поэтому если $G(x)$ есть функция предельного распределения, то она должна удовлетворять функциональному уравнению

$$G^k(x) = G(a_k x + b_k). \quad (14.24)$$

Этим условием характеризуются возможные формы предельного распределения. Для решения уравнения (14.24) мы воспользуемся методом, предложенным Дженкинсоном (1955). Положим

$$q(x) = -\log \{-\log G(x)\}. \quad (14.25)$$

Тогда, дважды логарифмируя (14.24), получим

$$\log k = q(x) - q(a_k x + b_k). \quad (14.26)$$

Разложим $q(x)$ и $q(a_k x + b_k)$ в ряд Тейлора в окрестности точки x_{0n} , для которой

$$q(x_{0n}) = 0. \quad (14.27)$$

Вместо (14.26) тогда получим

$$\log k = -q(a_k x_{0n} + b_k) + \sum_{r=1}^{\infty} \{q^{(r)}(x_{0n}) - q^{(r)}(a_k x_{0n} + b_k)\} \frac{(x - x_{0n})^r}{r!}. \quad (14.28)$$

Соотношение (14.28) является тождеством относительно x , поэтому

$$\log k = -q(a_k x_{0n} + b_k) \quad (14.29)$$

и

$$q^{(r)}(x_{0n}) = q^{(r)}(a_k x_{0n} + b_k), \quad r \geq 1. \quad (14.30)$$

Из (14.30) следует, что

$$\frac{q^{(r)}(a_k x_{0n} + b_k)}{\{q'(a_k x_{0n} + b_k)\}^r} = \frac{q^{(r)}(x_{0n})}{\{q'(x_{0n})\}^r} = c_r, \quad (14.31)$$

где c_r не зависят от k . Поскольку a_k и b_k — постоянные, зависящие от k , которое произвольно, (14.31) можно переписать в виде

$$q^{(r)}(x) = c_r \{q'(x)\}^r. \quad (14.32)$$

Дифференцируя формулу (14.32) по x и используя ее затем при $r=2$, получим

$$q^{(r+1)}(x) = r c_r \{q'(x)\}^{r-1} q''(x) = r c_r c_2 \{q'(x)\}^{r+1}. \quad (14.33)$$

Сравнение (14.32) и (14.33) показывает, что

$$c_{r+1} = r c_r \cdot c_2 = r! c_2^r, \quad r \geq 1. \quad (14.34)$$

Подставляя теперь (14.34) в (14.32), получим

$$q^{(r)}(x) = (r-1)! c_2^{r-1} \{q'(x)\}^r, \quad r \geq 1. \quad (14.35)$$

Если эту формулу использовать при разложении $q(x)$ в ряд Тейлора, то, учитывая (14.27), мы найдем

$$\begin{aligned} q(x) &= \sum_{r=1}^{\infty} c_2^{r-1} \{q'(x_{0n})\}^r (x - x_{0n})^r / r = \\ &= -\frac{1}{c_2} \log [1 - c_2 q'(x_{0n}) (x - x_{0n})]. \end{aligned} \quad (14.36)$$

Формулы (14.25) и (14.36) поэтому приводят к следующему результату:

$$G(x) = \exp [-\{1 - c_2 q'(x_{0n}) (x - x_{0n})\}^{1/c_2}]. \quad (14.37)$$

14.15 Из 14.13 следует, что функция распределения $G(x)$ может быть записана в виде

$$G(x) = \exp [-n \{1 - F(x)\}], \quad (14.38)$$

так что в силу (14.25)

$$q(x) = -\log [n \{1 - F(x)\}]. \quad (14.39)$$

В точке x_{0n} согласно (14.27) и (14.39)

$$q(x_{0n}) = -\log [n \{1 - F(x_{0n})\}] = 0.$$

Поэтому

$$n \{1 - F(x_{0n})\} = 1. \quad (14.40)$$

Левая часть в (14.40) представляет собой среднее число значений в выборке объема n , превышающих x_{0n} , так что x_{0n} есть то значение*) случайной величины, для которого среднее число наблюдений, превосходящих x_{0n} , равно 1.

Из (14.39) и (14.40) следует, что

$$q'(x) = F'(x) / \{1 - F(x)\}, \quad (14.41)$$

$$q'(x_{0n}) = nF'(x_{0n}). \quad (14.42)$$

Поэтому согласно (14.35) и (14.41)

$$c_2 = q''(x) / [q'(x)]^2 = \frac{d}{dx} \{-1/q'(x)\} = -\frac{d}{dx} \left\{ \frac{1 - F(x)}{F'(x)} \right\}. \quad (14.43)$$

Если

$$\lim_{x \rightarrow \infty} c_2 = 0, \quad (14.44)$$

то из (14.37), с учетом (14.42) и (14.44), получаем

$$\lim_{x \rightarrow \infty} G(x) = \exp[-\exp\{-nF'(x_{0n})(x - x_{0n})\}]. \quad (14.45)$$

Это предельное распределение было впервые предложено Фишером и Типпеттом (1928). Необходимость условия (14.44) для вывода распределения (14.45) была впервые установлена Мизесом (1936). Для экспоненциального распределения

$$dF(x) = e^{-x} dx, \quad 0 \leq x \leq \infty,$$

выражение в фигурных скобках в (14.43) тождественно по x равно нулю. Следовательно, для этого распределения условие (14.44) выполнено, а поэтому и справедливо соотношение (14.45). Распределения, для которых выполнено (14.44), были названы распределениями экспоненциального типа.

14.16 Если распределение $F(x)$ не экспоненциального типа, то (14.45) заменяется некоторой другой предельной формой. Из (14.43) следует, что

$$c_2 = 1 + F''(x) \{1 - F(x)\} / [F'(x)]^2. \quad (14.46)$$

Если

$$\lim_{x \rightarrow \infty} xF''(x) / \{1 - F(x)\} = k > 0, \quad (14.47)$$

то (14.46) и (14.47) дают при $x \rightarrow \infty$

$$c_2 \sim 1 + xF''(x) / \{kF'(x)\}. \quad (14.48)$$

Воспользовавшись правилом Лопиталю, из (14.47) получаем

$$1 + xF''(x) / F'(x) \sim -k. \quad (14.49)$$

*) Гамбел (1954) называет x_{0n} «наибольшим ожидаемым значением», хотя, как он сам отмечает, это может вызвать недоразумения, поскольку x_{0n} не есть математическое ожидание наибольшей порядковой статистики.

Поэтому из (14.48) и (14.49) следует, что

$$\lim_{x \rightarrow \infty} c_2 = -1/k. \quad (14.50)$$

С учетом (14.42), (14.47) и (14.50) из (14.37) получаем

$$\lim_{n \rightarrow \infty} G(x) = \exp \left[- \{ 1 + n(1 - F(x_{0n}))(x - x_{0n})/x_{0n} \}^{-k} \right],$$

что в силу (14.40) приводит к распределению

$$\lim_{n \rightarrow \infty} G(x) = \exp \left[- \left(\frac{x}{x_0} \right)^{-k} \right], \quad (14.51)$$

являющемуся вторым типом распределений, возникающих в теории крайних значений. Распределение (14.51) было получено Фреше (1927), Фишером и Типпеттом (1928). Условие (14.47) вытекает из соотношения

$$\lim_{x \rightarrow \infty} x^k \{ 1 - F(x) \} = A > 0, \quad (14.52)$$

что легко установить, если применить к (14.52) правило Лопиталя. Распределения, удовлетворяющие (14.52) и, следовательно, условию (14.47), не имеют моментов порядка выше k и поэтому были названы распределениями типа Коши.

14.17 Рассмотрим, наконец, распределение $F(x)$, сосредоточенное на $(-\infty, x_u]$, где $x_u < \infty$ и, следовательно, $F(x_u) = 1$. Если к тому же

$$F^{(r)}(x_u) = 0, \quad F^{(k)}(x_u) \neq 0 \quad (r = 1, 2, \dots, k-1), \quad (14.53)$$

то можно разложить экспоненту (14.38) в ряд Тейлора в окрестности точки x_u , причем первым ненулевым членом будет член, содержащий $F^{(k)}(x_u)$. Если $F^{(k+1)}(x)$ ограничена в окрестности точки x_u , то в силу малости $(x - x_u)$ можно пренебречь остаточным членом и мы получаем выражение

$$\lim_{n \rightarrow \infty} G(x) = \exp \left[- \{ (-nF^{(k)}(x_u)/k!)^{1/k} (x - x_u) \}^k \right], \quad (14.54)$$

которое является третьей предельной формой распределений крайних значений, найденной Фишером и Типпеттом (1928).

14.18 В 14.14—14.17 мы обсуждали возможные предельные распределения для $x_{(n)}$. Если мы поменяем знаки у переменных, то $-x_{(1)}$ станет $x_{(n)}$ и $G(x)$ перейдет в $1 - G(x)$. Поэтому функция распределения $H(x)$ для наименьшего значения получается очень просто:

$$H(x) = 1 - G(-x), \quad (14.55)$$

и результаты, установленные для $x_{(n)}$, применимы к $x_{(1)}$ с учетом этого изменения.

Пример 14.3

В случае нормального распределения

$$dF(x) = e^{-\frac{1}{2}x^2} dx / \sqrt{2\pi},$$

используя (5.68) в (14.43) для больших значений x , получаем

$$c_2 = -\frac{d}{dx} \{(1 - F(x))/F'(x)\} = -\frac{d}{dx} \left(\frac{1}{x} \right) = \frac{1}{x^2},$$

так что $c_2 \rightarrow 0$ при $x \rightarrow \infty$. Следовательно, условие (14.44) выполняется, и тем самым (14.45) является в этом случае предельной формой распределения. Чтобы его явно найти, мы воспользуемся тем фактом, что в точке x_{0n} выполнено условие (14.40). Это нам дает

$$1 = n \{1 - F(x_{0n})\} \sim nF'(x_{0n})/x_{0n},$$

откуда

$$nF'(x_{0n}) \sim x_{0n}.$$

Поэтому показатель экспоненты в (14.45) равен $-x_{0n}(x - x_{0n})$. Кроме того,

$$\log n + \log F'(x_{0n}) - \log x_{0n} \sim 0,$$

т. е.

$$\log n - \frac{1}{2} \log(2\pi) - \frac{1}{2} x_{0n}^2 - \log x_{0n} \sim 0,$$

что дает

$$x_{0n} \sim (2 \log n)^{1/2}.$$

Итак, в нормальном случае распределение (14.45) имеет вид

$$\exp[-\exp\{-(2 \log n)^{1/2}(x - (2 \log n)^{1/2})\}].$$

Асимптотическое распределение k -го значения

14.19 Возвратимся к распределению (14.1). Теперь мы от изучения предельных распределений крайних значений перейдем к общей задаче о предельных распределениях для порядковых статистик. При этом мы ограничимся только случаем наиболее важных распределений экспоненциального типа. Итак, пусть $x_{(n-m+1)}$ является m -й наибольшей порядковой статистикой.

Если в (14.1) мы положим $r = n - m + 1$ и продифференцируем логарифм плотности, то, аналогично 10.10, в точке, являющейся модой распределения $x_{(n-m+1)}$, получим соотношение

$$(n - m)F'/F - (m - 1)F'/(1 - F) + F''/F' = 0. \quad (14.56)$$

Из (14.41) и (14.43) следует, что если (14.44) выполнено, то

$$q''(x) = F''/(1-F) + (F')^2/(1-F)^2 \rightarrow 0,$$

так что поскольку $(1-F) \rightarrow 0$, то F'' также стремится к нулю. Применяя правило Лопиталя к последнему члену в левой части (14.56), получим

$$\lim (F''/F') = \lim \{F'/[-(1-F)]\}. \quad (14.57)$$

Поэтому (14.56) превращается в соотношение

$$(n-m)F'/F - mF'/(1-F) \approx 0, \quad (14.58)$$

следовательно, в точке $\tilde{x}$, являющейся модой распределения $x_{(n-m+1)}$,

$$F(\tilde{x}) \approx 1 - m/n. \quad (14.59)$$

Сравнение (14.59) с (14.40) показывает, что в экспоненциальном случае, когда $m=1$,

$$x_{0n} \approx \tilde{x},$$

т. е. мода распределения $x_{(n)}$ совпадает с тем значением, выше которого может оказаться в среднем лишь одно наблюдение из n .

Разложим $F(x)$ в ряд Тейлора в окрестности точки $\tilde{x}$:

$$F(x) = F(\tilde{x}) + (x - \tilde{x})F'(\tilde{x}) + (x - \tilde{x})^2 F''(\tilde{x})/2 + \dots \quad (14.60)$$

Поскольку в силу (14.57) и (14.59)

$$F'' \sim -(F')^2/(1-F) = -n(F')^2/m, \quad (14.61)$$

то, с учетом (14.59) и (14.61), вместо (14.60) мы получим

$$\begin{aligned} F(x) &= (1 - m/n) + (x - \tilde{x})F'(\tilde{x}) - n \{F'(\tilde{x})\}^2 (x - \tilde{x})^2 / 2m + \dots = \\ &= 1 - (m/n) \left[1 - \frac{(x - \tilde{x})F'(\tilde{x})}{m/n} + \left\{ \frac{(x - \tilde{x})F'(\tilde{x})}{m/n} \right\}^2 \frac{1}{2!} + \dots \right] \end{aligned} \quad (14.62)$$

или, приближенно,

$$F(x) \sim 1 - (m/n) \exp \{-(x - \tilde{x})F'(\tilde{x})/(m/n)\}. \quad (14.63)$$

Формулу (14.1) можно переписать в следующем виде:

$$dG_{n-m+1}(x) \propto \left(\frac{1-F}{F} \right)^{m-1} d(F^n). \quad (14.64)$$

Если воспользоваться обозначением $(-y_m)$ для экспоненты в (14.63) и подставить это значение в (14.64), то найдем

$$\left(\frac{1}{F} - 1\right)^{m-1} = \left(\frac{1}{1 - \frac{m}{n} e^{-y_m}} - 1\right)^{m-1} \sim \left(\frac{m}{n}\right)^{m-1} \exp\{-(m-1)y_m\}$$

и

$$d(F^n) = n \left(1 - \frac{m}{n} e^{-y_m}\right)^{n-1} \frac{m}{n} e^{-y_m} dy_m.$$

Поэтому из (14.64) при больших n получаем

$$dG_{n-m+1}(x) \propto \exp[-my_m - me^{-y_m}] dy_m,$$

откуда, после подсчета нормировочного множителя, следует, что

$$dG_{n-m+1}(x) = \frac{m^m}{(m-1)!} \exp[-my_m - me^{-y_m}] dy_m. \quad (14.65)$$

В частности, при $m=1$ получаем элемент распределения

$$dG_n(x) = \exp[-y - e^{-y}] dy, \quad (14.66)$$

соответствующий функции распределения (14.45).

14.20 Необычное предельное распределение (14.65), найденное Гамбелом (1934), очень далеко от нормального в случае умеренных или малых значений m . Его первый момент (для удобства индекс y опускается) выражается формулой

$$\mu'_1 = \frac{m^m}{(m-1)!} \int_{-\infty}^{\infty} \exp[-my - me^{-y}] y dy.$$

Полагая $e^{-y} = t/m$, получим

$$\begin{aligned} \mu'_1 &= -\frac{1}{(m-1)!} \int_0^{\infty} (\log t - \log m) t^{m-1} e^{-t} dt = \\ &= \log m - \frac{d}{dm} \{\log \Gamma(m)\} = \log m + \gamma - \sum_{r=1}^{m-1} \frac{1}{r}, \end{aligned} \quad (14.67)$$

где γ — константа Эйлера. Для r -го центрального момента имеет место формула

$$\begin{aligned} \mu_r &= \frac{(-1)^r}{(m-1)!} \int_0^{\infty} (\log t - \log m + \mu'_1)^r t^{m-1} e^{-t} dt = \\ &= \frac{(-1)^r}{\Gamma(m)} \left[\frac{d^r}{dq^r} \exp\{q(\mu'_1 - \log m)\} \Gamma(m+q) \right]_{q=0}. \end{aligned} \quad (14.68)$$

Эти формулы были получены и подробно исследованы Гамбелом, из работ которого заимствована следующая таблица:

m	Среднее	$\sqrt{\beta_1}$	$\beta_2 - 3$
1	0,577	1,139	2,400
3	0,176	0,621	0,763
5	0,103	0,468	0,437
10	0,051	0,324	0,212

Эти данные, соответствующие порядковым статистикам $x_{(n-m+1)}$ при малых m , ясно показывают, что предельное распределение сильно отличается от нормального. Аналогично устроено и распределение для $x_{(n-m+1)}$ при m , близких к n , причем его нечетные моменты, включая и само среднее, имеют те же самые значения, но противоположного знака, что и соответствующие моменты при малых m ; четные же моменты остаются теми же самыми.

Дальнейшее изучение асимптотических распределений порядковых статистик было проведено Дронкерсом (J. J. Dronkers, 1958, *Віом.* 45, 447), который вывел предельные формы для всех трех типов исходных распределений*). Им же были даны аппроксимации допредельных распределений. В экспоненциальном случае было установлено, что сходимость к предельной форме (14.65) является очень медленной (даже в случае исходного нормального распределения). Скорость сходимости может быть улучшена с помощью приема, использованного в 14.29 для решения одной аналогичной задачи.

Точные распределения порядковых статистик в нормальном случае

14.21 Многие авторы интересовались точным распределением порядковых статистик для того случая, когда результаты наблюдений нормально распределены. Наиболее полные результаты в этом направлении получил Рубин (1954), который установил общие рекуррентные соотношения для моментов статистик $x_{(r)}$ (в случае выборок объема n) через аналогичные моменты для выборок меньшего объема. Им были табулированы первые десять моментов для наибольшей порядковой статистики

*) Подробную и обстоятельную классификацию исходных типов и предельных форм дал Н. В. Смирнов (1949), *Труды Матем. ин-та АН СССР* 25, 1—60. (Прим. ред.)

при $n=1(1)50$. Пользуясь этими вычислениями, он нашел первые четыре центральных момента, а также коэффициенты асимметрии и эксцесса с точностью по меньшей мере до семи значащих цифр. Для наибольшей порядковой статистики Типпетт (1925) еще раньше вычислил ее распределение для ряда значений n (вплоть до 1000). Приводимые значения взяты из его таблиц:

n	Среднее	Стандартное отклонение	β_1	β_2
2	0,564	0,826	0,019	3,062
5	1,163	0,669	0,092	3,202
10	1,539	0,587	0,168	3,331
100	2,508	0,429	0,429	3,765
500	3,037	0,370	0,570	4,003
1000	3,241	0,351	0,618	4,088

В этой таблице значения β_1 и β_2 иллюстрируют то обстоятельство, что с ростом n распределение крайних значений все более отклоняется от нормального.

14.22 Хартли и Дэвид (1954) нашли верхнюю границу для среднего значения наибольшей порядковой статистики с помощью следующего приема. Согласно неравенству Коши — Буниковского

$$\int_0^1 nF^{n-1} \left(x + \frac{1}{a}\right) dF \leq \left[\left\{ \int_0^1 (nF^{n-1})^2 dF \right\} \left\{ \int_0^1 \left(x + \frac{1}{a}\right)^2 dF \right\} \right]^{1/2},$$

где F — любая функция распределения и $a = (n-1)/(2n-1)^{1/2}$. Если все интегралы конечны, то это неравенство превращается (без потери общности можно считать среднее значение равным нулю, а дисперсию — единице) в следующее:

$$\int_0^1 nF^{n-1} x dF + \frac{1}{a} \leq \left[\frac{n^2}{2n-1} \cdot \left(1 + \frac{1}{a^2}\right) \right]^{1/2}.$$

Интеграл слева есть не что иное, как среднее значение наибольшей порядковой статистики. Поэтому, используя значение a , получим

$$M(x_{(n)}) \leq (n-1)/(2n-1)^{1/2} = a. \quad (14.69)$$

Если исходное распределение симметрично, то $F(x) = 1 - F(-x)$. В этих условиях имеет место более сильный результат

$$M(x_{(n)}) \leq n \left[\left\{ 1 - 1/\binom{2n-2}{n-1} \right\} / \{2(2n-1)\} \right]^{1/2}, \quad (14.70)$$

принадлежащий Мориугути (1951). Им же были даны в симметричном случае аналогичные, но, правда, более сложные формулы для верхних и нижних границ дисперсии и коэффициентов вариации статистик $x_{(n)}$. Очевидная нулевая нижняя граница для $M(x_{(n)})$ не может вообще быть улучшена. Можно также доказать, что верхняя граница достигается в случае распределения

$$F(x) = \left(\frac{ax+1}{n} \right)^{1/(n-1)}, \quad -1/a \leq x \leq (n-1)/a. \quad (14.71)$$

Для наименьшей порядковой статистики аналогичный результат справедлив с заменой $M(x_{(n)})$ на $|M(x_{(1)})|$. Хартли и Дэвид (1954) получили следующий общий результат для порядковых статистик:

$$|M(x_{(m)})| \leq \left[\frac{B(2m-1, 2n-2m+1)}{\{B(m, n-m+1)\}^2} - 1 \right]^{1/2}. \quad (14.72)$$

Равенство здесь достигается только тогда, когда или $m=1$, или $n=1$, т. е. в том специальном случае, который был рассмотрен выше.

Совместное распределение двух порядковых статистик

14.23 Если в (14.2) мы сделаем преобразование

$$\begin{cases} y = nF(x_{(r)}), \\ z = n\{1 - F(x_{(s)})\}, \end{cases}$$

то получим

$$dH_{r,s}(y, z) = \frac{\left(\frac{y}{n}\right)^{r-1} \left(1 - \frac{y}{n} - \frac{z}{n}\right)^{s-r-1} \left(\frac{z}{n}\right)^{n-s} dy dz}{B(r, s-r) B(s, n-s+1) n^2}. \quad (14.73)$$

Пусть r и $n-s$ фиксированы и малы и пусть n стремится к бесконечности. Полагая временно $t = n - s + 1$, мы получим в пределе

$$dH = \frac{y^{r-1} e^{-y} dy}{\Gamma(r)} \cdot \frac{z^{t-1} e^{-z} dz}{\Gamma(t)}. \quad (14.74)$$

Поэтому r -й наименьший член и t -й наибольший член асимптотически независимы, что можно было, конечно, предвидеть. Каждое из этих значений имеет Γ -распределение. В частном случае $t=r$ имеем

$$dH = \frac{y^{r-1} e^{-y} dy}{\Gamma(r)} \cdot \frac{z^{n-r+1} e^{-z} dz}{\Gamma(n-r+1)}. \quad (14.75)$$

Сархан и Гринберг (1956) табулировали ковариации всех порядковых статистик в нормальном случае для $n=2(1)20$ с десятью десятичными знаками.

Размахи и середины размахов

14.24 О размахе выборки уже кратко речь шла в 2.17. Сейчас мы обобщим это определение, введя понятие m -размаха как разности

$$R_m = x_{(n-m+1)} - x_{(m)}, \quad (14.76)$$

т. е. R_m — расстояние между m -ми наибольшим и наименьшим значениями. Само понятие размаха мы получаем, если $m=1$.

Обобщая понятие середины размаха, введем так называемую m -ю середину размаха

$$M_m = \frac{1}{2} (x_{(n-m+1)} + x_{(m)}). \quad (14.77)$$

Введенная ранее середина размаха, очевидно, совпадает с M_1 .

Из (14.75) видно, что асимптотические распределения размахов и средин размахов совпадают с распределением разностей и сумм преобразованных независимых величин. Можно, однако, вывести точные распределения непосредственно из (14.2), положив $r=m=n-s+1$ и воспользовавшись формулами (14.76) и (14.77). На этом пути для совместного распределения R_m и M_m получим

$$dH_m \propto \left\{ F\left(M - \frac{1}{2} R\right) \right\}^{m-1} \left\{ F\left(M + \frac{1}{2} R\right) - F\left(M - \frac{1}{2} R\right) \right\}^{n-2m} \times \\ \times \left\{ 1 - F\left(M + \frac{1}{2} R\right) \right\}^{m-1} f\left(M - \frac{1}{2} R\right) f\left(M + \frac{1}{2} R\right) dR dM, \quad (14.78)$$

где индекс m в правой части опущен для удобства записи.

Распределение каждой из величин R_m и M_m получается интегрированием распределения (14.78). При этом нужно побеспокоиться о пределах интегрирования, поскольку они зависят в каждом случае от интервала, на котором сосредоточено исходное распределение F . Конечно, можно было бы получить эти распределения и прямо из формулы (14.2).

Точное распределение размаха

14.25 Положим в (14.78) $m=1$ и проинтегрируем по M . Тем самым мы получим распределение размаха

$$dG(R) = dR n(n-1) \int_{-\infty}^{\infty} \{F(x+R) - F(x)\}^{n-2} f(x+R) dF(x). \quad (14.79)$$

Интегрирование в элементарных функциях здесь удастся осуществить лишь в некоторых частных случаях. В большинстве слу-

чаев приходится прибегать к численному интегрированию. Из (14.79) для среднего значения размаха получим

$$\begin{aligned} M(R) &= n(n-1) \int_{-\infty}^{\infty} dF_n \int_{-\infty}^{x_{(n)}} (F_n - F_1)^{n-2} (x_{(n)} - x_{(1)}) dF_1 = \\ &= - \int_{-\infty}^{\infty} dx_{(n)} \int_{-\infty}^{x_{(n)}} \frac{d^2}{dx_{(1)} dx_{(n)}} (F_n - F_1)^n (x_{(n)} - x_{(1)}) dx_{(1)}. \quad (14.80) \end{aligned}$$

Интегрируя по частям под знаком второго интеграла, находим

$$\begin{aligned} M(R) &= - \int_{-\infty}^{\infty} dx_{(n)} \left\{ \left[- (x_{(n)} - x_{(1)}) \frac{d}{dx_{(n)}} F_n^n \right]_{x_{(1)} = -\infty} + \right. \\ &\quad \left. + \int_{-\infty}^{x_{(n)}} \frac{d}{dx_{(n)}} (F_n - F_1)^n dx_{(1)} \right\}. \quad (14.81) \end{aligned}$$

В силу того, что для верхнего предела интегрирования $(F_n - F_1)^n = 0$, порядок интегрирования и дифференцирования в последнем члене в (14.81) можно поменять местами. Это дает

$$\begin{aligned} M(R) &= \int_{-\infty}^{\infty} \left[(x_{(n)} - x_{(1)}) \frac{d}{dx_{(n)}} F_n^n \right]_{x_{(1)} = -\infty} dx_{(n)} - \int_{-\infty}^{\infty} (1 - F_1)^n dx_{(1)} = \\ &= \int_{-\infty}^{\infty} \left\{ \frac{d}{dx_{(n)}} [(x_{(n)} - x_{(1)}) F_n^n] - F_n^n \right\}_{x_{(1)} = -\infty} dx_{(n)} - \int_{-\infty}^{\infty} (1 - F_1)^n dx_{(1)} = \\ &= [(x_{(n)} - x_{(1)}) F_n^n]_{x_{(1)} = -\infty}^{x_{(n)} = \infty} - \int_{-\infty}^{\infty} F_n^n dx_{(n)} - \int_{-\infty}^{\infty} (1 - F_1)^n dx_{(1)} = \\ &= \int_{-\infty}^{\infty} \{1 - F^n - (1 - F)^n\} dx. \quad (14.82) \end{aligned}$$

Типпетт, которому принадлежит эта формула, установил также формулу и для дисперсии:

$$DR = 2 \int_{-\infty}^{\infty} \int_{-\infty}^{x_{(n)}} \{1 - F_n^n - (1 - F_1)^n - (F_n - F_1)^n\} dx_{(1)} dx_{(n)} - \{MR\}^2.$$

В случае нормального исходного распределения Типпеттом (1925) и Э. Пирсоном ((1926), (1932)) было проведено подробное исследование. Типпетт нашел первые четыре момента распределения размаха, табулировал средние значения для $n \leq 1000$ и дал номограмму для определения стандартных ошибок. Его таблицы средних значений (с пятью десятичными знаками) воспроизведены в *Biometrika Tables for Statisticians*, vol 1.

Нижеследующие значения показывают общее поведение распределения размаха, когда исходное распределение нормально.

n	Стандартное отклонение	β_1 (прибл.)	β_2 (прибл.)
2	0,853	0,99	3,87
10	0,797	0,16	3,20
100	0,605	0,21	3,39
500	0,524	0,39	3,50
1000	0,497	0,31	3,54

Из таблицы видно, что с ростом n распределение размаха все более и более отличается от нормального.

Пирсон и Хартли (1942) в нормальном случае табулировали с четырьмя десятичными знаками функцию распределения R/σ для выборок объема от 2 до 20. Эти результаты также воспроизведены *) в *Biometrika Tables for Statisticians*, vol 1. Хартер (H. L. Harter, 1960, *Ann. Math. Statist.* 31, 1122) дал с шестью десятичными знаками для $n=2(1)20(2)40(10)100$ значения R/σ , для которых функция распределения равна 0,0001, 0,0005, 0,001, 0,005, 0,01, 0,025, 0,05, 0,1(0,1)0,9, 0,95, 0,975, 0,99, 0,995, 0,999, 0,9995, 0,9999. Он также дал значения первых четырех моментов по меньшей мере с восемью значащими цифрами для $n=2(1)100$.

Коксом (1954) был рассмотрен эффект отклонения распределения размаха от нормального распределения в случае выборок малого объема. При этом им было отмечено, что эксцесс в большей степени, чем асимметрия, оказывает влияние на поведение распределения.

14.26 Из (14.70) следует, что в случае симметричных распределений (с нулевым средним и единичной дисперсией) для среднего значения размаха справедливы неравенства

$$0 \leq M(R) \leq n \left[2 \left\{ 1 - \frac{1}{\binom{2n-2}{n-1}} \right\} / (2n-1) \right]^{1/2} \sim \left(n + \frac{1}{2} \right)^{1/2},$$

*) Эти таблицы помещены также в сборнике *Таблицы математической статистики*. (Прим. ред.)

установленные Плэкеттом (1947), который доказал справедливость этого результата и для несимметричных распределений. Хартли и Дэвид (1954) также пришли к этому результату и нашли верхние и нижние границы $M(R)$ для распределений, сосредоточенных на конечных интервалах. В приводимой ниже таблице (заимствованной из работы Плэкетта (1947)) даются значения для верхней границы, ее асимптотические и точные значения в случае нормальных и равномерных исходных распределений.

n	Верхняя граница для $M(R)$	$(n + \frac{1}{2})^{1/2}$	Точные значения	
			нормальное распределение	равномерное распределение
2	1,155	1,581	1,128	1,155
3	1,732	1,871	1,693	1,732
4	2,084	2,121	2,059	2,078
8	2,921	2,915	2,847	2,694
12	3,539	3,536	3,258	2,931

Асимптотическое распределение размаха

14.27 Перейдем к изучению предельного при $n \rightarrow \infty$ распределения для R . Полагая в (14.2) $r = n - s + 1 = 1$, для совместного распределения наименьшего и наибольшего наблюдений получаем формулу

$$dG_{1, n}(x_{(1)}, x_{(n)}) = n(n-1)(F_n - F_1)^{n-2} dF_1 dF_n.$$

Введем новые величины с помощью преобразования

$$\left. \begin{aligned} u &= 2n \{F_1(1 - F_n)\}^{1/2}, \\ v &= \frac{1}{2} \log \{F_1/(1 - F_n)\}, \end{aligned} \right\} \quad (14.83)$$

якобиан которого равен $u/[2n^2 f_1 f_2]$. Совместное распределение для u и v выражается формулой

$$dF(u, v) = \left(\frac{n-1}{2n}\right) u \left(1 - \frac{u \operatorname{ch} v}{n}\right)^{n-2} du dv. \quad (14.84)$$

Если $n \rightarrow \infty$, то

$$\lim_{n \rightarrow \infty} dF(u, v) = \frac{1}{2} u \exp(-u \operatorname{ch} v) du dv. \quad (14.85)$$

В (14.84) область изменения переменных ограничена соотношениями

$$u \geq 0, \quad \operatorname{ch} v \leq n/u,$$

поэтому в (14.85) $\operatorname{ch} v$ и v могут изменяться до бесконечности. Следовательно, из (14.85) вытекает, что предельное распределение для u выражается формулой

$$\begin{aligned} dF(u) &= u \, du \int_0^{\infty} \exp(-u \operatorname{ch} v) \, dv = \\ &= u \, du \int_1^{\infty} \exp(-ut) (t^2 - 1)^{-\frac{1}{2}} \, dt. \end{aligned} \quad (14.86)$$

Переменную u можно переписать в виде

$$u = 2n \left\{ F\left(M - \frac{1}{2} R\right) \left[1 - F\left(M + \frac{1}{2} R\right)\right] \right\}^{\frac{1}{2}}. \quad (14.87)$$

Если теперь предположить, что середина размаха выборки M сходится по вероятности к некоторой известной величине μ , то u будет вести себя подобно

$$u' = 2n \left\{ F\left(\mu - \frac{1}{2} R\right) \left[1 - F\left(\mu + \frac{1}{2} R\right)\right] \right\}^{\frac{1}{2}}. \quad (14.88)$$

Если $\mu = 0$ и генеральная совокупность симметрична, то вместо (14.88) получим

$$u' = 2n \left\{ F\left(-\frac{1}{2} R\right) \right\}. \quad (14.89)$$

Отсюда следует, что в этом случае u будет стремиться просто к $F\left(-\frac{1}{2} R\right)$ (с точностью до постоянного множителя).

14.28 При выводе этих результатов, принадлежащих Элфвингу (1947), было использовано предположение о том, что середина размаха M сходится к некоторой известной постоянной μ по вероятности. Но даже тогда, когда исходное распределение симметрично и принадлежит экспоненциальному типу, это условие может не выполняться. Чтобы это показать, мы положим в (14.68) $m=1$ и $r=2$. Тогда можно найти, что дисперсия распределения (14.66) равна $\pi^2/6$. Согласно (14.45) переменной здесь является величина

$$y = nF'(x_{0n})(x - x_{0n}), \quad (14.90)$$

где $x = x_{(n)}$. Таким образом, здесь

$$Dy = \pi^2/6$$

и, следовательно, в случае больших выборок

$$Dx = \pi^2/[6 [nF'(x_{0n})]^2]. \quad (14.91)$$

Поскольку $x_{(n)}$ и $x_{(1)}$ асимптотически независимы, средина размаха, являющегося их полусуммой, имеет дисперсию

$$D(M) = \pi^2 / \{12 [nF'(x_{0n})]^2\}. \quad (14.92)$$

Пример 14.4

В случае нормального распределения (пример 14.3)

$$nF'(x_{0n}) \sim (2 \log n)^{1/2},$$

так что в силу (14.92)

$$D(M) = \pi^2 / (24 \log n).$$

Поэтому если исходное распределение нормально, то M сходится к своему среднему значению (которое в силу симметрии является средним значением исходного нормального распределения).

Пример 14.5

Двустороннее экспоненциальное распределение

$$dF = \frac{1}{2} \exp(-|x|) dx, \quad -\infty \leq x \leq \infty,$$

имеет конечные моменты всех порядков. Его функция распределения выражается формулой

$$F(x) = 1 - \frac{1}{2} e^{-x}, \quad x \geq 0. \quad (14.93)$$

Легко проверить, что здесь условие (14.44) удовлетворяется. Согласно (14.40) и (14.93) $x_{0n} = \log\left(\frac{1}{2}n\right)$, и поэтому

$$nF'(x_{0n}) = \frac{1}{2} n \exp\left(-\log \frac{1}{2}n\right) = 1.$$

Тем самым в этом случае согласно (14.92)

$$D(M) = \pi^2 / 12,$$

и, следовательно, M не сходится по вероятности к нулю, являющемуся здесь средним значением статистики M и исходной генеральной совокупности.

Редуцированные размах и средина размаха

14.29 Имея в виду замечание, сделанное в предшествующем пункте, мы рассмотрим теперь проблему, поставленную в **14.27**, и попытаемся дать более общее решение задачи об асимптотическом распределении размаха и срединны размаха или некоторых

простых функций от них. При этом мы сосредоточим наше внимание на симметричных распределениях экспоненциального типа.

Если мы приравняем правые части формул (14.38) и (14.45), то получим

$$1 - F(x_{(n)}) = \frac{1}{n} \exp \{-nF'(x_{0n})(x_{(n)} - x_{0n})\}$$

и аналогично, используя формулу (14.55),

$$F(x_{(1)}) = \frac{1}{n} \exp \{nF'(x_{0n})(x_{(1)} + x_{0n})\}.$$

Поэтому вместо (14.83) можно написать

$$\left. \begin{aligned} u &= 2 \exp \left\{ -\frac{1}{2} nF'(x_{0n})(x_{(n)} - x_{(1)} - 2x_{0n}) \right\}, \\ v &= \frac{1}{2} nF'(x_{0n})(x_{(n)} + x_{(1)}). \end{aligned} \right\} \quad (14.94)$$

Определим теперь редуцированный размах R^* и редуцированную средину размаха M^* формулами:

$$\left. \begin{aligned} R^* &= nF'(x_{0n})(R - 2x_{0n}), \\ M^* &= nF'(x_{0n})M. \end{aligned} \right\} \quad (14.95)$$

Тогда вместо (14.94) получим

$$u = 2 \exp \left(-\frac{1}{2} R^* \right), \quad (14.96)$$

$$v = M^*. \quad (14.97)$$

Используя формулу (14.96) и распределение u , выражающееся формулой (14.86), мы получим распределение для R^*

$$\begin{aligned} dF(R^*) &= 2 \exp(-R^*) dR^* \int_0^\infty \exp \left\{ -2 \exp \left(-\frac{1}{2} R^* \right) \operatorname{ch} v \right\} dv = \\ &= 2 \exp(-R^*) K_0 \left\{ 2 \exp \left(-\frac{1}{2} R^* \right) \right\} dR^*, \end{aligned} \quad (14.98)$$

где K_0 — функция Бесселя нулевого порядка от мнимого аргумента. Гамбел (1949), ранее получивший формулу (14.98) другим методом, дал таблицы плотности и функции распределения R^* и табулировал процентные точки, т. е. квантили, соответствующие вероятностям, выраженным в процентах. Этот результат был получен также Коксом (1948). Для получения новой аппроксимации им был использован метод перевала. Кроме того, Кокс провел сравнение результатов, даваемых новой аппро-

ксимацией, с (14.98) и результатом Элфвинга (14.86). Как и следовало ожидать, в нормальном случае приближение (14.86) и приближение, найденное с помощью метода перевала, оказались лучше, чем (14.98). Но, как мы увидим в следующем абзаце, даже (14.98) дает хорошую аппроксимацию для малых n .

Согласно (14.95) R является линейной функцией от R^* . Так же как в 14.28, асимптотически

$$\left. \begin{aligned} M(R^*) &= M(y_{(n)}) - M(y_{(1)}) = 2\gamma, \\ D(R^*) &= Dy_{(n)} + Dy_{(1)} = \pi^2/3. \end{aligned} \right\} \quad (14.99)$$

Из (14.99), используя (14.95), мы получим, что асимптотически

$$\left. \begin{aligned} M(R) &= 2\gamma/\{nF'(x_{0n})\} + 2x_{0n}, \\ D(R) &= \pi^2/[3\{nF'(x_{0n})\}^2]. \end{aligned} \right\} \quad (14.100)$$

Основываясь на этом результате и используя точные значения для $M(R)$ и $D(R)$ (табулированные Типпеттом (1925) и Э. Пирсоном (1926)) для малых n , Гамбел (1947) оценил x_{0n} и $F'(x_{0n})$ и исследовал точность аппроксимации распределения R , полученной из (14.98), путем сравнения ее результатов с точными значениями, вычисленными по таблице Пирсона и Хартли (1942). Он нашел, что функция распределения R хорошо аппроксимируется распределением (14.98) уже при $n=6$. Поэтому, по крайней мере в нормальном случае, формула (14.98) удовлетворительно решает задачу о распределении размаха.

14.30 Рассмотрим теперь редуцированную средину размаха M^* . В силу (14.97) мы видим, что ее асимптотическое распределение, полученное интегрированием по u распределения (14.85), выражается формулой

$$dG(M^*) = \frac{1}{2} dM^* \int_0^\infty u \exp(-u \operatorname{ch} M^*) du. \quad (14.101)$$

Полагая $z = u \operatorname{ch} M^*$, получим

$$dG(M^*) = dM^*/(2 \operatorname{ch}^2 M^*). \quad (14.102)$$

С помощью формул (14.95) и (14.102) можно получить непосредственно и распределение средины размаха M . Это распределение симметрично, имеет нулевое среднее. Поэтому для симметричных распределений экспоненциального типа M является несмещенной оценкой своего среднего значения (среднего значения совокупности), равного нулю. Но пример 14.5 показывает, что четные моменты M сильно зависят от формы исходного распределения.

Таблицы для распределений крайних значений

14.31 Сборник *Probability Tables for the Analysis of Extreme — Value Data* (National Bureau of Standards, Applied Mathematics, Series 22, Washington, 1953) позволяет находить:

(1) функцию и плотность распределений крайних значений в случае исходных распределений экспоненциального типа (см. формулы (14.45) и (14.66)) с семью десятичными знаками в интервале от $-3,0$ до $17,0$, включая $-2,40(0,05)0(0,1)4,0$, а также квантили крайних значений*) с пятью десятичными знаками, соответствующие вероятностям $0(0,0001)0,005(0,001)0,988(0,0001)0,9994(0,00001)1$. Кроме того, с пятью десятичными знаками здесь же даны значения плотности, соответствующие значениям ф. р. $0(0,0001)0,010(0,001)0,999$;

(2) процентные точки распределения m -х порядковых статистик (см. (14.65)) для $m=1(1)15(5)50$ и двенадцать процентных точек размаха для вероятностей, лежащих между $0,005$ и $0,995$;

(3) функцию и плотность распределения редуцированного размаха (см. (14.98)) (в случае исходных распределений экспоненциального типа) для значений $R^*=1,0(0,05)11,0(0,5)20,0$ с семью десятичными знаками, а также квантили редуцированного размаха (с тремя и четырьмя десятичными знаками) для некоторых избранных значений вероятностей.

14.32 Когда в последующих главах (второй том) мы будем рассматривать теорию оценок и теорию различения гипотез, нам придется обратиться к изучению некоторых других выборочных свойств порядковых статистик.

УПРАЖНЕНИЯ

14.1 Показать, что для распределения Коши

$$dF = \frac{dx}{\pi(1+x^2)}, \quad -\infty \leq x \leq \infty,$$

дисперсия выборочной медианы в случае выборок нечетного объема n приближенно равна

$$\frac{1}{4} \pi^2 (n+2)^{-1} + \frac{1}{8} \pi^4 (n+2)^{-2}.$$

(Дэвид и Джонсон, 1954.)

14.2 Для «логистического» распределения с ф. р.

$$F(x) = (1 + e^{-x})^{-1}$$

*) Таблица квантилей воспроизведена в книге: И. В. Смирнов, И. В. Дунин-Барковский (1959), *Краткий курс математической статистики для технических приложений*, М., Физматгиз, 1959. См. также *Сборник статистических таблиц*. (Прим. ред.)

установить прямым методом, что предельная ф. р. статистики $z = x_{(n)} - \log n$ выражается формулой

$$G(z) = \exp(-e^{-z}).$$

Показать, что в рассматриваемом случае этот результат эквивалентен (14.45).

14.3 Пользуясь обозначениями из 14.9, показать, что в выборках нечетного объема n , с точностью до $(n+2)^{-2}$, для выборочной медианы выполнены соотношения:

$$\begin{aligned} M(m_e) &= X_{r+1} + \frac{1}{8(n+2)} X''_{r+1} + \frac{1}{128(n+2)^2} X'''_{r+1}, \\ D(m_e) &= \frac{1}{4(n+2)} (X'_{r+1})^2 + \frac{1}{32(n+2)^2} \{2X'_{r+1}X'''_{r+1} + (X''_{r+1})^2\}, \\ \gamma_1(m_e) &= \frac{3}{2(n+2)^{1/2}} \frac{X''_{r+1}}{X'_{r+1}}, \\ \gamma_2(m_e) &= \frac{1}{n+2} \left\{ \frac{X'''_{r+1}}{X'_{r+1}} + \frac{3(X''_{r+1})^2}{(X'_{r+1})^2} - 6 \right\}. \end{aligned}$$

(Дэвид и Джонсон, 1954.)

14.4 Используя х. ф. упражнений 4.21 и 14.21, установить результат (14.10) для предельного распределения редуцированной середины размаха.

14.5 Пусть $G_{r,n}$ — функция распределения r -й порядковой статистики в выборке объема n (см. (14.1)). Показать, что

$$(n+1) dG_{r,n} = r dG_{r+1,n+1} + (n-r+1) dG_{r,n+1}$$

и что, следовательно, х. ф. и моменты $x_{(r)}$ удовлетворяют аналогичному рекуррентному соотношению. Аналогично если $G_{r,s,n}$ — функция распределения (14.2), то

$$(n+1) dG_{r,s,n} = r dG_{r+1,s+1,n+1} + (s-r) dG_{r,s+1,n+1} + (n-s+1) dG_{r,s,n+1}.$$

14.6 Переписав (14.20) в виде $V_p = p(1-p)/(nf_p^2)$, $f_p > 0$, показать, что

$$\begin{aligned} \frac{\partial \log V_p}{\partial p} &= \frac{1-2p}{p(1-p)} - 2 \frac{f'_p}{f_p^2}, \\ \frac{\partial^2 \log V_p}{\partial p^2} &= -\left(\frac{1}{p^2} + \frac{1}{(1-p)^2} \right) + 2 \left(\frac{f'_p}{f_p^2} \right)^2 - 2 \frac{f''_p}{f_p^2}. \end{aligned}$$

Вывести отсюда, что V_p имеет экстремум при $p=0,5$ тогда и только тогда, когда в этой точке f_p также имеет экстремум; кроме того, показать, что $f_{0,5}$ — минимальное значение, если $V_{0,5}$ — максимальное значение; что если $f_{0,5}$ — максимальное значение, то $V_{0,5}$ — минимальное значение при $|f''_{0,5}| > 4f_{0,5}^2$ и максимальное значение в обратном случае; что если точка 0,5 является для f_p точкой перегиба, то $V_{0,5}$ — максимальное значение. Показать также, что если f_p имеет экстремум в точке $p_0 \neq 0,5$, то V_p не имеет точек экстремума ни в точке p_0 , ни в точке 0,5.

Сэн (Sen, 1961, *J. Roy. Statist. Soc. B23*, 453.).

14.7 Для непрерывного симметричного распределения, плотность которого имеет абсолютный максимум в точке ξ , являющейся медианой, показать, что дисперсия медианы выборки нечетного объема удовлетворяет неравенству

$$Dm_e \geq [4 \{f(\xi)\}^2 (n+2)]^{-1},$$

и установить, что равенство достигается на равномерном распределении. Отсюда вывести, что асимптотическое значение дисперсии в действительности представляет собой нижнюю границу для дисперсии в случае выборок малого объема.

(Чу, 1955.)

14.8 Показать, что в выборке объема n из совокупности с распределением

$$dF = e^{-x} dx, \quad 0 \leq x \leq \infty,$$

дисперсия наименьшего выборочного значения равна $1/n^2$.

14.9 Доказать, что предельные выражения, задаваемые уравнениями (14.45), (14.51) и (14.54), удовлетворяют функциональному соотношению (14.24).

14.10 Рассмотрим выборку объема n из нормальной совокупности с единичной дисперсией и выберем то выборочное значение, которое ближе всего к истинному среднему. Показать, что дисперсия этого значения выражается формулой

$$v_n = 1 - \frac{n(n-1)}{\pi} + \frac{1}{2} n(n-1)(n-2) \left(\frac{2}{\pi}\right)^{\frac{1}{2}} n \times \\ \times \int_0^{\infty} \exp\left(-\frac{3}{2}x^2\right) \left\{ \int_x^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt \right\}^{n-3} dx.$$

(Кендалл, 1954.)

14.11 В условиях предыдущего упражнения показать, что

$$v_5 = 1 - \frac{20}{\pi} + \frac{240}{\pi^2 \sqrt{3}} \arctg\left(\sqrt{\frac{5}{3}} - \frac{\pi}{6}\right)$$

(это точное равенство) и приближенно $v_n = \pi/n^2$.

(Кендалл, 1954.)

14.12 Показать, что элемент совместного распределения размаха и средней размаха выборки из совокупности с равномерным в интервале $\left(-\frac{1}{2}, \frac{1}{2}\right)$ распределением, выражается формулой

$$dF(R, M) = n(n-1) R^{n-2} dR dM, \quad 0 \leq R \leq 1 - 2|M| \leq 1.$$

14.13 Пользуясь результатом предыдущего упражнения, показать, что элемент маргинального распределения M выражается формулой

$$dG(M) = n(1 - 2|M|)^{n-1} dM, \quad |M| \leq \frac{1}{2}.$$

и что M имеет моменты

$$\mu_{2k-1} = 0, \quad \mu_{2k} = \left\{ 2^{2k} \binom{2k+n}{2k} \right\}^{-1}, \quad k = 1, 2, \dots$$

В частности, $\mu_2 = \{2(n+1)(n+2)\}^{-1}$.

14.14 Показать, что в выборке объема $n \geq 2$ из нормальной совокупности разности порядковых статистик и выборочного среднего не зависят от выборочного среднего. Отсюда вывести, что

$$x_r(x_{(j)} - \bar{x}) = \begin{cases} x_r(x_{(j)}) & r > 2, \\ x_2(x_{(j)}) - n^{-1} & r = 2. \end{cases}$$

(Мак Кей, 1935.)

14.15 Показать, что среднее значение размаха выборки объема n выражается равенством

$$M(R) = \int_{-\infty}^{\infty} x d\{F^n + (1-F)^n - 1\}.$$

Отсюда вывести, что $M(R)$ существует тогда и только тогда, когда среднее значение генеральной совокупности конечно. В этом случае $M(R)$ также равно

$$\int_{-\infty}^{\infty} \{1 - F^n - (1-F)^n\} dx.$$

(Кокс, 1954.)

14.16 Используя результат упражнения 14.12, показать, что маргинальное распределение R выражается формулой

$$dH(R) = n(n-1) R^{n-2} (1-R) dR, \quad 0 \leq R \leq 1,$$

и имеет моменты

$$\mu'_k = n(n-1)/\{(n+k)(n+k-1)\}.$$

В частности,

$$\mu'_1 = (n-1)/(n+1),$$

$$\mu'_2 = 2(n-1)/\{(n+1)^2(n+2)\}.$$

14.17 Пользуясь результатом упражнения 14.12, показать, что отношение M/R подчиняется распределению

$$dF\left(\frac{M}{R}\right) = (n-1) \left\{ 1 + 2 \left| \frac{M}{R} \right| \right\}^{-n} d\left(\frac{M}{R}\right),$$

моменты которого выражаются равенствами

$$\mu_{2k-1} = 0, \quad \mu_{2k} = 1 / \left\{ 2^{2k} \binom{n-2}{2k} \right\}, \quad k = 1, 2, \dots$$

(Карлтон, 1946.)

14.18 Полагая в упражнении 14.13 $t = nM$, показать, что

$$\lim_{n \rightarrow \infty} dF(t) = \exp(-2|t|) dt.$$

Отсюда вывести, что разность между двумя так редуцированными срединами размахов, $z=t_1-t_2$, подчиняется распределению

$$dG(z) = \left(\frac{1}{2} + |z|\right) \exp(-2|z|) dz,$$

моменты которого выражаются равенствами

$$\mu_{2k-1} = 0, \quad \mu_{2k} = (2k)! (k+1)/2^{2k}, \quad k = 1, 2, \dots$$

(Карлтон, 1946.)

14.19 Показать, что функция распределения размаха выражается формулой

$$G(R) = n \int_{-\infty}^{\infty} \{F(x+R) - F(x)\}^{n-1} dF(x).$$

(Гамбел, 1947.)

14.20 Показать, что случайная величина u с функцией распределения (14.86) имеет среднее $\frac{1}{2}\pi$ и дисперсию $\left(4 - \frac{1}{4}\pi^2\right)$.

(Элфвинг, 1947.)

14.21 Показать, что распределение (14.66) имеет своей х. ф. $\Gamma(1-it)$. Следовательно,

$$\begin{aligned} \kappa_1 &= \gamma \text{ (постоянная Эйлера, } = 0,577\dots); \\ \kappa_2 &= \sum n^{-2} = \pi^2/6 = 1,645; \quad \kappa_3 = 2 \sum n^{-3} = 2,404; \\ \kappa_4 &= 6 \sum n^{-4} = \pi^4/15 = 6,494. \end{aligned}$$

Вывести отсюда, что $\beta_1 = 1,299$, $\beta_2 = 5,4$.

МНОГОМЕРНОЕ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ И КВАДРАТИЧНЫЕ ФОРМЫ

15.1 Об одномерном нормальном распределении

$$dF = (2\pi)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\} \frac{dx}{\sigma}, \quad -\infty \leq x \leq \infty,$$

много говорилось в предыдущих главах. Если отвлечься от параметров сдвига и масштаба μ и σ , то показатель экспоненты в его функции плотности будет равен просто $-\frac{1}{2}x^2$. Нашей те-
перешней задачей является обобщение одномерного нормального распределения на двумерный и вообще многомерный случай. Естественно искать это обобщение в виде функции плотности, являющейся экспонентой квадратичной формы от p переменных, т. е. в виде

$$dF \propto \exp \left\{ -\frac{1}{2} \sum_{j=1}^p \sum_{k=1}^p a_{jk} \left(\frac{x_j - \mu_j}{\sigma_j} \right) \left(\frac{x_k - \mu_k}{\sigma_k} \right) \right\} \prod_{j=1}^p \frac{dx_j}{\sigma_j}, \quad (15.1)$$

$$-\infty \leq x_j \leq \infty.$$

Здесь, как и в одномерном случае, μ_j и σ_j — параметры сдвига и масштаба, которые с помощью нормирующего преобразования можно сделать равными соответственно 0 и 1.

В этой главе, как всегда при рассмотрении многомерного случая, нам будет удобно использовать векторные и матричные обозначения; благодаря этому достигается экономность записи и большая ясность. Жирные прописные буквы будут обозначать матрицы, жирные строчные буквы — векторы-столбцы; штрихом обозначается операция транспонирования, так что жирные строчные буквы со штрихом представляют векторы-строки. При этих

соглашениях (15.1) принимает вид

$$dF \propto \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{x} - \boldsymbol{\mu}) \right\} \prod_j dx_j, \quad (15.2)$$

где $\mathbf{A}$ — симметрическая $(p \times p)$ -матрица.

Характеристическая функция и моменты

15.2 Соотношение (15.2) действительно определяет функцию плотности, если квадратичная форма $(\mathbf{x} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{x} - \boldsymbol{\mu})$ неотрицательно определена. В самом деле, в этом случае (и только в этом случае) можно найти действительное линейное преобразование, переводящее показатель экспоненты в сумму квадратов со знаком минус, и интеграл от dF будет сходиться. Более того, существует ортогональное преобразование

$$\mathbf{x} = \mathbf{B}\mathbf{y}, \quad (15.3)$$

преобразующее показатель экспоненты во взятую со знаком минус сумму квадратов с коэффициентами, являющимися корнями характеристического уравнения матрицы $\mathbf{A}$. Ограничиваясь невырожденным случаем, когда $\text{rang } \mathbf{A}$ равен *) p , и приписывая за начало отсчета вектор $\boldsymbol{\mu}$, мы воспользуемся этим преобразованием для нахождения характеристической функции распределения (15.2).

Попутно будет вычислена константа, на которую нужно умножить правую часть (15.2), чтобы интеграл от dF был равен единице.

Характеристическая функция в наших новых обозначениях выражается формулой (ср. (4.27))

$$\varphi(\mathbf{t}) \propto \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left(i\mathbf{t}'\mathbf{x} - \frac{1}{2} \mathbf{x}'\mathbf{A}\mathbf{x} \right) \prod_j dx_j, \quad (15.4)$$

или, после подстановки (15.3),

$$\varphi(\mathbf{t}) \propto \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left(i\mathbf{t}'\mathbf{B}\mathbf{y} - \frac{1}{2} \mathbf{y}'\mathbf{C}\mathbf{y} \right) \prod_j dy_j; \quad (15.5)$$

в силу ортогональности $\mathbf{B}$ якобиан равен единице. В (15.5) мы приняли обозначение

$$\mathbf{C} = \mathbf{B}'\mathbf{A}\mathbf{B}. \quad (15.6)$$

*) Если $\mathbf{A}$ имеет ранг $r < p$, то распределение вырождается и оказывается сосредоточенным на подпространстве размерности $r < p$, т. е. одна или несколько переменных излишни.

Так как C — диагональная матрица, то правая часть (15.5) распадается в произведение p однократных интегралов вида

$$I_j = \int_{-\infty}^{\infty} \exp\left(iu_j y_j - \frac{1}{2} c_{jj} y_j^2\right) dy_j,$$

где u_j — линейные функции от переменных t_j . Значение I_j легко определяется по формуле характеристической функции одномерного нормального распределения (пример 3.11). Это значение равно

$$I_j = \left(\frac{2\pi}{c_{jj}}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{u_j^2}{c_{jj}}\right). \quad (15.7)$$

Применяя формулу (15.7) для всех сомножителей в (15.5), получаем

$$\varphi(t) \propto (2\pi)^{\frac{p}{2}} \left(\prod_j c_{jj}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \sum_j \frac{u_j^2}{c_{jj}}\right). \quad (15.8)$$

Детерминант диагональной матрицы C есть

$$|C| = \prod_j c_{jj},$$

а детерминант обратной к ней матрицы C^{-1} равен

$$|C^{-1}| = |C|^{-1} = \left(\prod_j c_{jj}\right)^{-1}. \quad (15.9)$$

Подставляя (15.9) в (15.8) и переписывая показатель экспоненты в матричной форме, имеем

$$\varphi(t) \propto (2\pi)^{\frac{1}{2}p} |C^{-1}|^{\frac{1}{2}} \exp\left(-\frac{1}{2} t' B C^{-1} B' t\right). \quad (15.10)$$

Так как в силу (15.6)

$$B C^{-1} B' = B (B' A B)^{-1} B' = B B^{-1} A^{-1} (B')^{-1} B' = A^{-1} \quad (15.11)$$

и матрица B ортогональна, то

$$|A^{-1}| = |B C^{-1} B'| = |B| |C^{-1}| |B'| = |C^{-1}|. \quad (15.12)$$

Используя (15.11) и (15.12), мы можем переписать теперь (15.10) в виде

$$\varphi(t) \propto (2\pi)^{\frac{1}{2}p} |A^{-1}|^{\frac{1}{2}} \exp\left(-\frac{1}{2} t' A^{-1} t\right). \quad (15.13)$$

Полагая в (15.13) $t=0$, находим, что константа в плотности многомерного нормального распределения равна

$$(2\pi)^{-\frac{1}{2}P} |A^{-1}|^{-\frac{1}{2}} = (2\pi)^{-\frac{1}{2}P} |A|^{\frac{1}{2}}.$$

Таким образом, уточняя (15.2), можем написать

$$dF = (2\pi)^{-\frac{1}{2}P} |A|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' A (\mathbf{x} - \boldsymbol{\mu}) \right\} \Pi dx_j. \quad (15.14)$$

Соответствующая характеристическая функция имеет вид

$$\varphi(t) = \exp \left(-\frac{1}{2} t' A^{-1} t \right) \exp(it' \boldsymbol{\mu}). \quad (15.15)$$

Второй множитель в правой части (15.15) появляется в том случае, когда мера сдвига $\boldsymbol{\mu}$ не равна нулю. В этом легко убедиться, замстив, что именно на этот множитель умножается (15.5), если во втором слагаемом показателя экспоненты в (15.4) и в (15.3) $\mathbf{x}$ заменить на $\mathbf{x} - \boldsymbol{\mu}$.

15.3 Из формулы для характеристической функции нормального распределения (15.15) легко найти его моменты n , еще проще, семиинварианты. Производящая функция семиинвариантов имеет вид

$$\psi(t) = -\frac{1}{2} t' A^{-1} t + it' \boldsymbol{\mu}. \quad (15.16)$$

Так как правая часть (15.16) не содержит членов степени выше второй, то, как и в одномерном случае, все семиинварианты порядка выше второго равны нулю. Среднее значение j -й случайной величины равно

$$\left[\frac{\partial \psi(t)}{\partial (it_j)} \right]_{t=0} = \mu_j,$$

так что вектор сдвига $\boldsymbol{\mu}$ есть вектор средних значений случайных величин. Дисперсия j -й величины равна

$$\left[\frac{\partial^2 \psi(t)}{\partial (it_j)^2} \right]_{t=0} = a_{jj}^{-1}, \quad (15.17)$$

т. е. совпадает с элементом, стоящим в j -й строке и j -м столбце матрицы A^{-1} . Ковариация j -й и k -й величин равна

$$\left[\frac{\partial^2 \psi(t)}{\partial (it_j) \partial (it_k)} \right]_{t=0} = a_{jk}^{-1}. \quad (15.18)$$

Формулы (15.17) и (15.18) показывают, таким образом, что матрица центральных вторых моментов случайных величин, совместное распределение которых имеет х. ф. (15.15), есть A^{-1} . Эту матрицу будем называть *матрицей рассеяния*, или *диспер-*

сионной матрицей. Теперь мы можем принять более удобное обозначение, заменив A^{-1} на V , и в результате получим окончательное выражение для p -мерного нормального распределения в терминах его первых и вторых моментов:

$$dF = (2\pi)^{-\frac{1}{2}p} |V^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' V^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \Pi d\mathbf{x}. \quad (15.19)$$

Соответствующая характеристическая функция принимает вид

$$\varphi(\mathbf{t}) = \exp \left(-\frac{1}{2} \mathbf{t}' V \mathbf{t} \right) \exp (i \mathbf{t}' \boldsymbol{\mu}). \quad (15.20)$$

Пример 15.1 Двумерное нормальное распределение

Мы уже встречались с двумерным вариантом формул (15.19) и (15.20) в примере 3.17 и в ряде других мест. В двумерном случае матрица вторых моментов имеет вид

$$V = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix},$$

где ρ — коэффициент корреляции, равный $\mu_{11}/\sigma_1\sigma_2$. Следовательно,

$$V^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{pmatrix},$$

так что если средние значения равны нулю, то (15.19) можно записать в виде

$$dF = \frac{dx_1 dx_2}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2} \right) \right\},$$

а (15.20) — в виде

$$\varphi(t_1, t_2) = \exp \left\{ -\frac{1}{2} (\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2) \right\}.$$

Линейные функции от нормально распределенных случайных величин

15.4 В примере 11.12 мы видели, что линейная функция от независимых нормально распределенных случайных величин сама распределена нормально. Теперь мы в состоянии получить много более общий результат, а именно, совместное распределение всякого множества линейных функций от совокупности случайных

величин с нормальным совместным распределением нормально. Характеристическая функция невырожденного p -мерного нормального распределения выражается формулой (15.20), которую при $\mu=0$ можно записать в виде

$$M\{\exp(it'x)\} = \exp\left(-\frac{1}{2}t'Vt\right). \quad (15.21)$$

Пусть

$$t' = s'A, \quad (15.22)$$

где s' есть q -компонентный вектор, t' — вектор с p компонентами и A — некоторая $(q \times p)$ -матрица. Подставляя (15.22) в (15.21), получаем

$$M\{\exp(is'A x)\} = \exp\left(-\frac{1}{2}s'AVA's\right). \quad (15.23)$$

Соотношение (15.23) показывает, что распределение множества q линейных функций от p случайных величин с нормальным совместным распределением, определяемых соотношением $y = Ax$, имеет характеристическую функцию многомерного нормального распределения с матрицей рассеяния AVA' . Если AVA' — невырожденная матрица, то многомерная формула обращения (4.17) дает многомерную плотность (15.19) для y с AVA вместо V .

Если матрица AVA' вырождена, то обращение невозможно, так как тогда матрица $(AVA')^{-1}$ не определена. Тем не менее и в этом случае (15.23) есть характеристическая функция (вырожденного) многомерного нормального распределения. В частности, так будет при $q > p$. С этой точки зрения ясно, что всякое вырожденное многомерное нормальное распределение можно рассматривать как результат некоторого линейного преобразования невырожденного нормального распределения. В практике вырожденные распределения не вызывают никаких трудностей, поскольку вырожденное распределение всегда можно свести к невырожденному, элиминируя излишние переменные.

Важным следствием результатов этого пункта является тот факт, что любое частное (маргинальное) распределение нормального распределения нормально. Чтобы в этом убедиться, достаточно в качестве A взять $p \times p$ -матрицу, у которой все элементы равны нулю, за исключением q , $q \leq p-1$, элементов на главной диагонали, равных единице. Матрицей вторых моментов частного распределения является соответствующая часть матрицы вторых моментов исходного распределения. Однако обратная к ней матрица, определяющая коэффициенты показателя экспоненты в функции плотности, не будет, конечно, соответствующей частью аналогичной матрицы для исходного распределения.

Многомерный нормальный интеграл

15.5 Помимо параметров сдвига и масштаба, функция плотности (15.19) содержит еще $\frac{1}{2}p(p-1)$ параметров — ковариации p величин. Ясно, что табулирование соответствующей функции распределения при $p > 2$ — дело чрезвычайно сложное. В *Tables of the Bivariate Normal Distribution and Related Functions* (National Bureau of Standards, Applied Mathematics Series, 50, Washington, 1959) приводятся таблицы функции двумерного нормального распределения (с шестью или семью десятичными знаками) для $\pm\rho=0(0,05)0,95(0,01)1$ и значений основных переменных $0(0,1)4$. Зелен и Северо (Zelen and Severo, 1960, *Ann. Math. Statist.* 31, 619) построили номограмму для нахождения значений двумерного нормального интеграла с ошибкой, не превосходящей одного процента.

Кендалл (1941) нашел выражение функции нормального распределения в терминах бесконечного ряда, однако сходимость его слишком медленная, чтобы выражение это можно было широко использовать. В общем случае Плэкеттом (1954) была предложена формула редукции, позволяющая удовлетворительно табулировать нормальное распределение в трех- и четырехмерных случаях с помощью табулированных интегралов.

Оуэн (D. B. Owen (1957), *The Bivariate Normal Probability Integral*, Sandia Corporation, Washington, D. C., а также (1956) *Ann. Math. Statist.* 27, 1075) создал таблицы для вычисления двумерного нормального интеграла по областям, ограниченным прямыми линиями, а также по некоторым другим областям.

Стек (G. P. Steck (1958), *Ann. Math. Statist.* 29, 780) предложил таблицы для вычисления трехмерного интеграла *).

15.6 В статистике часто возникает специальная задача вычисления многомерного нормального интеграла по области, в которой все p компонент положительны. Хотя эта задача значительно проще общей, решение ее при $p > 3$ по-прежнему сталкивается со значительными аналитическими трудностями, для преодоления которых в отдельных случаях было предложено много специальных приемов.

*) Таблицы Оуэна предназначены для вычисления двумерного нормального интеграла по многоугольнику. Вместе с таблицами Стека они воспроизведены в книге *Сборник статистических таблиц*. Более подробно функция Оуэна табулирована в книге: Н. В. Смирнов, Л. Н. Большев (1962), *Таблицы для вычисления функции двумерного нормального распределения*, М., Изд-во АН СССР. (Прим. ред.)

Сначала мы рассмотрим двух- и трехмерные случаи — только в этих случаях известны аналитические решения рассматриваемой задачи.

Пусть число измерений равно двум. Интеграл от функции плотности по положительному квадранту, если ее представить в виде обратного преобразования Фурье соответствующей ей характеристической функции, имеет вид

$$P_2^0 = \int_0^\infty \int_0^\infty \left\{ \frac{1}{(2\pi)^2} \int_{-\infty}^\infty \int_{-\infty}^\infty \exp\left(-it'x - \frac{1}{2}t'Vt\right) dt_1 dt_2 \right\} dx_1 dx_2; \quad (15.24)$$

векторы и матрица здесь — двумерные *). Изменив в (15.24) порядок интегрирования и произведя затем интегрирование по x_1 и x_2 , получим

$$4\pi^2 P_2^0 = \int_{-\infty}^\infty \int_{-\infty}^\infty \exp\left(-\frac{1}{2}t'Vt\right) \frac{dt_1}{it_1} \frac{dt_2}{it_2}. \quad (15.25)$$

Коэффициент при $(-t_1 t_2)$ в показателе экспоненты, входящей в подынтегральное выражение, равен ковариации случайных величин, т. е. равен $\rho\sigma_1\sigma_2$ (см. пример 15.1). Дифференцируя (15.25) по ρ , находим

$$4\pi^2 \frac{\partial P_2^0}{\sigma_1\sigma_2 \partial \rho} = \int_{-\infty}^\infty \int_{-\infty}^\infty \exp\left(-\frac{1}{2}t'Vt\right) dt_1 dt_2. \quad (15.26)$$

Подынтегральное выражение правой части соотношения (15.26) с точностью до постоянного множителя представляет собой функцию плотности двумерного нормального распределения. Интеграл в (15.26) равен поэтому единице, деленной на этот постоянный множитель. Следовательно (см. пример 15.1),

$$\frac{\partial P_2^0}{\sigma_1\sigma_2 \partial \rho} = \frac{1}{4\pi^2} \frac{2\pi}{|V|^{\frac{1}{2}}} = \frac{1}{2\pi |V|^{\frac{1}{2}}} = \left\{ 2\pi\sigma_1\sigma_2 (1 - \rho^2)^{\frac{1}{2}} \right\}^{-1}. \quad (15.27)$$

Интегрируя уравнение (15.27), получаем

$$P_2^0 = \arcsin \rho / (2\pi) + c.$$

*) Здесь предполагается, что средние компонент равны нулю. (Прим. перев.)

Положив $\rho=0$, убеждаемся, что $c=\frac{1}{4}$. Таким образом *),

$$P_2^0 = \frac{1}{4} + \arcsin \rho / (2\pi). \quad (15.28)$$

Соотношение (15.28) называют обычно теоремой Шеппарда (1898) о медианной дихотомии. В силу симметрии то же значение имеет интеграл по отрицательному квадранту. Интеграл по каждому из остальных двух квадрантов равен, очевидно, $(\frac{1}{2} - P_2^0)$. National Bureau of Standards (1959) табулировало величину P_2^0 для $\rho=0(0,01)1$ (см. ссылку в 15.5).

15.7 Используя полученный результат, легко найти значение интересующего нас интеграла в трехмерном случае. Как известно (упражнение 7.1), вероятность появления по меньшей мере одного из совокупности p событий равна

$$\sum P_1 - \sum P_2 + \sum P_3 - \dots + (-1)^{p-1} P_p. \quad (15.29)$$

Здесь первая из сумм есть сумма вероятностей отдельных событий, вторая — сумма вероятностей одновременного осуществления (всевозможных) пар событий и т. д., последний член P_p есть вероятность одновременного осуществления всех p событий.

Пусть каждое из этих p событий осуществляется тогда и только тогда, когда положительна соответствующая случайная величина, причем все эти p величин имеют совместное нормальное распределение. Событие, состоящее в том, что по меньшей мере одна из величин положительна, является дополнительным к событию, состоящему в том, что ни одна из величин не положительна. В силу симметрии вероятность этого последнего события совпадает с искомой вероятностью неотрицательности всех величин. Отсюда и из (15.29) вытекает, следовательно, что

$$1 - P_p^0 = \sum P_1^0 - \sum P_2^0 + \sum P_3^0 - \dots + (-1)^{p-1} P_p^0 \quad (15.30)$$

(как и выше, мы здесь приписываем к вероятностям индекс 0 сверху). Если p чётно, то последний член справа отрицателен, и вероятность P_p^0 из (15.30) определить невозможно; если p

*) Данный здесь вывод соотношения (15.28) нестрогий: в нем используются расходящиеся интегралы. Строгий вывод этого соотношения можно найти в книге: Г. Крамёр, Математические методы статистики, стр. 320—321. (Прим. перев.)

нечетно, то (15.30) приводится к виду

$$P_p^0 = \frac{1}{2} \left\{ 1 - \sum P_1^0 + \sum P_2^0 - \dots + (-1)^{p-1} \sum P_{p-1}^0 \right\}. \quad (15.31)$$

Соотношение (15.31) позволяет находить при нечетных p вероятность P_p^0 по соответствующим вероятностям низших порядков. В частности, при $p=3$ имеет место следующее простое обобщение формулы Шеппарда:

$$\begin{aligned} P_3^0 &= \frac{1}{2} \left\{ 1 - \sum P_1^0 + \sum P_2^0 \right\} = \\ &= \frac{1}{2} \left\{ 1 - 3 \cdot \frac{1}{2} + \sum \left(\frac{1}{4} + \arcsin \rho / (2\pi) \right) \right\} = \\ &= \frac{1}{8} + \frac{1}{4\pi} (\arcsin \rho_{12} + \arcsin \rho_{13} + \arcsin \rho_{23}). \end{aligned} \quad (15.32)$$

15.8 При $p > 3$ таких простых результатов не получается. В общем случае наилучший из найденных к настоящему времени результатов принадлежит Кендаллу (1941) и состоит в разложении P_p^0 в степенной ряд по ковариациям. Без ограничения общности можно предположить дисперсии всех маргинальных одномерных распределений равными единице. Тогда обратное преобразование Фурье характеристической функции (ср. (15.24)) можно разложить в степенной ряд по коэффициентам корреляции ρ_{ij} . Именно *),

$$\begin{aligned} P_p^0 &= \int_0^\infty \dots \int_0^\infty \left\{ \frac{1}{(2\pi)^p} \int_{-\infty}^\infty \dots \int_{-\infty}^\infty \exp \left(-it'x - \frac{1}{2} t'Vt \right) dt_1 \dots dt_p \right\} \times \\ &\times dx_1 \dots dx_p = \int_0^\infty \dots \int_0^\infty \left\{ \frac{1}{(2\pi)^p} \int_{-\infty}^\infty \dots \int_{-\infty}^\infty \exp \left(-it'x - \frac{1}{2} t't \right) \times \right. \\ &\times \sum \left[(-1)^{n_{..}} \frac{\rho_{12}^{n_{12}} \rho_{13}^{n_{13}} \dots}{n_{12}! n_{13}! \dots} \right] t_1^{n_{11}} t_2^{n_{22}} \dots t_p^{n_{pp}} dt_1 \dots dt_p \Big] dx_1 \dots dx_p, \end{aligned} \quad (15.33)$$

где суммирование ведется по всем возможным множествам величин ρ_{ij} и всем целым неотрицательным значениям n_{ij} ; $n_{i.} = \sum_j (n_{ij} + n_{ji})$, $n_{..} = \sum_i n_{i.}$. Меняя в (15.33) порядок суммиро-

*) Приводимое ниже рассуждение, подобно рассуждению в 15.6, нестрогое. (Прим. перев.)

вания и интегрирования, получаем

$$P_p^0 = \sum \left[(-1)^{n \dots} \frac{\rho_{12}^{n_{12}} \rho_{13}^{n_{13}} \dots}{n_{12}! n_{13}! \dots} \right] \prod_{j=1}^p G_{n_j}, \quad (15.34)$$

где

$$G_{n_j} = \frac{1}{2\pi} \int_0^\infty \int_{-\infty}^\infty t_j^{n_j} \exp \left(-it_j x_j - \frac{1}{2} t_j^2 \right) dt_j dx_j. \quad (15.35)$$

Ясно, что $G_{n_j} = 1/2$ при $n_j = 0$. Когда n_j — целое положительное, интегрируя в (15.35) по x_j , как мы уже делали в (15.24), находим

$$G_{n_j} = \frac{1}{2\pi i} \int_{-\infty}^\infty t_j^{n_j-1} e^{-\frac{1}{2} t_j^2} dt_j.$$

Применение формул для моментов одномерного нормального распределения (5.60) приводит теперь к следующим соотношениям:

$$G_{n_j} = \begin{cases} 0 & \text{при } n_j \text{ четном,} \\ \frac{(2m)!}{\sqrt{2\pi} i 2^m m!} & \text{при } n_j \text{ нечетном, } n_j = 2m + 1. \end{cases}$$

Таким образом, окончательно

$$G_{n_j} = \begin{cases} \frac{1}{2} & \text{при } n_j = 0, \\ 0 & \text{при } n_j = 2m \quad (m = 1, 2, \dots), \\ \frac{(2m)!}{\sqrt{2\pi} i 2^m m!} & \text{при } n_j = 2m + 1 \quad (m = 0, 1, 2, \dots). \end{cases} \quad (15.36)$$

Из этих формул следует, что в разложении (15.34) те члены, которым отвечает хотя бы одно четное n_j , равны нулю.

Формально соотношения (15.34) и (15.36) дают решение задачи. Однако хотя ряд (15.34) всегда сходится, сходимость его медленная, если ρ_{ij} не малы. Более того, не все члены этого ряда положительны. В четырехмерном случае (см., например, Морэн (Morgan, 1948), *Biometrika* 35, 203) первые члены в (15.34) имеют вид

$$P_4^0 = \frac{1}{16} + \frac{1}{8\pi} \sum_{i,j} \rho_{ij} + \frac{1}{4\pi^2} \sum_{ijkl} \rho_{ij} \rho_{kl} + \dots,$$

а коэффициент при $\rho_{12}\rho_{13}\rho_{14}$ отрицателен и равен $-1/(4\pi)^2$.

15.9 В частном случае, когда все коэффициенты корреляции равны между собой, задача допускает совсем простое решение.

Пусть $x_0, x_1, \dots, x_p$ есть $p+1$ независимых нормально распределенных случайных величин с нулевыми средними и единичными дисперсиями. Рассмотрим величины

$$y_i = x_i - b x_0, \quad i = 1, \dots, p, \quad b > 0.$$

Совместное распределение их нормально, каждая из них имеет среднее нуль, дисперсию $1+b^2$, и ковариация любых двух из них равна b^2 . Коэффициент корреляции для всякой пары игроков равен, следовательно,

$$\rho = b^2/(1+b^2). \quad (15.37)$$

Когда b изменяется от 0 до бесконечности, ρ изменяется от 0 до 1. Поэтому, выбирая соответствующим образом b , в качестве распределения игроков можно получить любое p -мерное нормальное распределение, для которого коэффициенты корреляции одинаковы и положительны. Значение соответствующей вероятности P_p^0 легко находится из самой конструкции игроков; она равна вероятности того, что p независимых величин с нормированным нормальным распределением превосходят bx , проинтегрированной по нормированному нормальному распределению:

$$P_p^0 = \int_{-\infty}^{\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \left\{ \int_{bx}^{\infty} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}t^2} dt \right\}^p dx. \quad (15.38)$$

Стоящий здесь справа интеграл является интегралом типа Хойо (14.5), и с небольшой ошибкой его можно просто вычислить, заменяя интегрирование суммированием. Очевидно, вышеприведенные рассуждения, по существу, не изменятся, если некоторые из y_i определять равенством

$$y_i = x_i + bx_0;$$

при этом некоторые пары y_i будут иметь отрицательный коэффициент корреляции, равный $-\rho$. В этом случае (15.38) должно быть, конечно, соответствующим образом изменено.

Рубин (1954) табулировал величину P_p^0 в случае одинаковых коэффициентов корреляции для $\rho=1/j$, $p=1, 2, \dots, 51-j$, где $j=2, 3, \dots, 12$.

Пример 15.2

Чтобы проиллюстрировать аппроксимацию интеграла (15.38) суммами, проведем вычисления при $p=2$ и $b=1$. Здесь $\rho=\frac{1}{2}$. Заменяя интеграл суммой значений подынтегральной функции в точках $x=\pm 3, \pm 2, \pm 1, 0$ и используя таблицы 1 и 2 приложения, получаем результаты, указанные на стр. 489.

Таким образом, приближенное значение интеграла при этой аппроксимации с точностью до четвертого знака равно 0,3331. Точное его значение есть $1/3$. Можно было бы достигнуть большей точности, беря сумму значений подынтегральной функции в большем числе точек. Например, если добавить к рассмотрен-

x	$(1) = \int_x^\infty (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}t^2} dt$	$(2) = \{(1)\}^2$	$(3) = \text{сумма чисел (2)}$	$(4) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2}$	$(3) \times (4)$
-3	0,99865	0,99730	0,99730	0,00443	0,00442
+3	0,00135	0,00000			
-2	0,97725	0,95502	0,95554	0,05399	0,05159
+2	0,02275	0,00052			
-1	0,84134	0,70785	0,73302	0,24197	0,17737
+1	0,15866	0,02517			
0	0,50000	0,25000	0,25000	0,39894	0,09973
					0,33311

ной сумме еще пару членов, соответствующих $x = \pm 4$, то она увеличится на 0,00013, и в результате получится 0,33324. Добавление новых членов, соответствующих значениям $x = \pm 5$, не увеличивает точности, если ограничиться пятью знаками, однако ее можно сколь угодно сильно повысить, беря суммы, отвечающие более мелким интервалам разбиения области интегрирования (например, беря интервалы длиной 0,5 вместо 1).

Морэном (1956) была найдена граница для относительной ошибки при аппроксимации интеграла (15.38) суммами. В наших обозначениях она равна

$$2 \exp \left\{ -\frac{\pi^2}{h^2 (1 + pb^2)} \right\},$$

где h — длина интервалов разбиения области интегрирования. В рассмотренном примере значение этой границы есть

$$2 \exp \left(-\frac{\pi^2}{3h^2} \right).$$

При $h=1$ граница значительно превышает ошибку, имевшую место выше при аппроксимации суммой с девятью членами и равную примерно 0,0003. С убыванием h граница резко уменьшается и, по-видимому, становится близкой к истинной ошибке.

Квадратичные формы от нормально распределенных случайных величин

15.10 Пусть вектор $\mathbf{x}$ имеет нормальное распределение

$$dF \propto \exp \left(-\frac{1}{2} \mathbf{x}' \mathbf{A} \mathbf{x} \right) \prod_{j=1}^p dx_j$$

и матрица $\mathbf{A}$ не обязательно невырождена. Рассмотрим распределение самой квадратичной формы $\mathbf{x}' \mathbf{A} \mathbf{x}$. Для соответствующей

характеристической функции имеем формулу

$$\varphi(t) \propto \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \mathbf{x}' \mathbf{A} \mathbf{x} (1 - \theta) \right\} \prod_j dx_j, \quad (15.39)$$

где $\theta = it$. Сделаем, как в 15.2, преобразование переменных (15.3), переводящее квадратичную форму $\mathbf{x}' \mathbf{A} \mathbf{x}$ к диагональному виду. При этом (15.39) перейдет в формулу

$$\varphi(t) \propto \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \mathbf{y}' \mathbf{C} \mathbf{y} (1 - \theta) \right\} \prod_j dy_j. \quad (15.40)$$

Число ненулевых компонент у вектора $\mathbf{y}$ в точности равно числу ненулевых корней r характеристического уравнения матрицы $\mathbf{A}$. Если в (15.40) проинтегрировать по $p - r$ переменным, соответствующим нулевым корням характеристического уравнения, то стоящая под знаком интеграла функция не изменится. Произведем это интегрирование и в оставшемся r -мерном интеграле произведем замену $z_j = y_j c_{jj} (1 - \theta)^{\frac{1}{2}}$. В результате получим

$$\varphi(t) \propto (1 - \theta)^{-\frac{1}{2}r} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left(-\frac{1}{2} \mathbf{z}' \mathbf{z} \right) \prod_j dz_j. \quad (15.41)$$

Так как интеграл в (15.41) не зависит от θ , то

$$\varphi(t) = (1 - \theta)^{-\frac{1}{2}r}, \quad (15.42)$$

поскольку согласно определению характеристической функции $\varphi(0) = 1$. Из формулы обращения, примера 3.6 и равенства (15.42) вытекает, что распределение квадратичной формы, стоящей в показателе экспоненты функции плотности многомерного нормального распределения, есть χ^2 -распределение с r степенями свободы, задаваемое формулой (11.36) с $n - p = r$ (см. также главу 16), где r — ранг (квадратичной формы) распределения. Полученный результат является обобщением результата из 11.2, где, по существу, рассматривался случай, когда матрица $\mathbf{A}$ диагональна и невырождена.

Пример 15.3 Критерий согласия и распределение χ^2

Пусть имеется выборка объема n из генеральной совокупности с мультиномиальным распределением (см. 5.30) задаваемым вероятностями p_i ($i = 1, \dots, k$). Из многомерной центральной предельной теоремы следует, что распределение соответ-

ствующих k группам чисел x_i будет стремиться при возрастании n к многомерному нормальному распределению со средними np_i и матрицей рассеяния (см. (5.80))

$$V = n \begin{pmatrix} p_1(1-p_1) & -p_1p_2 & -p_1p_3 & \dots & -p_1p_k \\ -p_2p_1 & p_2(1-p_2) & -p_2p_3 & \dots & -p_2p_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -p_kp_1 & -p_kp_2 & \dots & \dots & p_k(1-p_k) \end{pmatrix}.$$

В силу имеющегося линейного соотношения $\sum_{j=1}^k x_j = n$ это нормальное распределение является вырожденным и имеет ранг $k-1$, так что $|V|=0$ и матрица V необратима. Но одну из переменных x_i (скажем x_k) можно отбросить как излишнюю. Оставшиеся $k-1$ переменных x_i распределены нормально со средними np_i и матрицей рассеяния V^* , получающейся из V удалением последнего столбца и последней строки. Простым перемножением легко убедиться в том, что

$$(V^*)^{-1} = \frac{1}{n} \begin{pmatrix} \frac{1}{p_1} + \frac{1}{p_k} & \frac{1}{p_k} & \frac{1}{p_k} & \dots & \frac{1}{p_k} \\ \frac{1}{p_k} & \frac{1}{p_2} + \frac{1}{p_k} & \frac{1}{p_k} & \dots & \frac{1}{p_k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{p_k} & \frac{1}{p_k} & \frac{1}{p_k} & \dots & \frac{1}{p_{k-1}} + \frac{1}{p_k} \end{pmatrix}.$$

Квадратичная форма в экспоненте функции плотности этого распределения равна, следовательно,

$$\begin{aligned} (x - \mu)' (V^*)^{-1} (x - \mu) &= \\ &= \sum_{i=1}^{k-1} \frac{(x_i - np_i)^2}{np_i} + \frac{1}{np_k} \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} (x_i - np_i)(x_j - np_j) = \\ &= \sum_{i=1}^{k-1} \frac{(x_i - np_i)^2}{np_i} + \frac{1}{np_k} \left\{ \sum_{i=1}^{k-1} (x_i - np_i) \right\}^2 = \sum_{i=1}^k \frac{(x_i - np_i)^2}{np_i}. \end{aligned}$$

Согласно 15.10 эта квадратичная форма имеет χ^2 -распределение с $k - 1$ степенями свободы. Полученный результат, принадлежащий Карлу Пирсону (1900), служит основой критерия согласия χ^2 , который будет обсуждаться во втором томе.

15.11 Предположим теперь, что компоненты p -мерного вектора $\mathbf{x}$ независимы и имеют нормированное нормальное распределение. Рассмотрим распределение квадратичной формы общего вида

$$Q = \mathbf{x}' \mathbf{A} \mathbf{x}.$$

Ортогональным преобразованием (15.3), переводящим $\mathbf{x}$ в вектор $\mathbf{y}$ с независимыми и нормированно нормально распределенными компонентами, Q можно преобразовать к виду

$$Q = \sum_{i=1}^r a_i y_i^2, \quad r \leq p, \quad (15.43)$$

где a_i — ненулевые корни характеристического уравнения матрицы $\mathbf{A}$. Так как величина $a_i y_i^2$ пропорциональна случайной величине, имеющей χ^2 -распределение с одной степенью свободы, то соответствующая ей производящая функция семиинвариантов равна (ср. 15.10 и главу 16)

$$\psi_i(t) = -\frac{1}{2} \log(1 - 2a_i \theta), \quad (15.44)$$

где $\theta = it$. Производящая функция семиинвариантов формы Q в силу аддитивного свойства этой функции (15.43) и (15.44) есть

$$\psi(t) = -\frac{1}{2} \sum_{i=1}^r \log(1 - 2a_i \theta). \quad (15.45)$$

С помощью формулы (15.45) легко найти общую формулу для семиинвариантов распределения формы Q :

$$\kappa_s = 2^{s-1} (s-1)! \sum_{i=1}^r a_i^s. \quad (15.46)$$

Выражение $\sum_{i=1}^r a_i^s$, равное сумме элементов главной диагонали s -й степени матрицы $\mathbf{A}$, будем записывать в виде $\text{tr } \mathbf{A}^s$. При этом (15.46) переходит в

$$\kappa_s = 2^{s-1} (s-1)! \text{tr } \mathbf{A}^s. \quad (15.47)$$

Из (15.46) ясно, что распределение формы Q зависит только от значений ненулевых корней характеристического уравнения матрицы $\mathbf{A}$.

Из сравнения (15.46) с семинвариантами χ^2 -распределения (они приведены в главе 16; их легко можно вычислить, исходя из (15.42)) видно, что Q имеет χ^2 -распределение с r степенями свободы тогда и только тогда, когда все $a_i = 1$. (При этом A — идемпотентная матрица, т. е. $A = A^s$ для всех s .)

15.12 Характеристическая функция формы Q , соответствующая производящей функции семинвариантов (15.45), выражается формулой

$$\begin{aligned} \varphi(t) &= M \{ \exp(0x'Ax) \} = \\ &= \frac{1}{(2\pi)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} x' [I - 2\theta A] x \right\} dx_1 \dots dx_n. \end{aligned} \quad (15.48)$$

Правая часть (15.48) с точностью до множителя $|V|^{-\frac{1}{2}}$ представляет собой интеграл типа интеграла от плотности нормального распределения с $V^{-1} = I - 2\theta A$. Следовательно, значение правой части (15.48) есть $|V|^{\frac{1}{2}}$, так что

$$\varphi(t) = |I - 2\theta A|^{-\frac{1}{2}}. \quad (15.49)$$

Аналогично выводится формула характеристической функции совместного распределения двух квадратичных форм

$$Q_1 = x'Ax, \quad Q_2 = x'Bx.$$

Вычисляя математическое ожидание величины $\exp(\theta_1 Q_1 + \theta_2 Q_2)$, получаем

$$\psi(t_1, t_2) = |I - 2\theta_1 A - 2\theta_2 B|^{-\frac{1}{2}}. \quad (15.50)$$

Соответствующая производящая функция семинвариантов равна

$$\psi(t_1, t_2) = \sum_{s=1}^{\infty} 2^{s-1} \operatorname{tr}(\theta_1 A + \theta_2 B)^s / s. \quad (15.51)$$

В 4.16 мы видели, что для независимости случайных величин необходимо и достаточно, чтобы характеристическая функция их совместного распределения совпадала с произведением характеристических функций каждой из этих величин в отдельности. Сравнивая (15.49) и (15.50), убеждаемся, что формы Q_1 и Q_2 независимы тогда и только тогда, когда

$$|I - 2\theta_1 A| |I - 2\theta_2 B| = |I - 2\theta_1 A - 2\theta_2 B| \quad (15.52)$$

или, эквивалентно,

$$|\lambda I - 2\theta_1 A| |\lambda I - 2\theta_2 B| = \lambda^p |\lambda I - 2\theta_1 A - 2\theta_2 B|. \quad (15.53)$$

Этот результат принадлежит Кокрэнгу (1934).

15.13 Другой способ нахождения условий независимости двух квадратичных форм Q_1 и Q_2 основывается на признаке независимости двух случайных величин, состоящем в равенстве нулю их смешанных семинвариантов (пример 12.7). Независимость Q_1 и Q_2 равносильна равенству нулю коэффициентов при $\theta_1^k \theta_2^l$ в производящей функции семинвариантов (15.51) при $k, l \geq 1$, а это в свою очередь равносильно равенствам *)

$$\text{tr}(A^k B^l A^m B^n \dots) = 0 \quad (\text{все } k, l, m, n, \dots \geq 1). \quad (15.54)$$

Докажем теперь теорему Крэйга (1943): *квадратичные формы Q_1 и Q_2 независимы тогда и только тогда, когда*

$$AB = 0. \quad (15.55)$$

Достаточность условия (15.55) для выполнения (15.54) очевидна. Для доказательства его необходимости положим в (15.54) $k=l=2, m=n=\dots=0$. Поскольку матрицы A и B симметричны, имеем

$$0 = \text{tr}(A^2 B^2) = \text{tr}(B A^2 B) = \text{tr}\{(AB)'(AB)\}. \quad (15.56)$$

Элементы главной диагонали матрицы $(AB)'(AB)$ представляют собой сумму квадратов элементов соответствующего столбца матрицы AB , и все они, конечно, действительны. Таким образом, из (15.56) вытекает, что все элементы матрицы AB равны нулю, т. е. что выполняется (15.55). Замечательно, что уже одно соотношение (15.56) влечет (15.55): никакие другие соотношения типа (15.54) в доказательстве не использовались. Приведенное доказательство принадлежит Лэнкастеру (1954).

*) Следующее далее доказательство теоремы Крэйга основывается на этой равносильности, которую едва ли можно признать очевидной. Докажем здесь теорему Крэйга иным способом. Если формы Q_1 и Q_2 независимы, то в (15.51) коэффициенты при $\theta_1^k \theta_2^l$ ($k, l \geq 1$) равны нулю, и, в частности, равен нулю коэффициент при $\theta_1^2 \theta_2^2$: $4\{2 \text{tr}(A^2 B^2) + \text{tr}(AB)^2\} = 0$. Так как $\text{tr}(A^2 B^2) = \text{tr}\{(AB)'(AB)\}$ (см. (15.56)), то, обозначив $C = AB$, получим

$$\begin{aligned} 0 &= 2 \text{tr}(A^2 B^2) + \text{tr}(AB)^2 = 2 \text{tr} C C' + \text{tr} C^2 = \\ &= 2 \sum_{i,j} c_{ij}^2 + \sum_{i,j} c_{ij} c_{ji} = \sum_{i,j} c_{ij}^2 + \frac{1}{2} \sum_{i,j} (c_{ij} + c_{ji})^2. \end{aligned}$$

Следовательно, $AB = C = 0$. Доказательство достаточности (15.55) для независимости Q_1 и Q_2 не вызывает затруднений. (Прим. перев.)

15.14 Далее, так как матрицы A и B симметричны, их можно записать в виде

$$A = TT', \quad B = UU',$$

откуда

$$AB = TT'UU'.$$

Следовательно,

$$\text{tr}(AB) = \text{tr}(TT'UU') = \text{tr}(U'TT'U) = \text{tr}\{T'U\}'T'U\}.$$

Но след матрицы $(T'U)'T'U$ равен сумме квадратов всех элементов матрицы $T'U$, которые, вообще говоря, комплексны. Если же матрицы A и B неотрицательны, то матрица $T'U$ вещественна и из равенства

$$\text{tr}(AB) = \text{tr}\{T'U\}'T'U\} = 0 \quad (15.57)$$

вытекает, что

$$T'U = 0$$

и

$$AB = TT'UU' = 0.$$

Таким образом, для неотрицательных A и B соотношение (15.57) влечет (15.55) и, следовательно, независимость Q_1 и Q_2 . Этот результат был найден Матёрном (1949).

В частности, A и B неотрицательны, если Q_1 и Q_2 имеют распределение χ^2 . В этом случае эквивалентность (15.57) и (15.55) можно установить непосредственно, заметив, что в силу идемпотентности A и B

$$AB = A^2B^2,$$

так что (15.57) влечет (15.56), откуда вытекает (15.55) и независимость Q_1 и Q_2 .

Пример 15.4 Независимость среднего значения и дисперсии выборки из нормальной генеральной совокупности.

В примерах 11.3 и 11.7 мы видели, что для выборки объема n из одномерной генеральной совокупности с нормальным распределением (имеющим, без ограничения общности, среднее нуль и дисперсию единица) среднее $\bar{x}$ распределено нормально со средним значением, равным нулю, и дисперсией, равной $1/n$,

а сумма квадратов $\sum_{i=1}^n (x_i - \bar{x})^2$ имеет распределение χ^2 с $(n-1)$ степенями свободы. Ясно, что величина $n\bar{x}^2$ имеет χ^2 -распределение с одной степенью свободы. Матрицы, соответствующие квадратичным формам $n\bar{x}^2$ и $\sum_{i=1}^n (x_i - \bar{x})^2$, имеют вид

$$A = \frac{1}{n} U, \quad B = I - \frac{1}{n} U,$$

где U — матрица, все элементы которой равны единице. В силу идемпотентности A

$$\text{tr}(AB) = \text{tr}(A - A^2) = 0.$$

Отсюда и из 15.14 следует, что формы $n\bar{x}^2$ и $\sum_{i=1}^n (x_i - \bar{x})^2$ распределены независимо. Этот результат был уже получен нами в несколько иной форме в примере 11.3.

15.15 Из условия Крэйга (15.55) можно вывести необходимое и достаточное условие независимости квадратичной формы $Q = \mathbf{x}'A\mathbf{x}$ и линейной формы $L = \mathbf{b}'\mathbf{x}$. Формы Q и L независимы тогда и только тогда, когда независимы формы Q и $L^2 = \mathbf{x}'\mathbf{b}\mathbf{b}'\mathbf{x}$. Согласно (15.55) для независимости форм Q и L^2 необходимо и достаточно, чтобы имело место равенство

$$A\mathbf{b}\mathbf{b}' = 0. \quad (15.58)$$

Из (15.58) следует, что

$$A\mathbf{b}\mathbf{b}'\mathbf{b} = 0, \quad (15.59)$$

откуда, отбрасывая ненулевое скалярное произведение $\mathbf{b}'\mathbf{b}$, получаем

$$A\mathbf{b} = 0. \quad (15.60)$$

Таким образом, (15.60) необходимо для выполнения (15.58). Достаточность (15.60) для (15.58) очевидна.

Пример 15.5

Применим найденное в 15.15 условие для получения результата примера 15.4. В данном случае нас интересуют квадратичная форма

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \mathbf{x}' \left(I - \frac{1}{n} U \right) \mathbf{x}$$

и линейная форма

$$\bar{x} = \frac{1}{n} \mathbf{1}'\mathbf{x},$$

где $\mathbf{1}$ — вектор с единичными компонентами. Произведение соответствующих матриц равно

$$\left(I - \frac{1}{n} U \right) \cdot \frac{1}{n} \mathbf{1} = \frac{1}{n} \mathbf{1} - \frac{1}{n^2} n \mathbf{1} = 0,$$

так что (15.60) здесь выполнено, и формы $\sum_{i=1}^n (x_i - \bar{x})^2$ и $\bar{x}$ независимы.

Разложение суммы квадратов независимых нормально распределенных случайных величин

15.16 Теперь мы докажем замечательный результат Кокрэна (1934), усиленный впоследствии Джеймсом (1952). Пусть, как и выше, $\mathbf{x}$ — вектор, образованный p независимыми случайными величинами с нормированным нормальным распределением. Предположим, что сумма квадратов $\mathbf{x}'\mathbf{x}$ разложена на сумму k квадратичных форм $Q_i = \mathbf{x}'\mathbf{A}_i\mathbf{x}$ ранга r_i

$$\sum_{i=1}^k \mathbf{x}'\mathbf{A}_i\mathbf{x} = \mathbf{x}'\mathbf{I}\mathbf{x}. \quad (15.61)$$

Тогда любое из следующих трех условий влечет два остальных:

- (а) сумма рангов форм Q_i равна p ;
- (б) каждая из форм Q_i имеет χ^2 -распределение;
- (в) формы Q_i попарно независимы.

Используя метод, принадлежащий Лэнкастеру (1954), мы покажем, что каждое из условий (а) и (в) влечет (б) и что (б) влечет (а) и (в).

15.17 Возьмем какую-нибудь форму Q_i , например Q_1 . Производя диагонализующее $\mathbf{A}_1$ ортогональное преобразование (15.3), получим

$$\mathbf{y}'\mathbf{B}'\mathbf{A}_1\mathbf{B}\mathbf{y} + \mathbf{y}'\mathbf{B}'(\mathbf{I} - \mathbf{A}_1)\mathbf{B}\mathbf{y} = \mathbf{y}'\mathbf{I}\mathbf{y}, \quad (15.62)$$

поскольку сумма квадратов инвариантна относительно ортогонального преобразования. Так как матрицы $\mathbf{B}'\mathbf{A}_1\mathbf{B}$ и $\mathbf{I}$ диагональны, то диагональна и матрица $\mathbf{B}'(\mathbf{I} - \mathbf{A}_1)\mathbf{B}$. Более того, так как $(p - r_1)$ элементов главной диагонали матрицы $\mathbf{B}'\mathbf{A}_1\mathbf{B}$ суть нули, то соответствующие элементы матрицы $\mathbf{B}'(\mathbf{I} - \mathbf{A}_1)\mathbf{B}$ равны единице, и поскольку в силу (а) эта последняя матрица имеет ранг $p - r_1$, остальные элементы ее главной диагонали суть нули и соответствующие им элементы матрицы $\mathbf{B}'\mathbf{A}_1\mathbf{B}$ равны единице. Следовательно, матрица $\mathbf{A}_1$ идемпотентна, и из 15.11 следует, что Q_1 имеет χ^2 -распределение. Для других Q_i этот факт устанавливается аналогично. Таким образом, (а) влечет (б).

15.18 Пусть выполняется условие (б). Согласно (15.61),

$$\mathbf{I} = \sum_i \mathbf{A}_i. \quad (15.63)$$

Так как матрицы $\mathbf{A}_i$ идемпотентны, то после возведения (15.63) в квадрат получим

$$\mathbf{I} = \sum_i \mathbf{A}_i + \sum_{i \neq j} \mathbf{A}_i\mathbf{A}_j,$$

или

$$\sum_{i \neq j} A_i A_j = 0, \quad (15.64)$$

и, следовательно,

$$\sum_{i \neq j} \operatorname{tr} (A_i A_j) = 0. \quad (15.65)$$

В силу идемпотентности A_i и инвариантности характеристических чисел и следов относительно ортогональных преобразований, имеем

$$\operatorname{tr} (A_i A_j) \geq 0 \quad (i \neq j). \quad (15.66)$$

Из (15.65) и (15.66) вытекает, что

$$\operatorname{tr} (A_i A_j) = 0, \quad (15.67)$$

откуда согласно 15.14 $A_i A_j = 0$ (для всех $i \neq j$), и, следовательно, Q_i не зависит от Q_j . Таким образом, из (б) вытекает (в).

15.19 Если выполняется (б), то r_i характеристических корней матрицы A_i равны единице, а остальные — нулю. Беря след левой и правой частей (15.63), получаем

$$p = \sum_{i=1}^k r_i,$$

так что из (б) следует (а).

Если выполняется (в), то, возводя (15.63) в степени, получим

$$\sum_i A_i^s = I \quad (\text{для всех } s), \quad (15.68)$$

откуда

$$\operatorname{tr} \sum A_i^s = p \quad (\text{для всех } s). \quad (15.69)$$

Но (15.69) может иметь место лишь в том случае, когда каждый ненулевой характеристический корень любой матрицы A_i равен единице, т. е. когда все Q_i имеют χ^2 -распределение. Следовательно, (в) влечет (б).

Мы доказали, что из (а) вытекает (б) и обратно и что из (б) вытекает (в) и обратно. Таким образом, выполнение каждого из этих условий достаточно для выполнения двух остальных. См. также упражнение 15.16.

Из того, что сумма двух случайных величин, имеющих распределение χ^2 , тоже имеет распределение χ^2 , причем число степеней свободы распределения суммы равно сумме чисел степеней свободы распределений слагаемых, не следует, вообще говоря, независимость этих случайных величин. Пример 7.6 содержит соответствующий противоречащий пример, построен-

ный Джеймсом (1952). Таким образом, ограничение, состоящее в том, что рассматривались только квадратичные формы от независимых случайных величин с нормированным нормальным распределением, существенно.

Пример 15.6

Вернемся еще раз к среднему и дисперсии выборки из нормальной генеральной совокупности, обсуждавшимся в примере 15.4. Разложение суммы квадратов

$$\sum_{i=1}^n (x_i - \bar{x})^2 + n\bar{x}^2 = \sum_{i=1}^n x_i^2$$

удовлетворяет условиям (а), (б) и (в) теоремы Кокрэна. Выполнимость условия (б) была установлена в примерах 11.3 и 11.7, и, следовательно, (а) и (в) тоже выполняются. Легко также убедиться в том, что ранги квадратичных форм этого разложения равны $n-1$ и 1 (в соответствии с числами степеней свободы их χ^2 -распределений), так что (а) выполняется, и поэтому обе эти квадратичные формы имеют χ^2 -распределение и независимы.

15.20 Всюду, начиная с 15.11, мы предполагали, что $\mathbf{x}$ — вектор с независимыми нормированными нормально распределенными компонентами, однако изложенные результаты, по существу, не изменятся, если предположить, что $\mathbf{x}$ имеет многомерное невырожденное нормальное распределение со средним $\mathbf{0}$ и матрицей рассеяния $\mathbf{V}$. В самом деле, так как матрица $\mathbf{V}$ положительно определенная, то ее можно представить в виде

$$\mathbf{V} = \mathbf{T}\mathbf{T}', \quad (15.70)$$

где $\mathbf{T}$ — действительная матрица, и преобразование $\mathbf{x} = \mathbf{T}\mathbf{y}$ согласно (15.70) переводит показатель экспоненты, соответствующей плотности распределения $\mathbf{x}$, из $\mathbf{x}'\mathbf{V}\mathbf{x}$ в

$$\mathbf{y}'\mathbf{T}'\mathbf{V}^{-1}\mathbf{T}\mathbf{y} = \mathbf{y}'\mathbf{T}'(\mathbf{T}')^{-1}\mathbf{T}^{-1}\mathbf{T}\mathbf{y} = \mathbf{y}'\mathbf{y}.$$

Тем самым общий случай сводится к изученному случаю независимых компонент. В результате применения преобразования $\mathbf{T}$ квадратичные формы $\mathbf{x}'\mathbf{A}\mathbf{x}$ и $\mathbf{x}'\mathbf{B}\mathbf{x}$ переходят соответственно в $\mathbf{y}'\mathbf{T}'\mathbf{A}\mathbf{T}\mathbf{y}$ и $\mathbf{y}'\mathbf{T}'\mathbf{B}\mathbf{T}\mathbf{y}$, и условие Крэйга (15.55) принимает вид

$$\mathbf{0} = \mathbf{T}'\mathbf{A}\mathbf{T}\mathbf{T}'\mathbf{B}\mathbf{T} = \mathbf{T}'\mathbf{A}\mathbf{V}\mathbf{B}\mathbf{T}$$

или

$$\mathbf{A}\mathbf{V}\mathbf{B} = \mathbf{0}. \quad (15.71)$$

Это обобщение результата Крэйга принадлежит Эйткину (1950). Подобным образом, условие Лэнкастера (15.56).

переходит в

$$0 = \text{tr} \{ (T'AT)^2 (T'BT)^2 \} = \text{tr} \{ (AV)^2 (BV)^2 \}, \quad (15.72)$$

а условие Матérна (15.57) для неотрицательно определенных форм — в

$$0 = \text{tr} \{ (T'AT)(T'BT) \} = \text{tr} (AVBV). \quad (15.73)$$

15.21 Теорема Кокрэна (см. 15.16) имеет место в более общей ситуации, когда в (15.61) справа стоит любая квадратичная форма с распределением χ^2 . В самом деле, как мы знаем, такую форму можно преобразовать ортогональным преобразованием к сумме квадратов, и при этом ранги матриц, получающихся после преобразования стоящих в (15.61) слева вырожденных матриц, будут теми же самыми. Условия (а), (б) и (в) из 15.16 для преобразованных форм эквивалентны тем же условиям для непреобразованных форм, так что формулировка теоремы для более общего случая остается той же.

Если величины x_i не независимы, то теорема Кокрэна по-прежнему имеет место с очевидными модификациями (упражнение 15.17).

Характеризации нормального распределения

15.22 В примере 12.7 мы видели, что независимость среднего и дисперсии выборки из одномерной нормальной генеральной совокупности характеризует нормальное распределение: никакое другое распределение с конечными семинвариантами всех порядков не обладает этим свойством. Аналогичный факт имеет место и в многомерном случае: независимость вектора среднего значения и матрицы рассеяния выборки характеризует многомерное нормальное распределение среди всех многомерных распределений с конечными матрицами рассеяния. Лукачем (1942) было замечено, что для справедливости этого результата существование семинвариантов высших порядков не обязательно. Ниже мы приводим доказательство Лукача для одномерного случая и вкратце указываем, как оно переносится на многомерный.

15.23 Предположим, что распределение генеральной совокупности имеет конечные среднее μ и дисперсию σ^2 , и пусть $\varphi(t)$ — характеристическая функция (х.ф.) этого распределения. Х.ф. совместного распределения среднего $\bar{x}$ и дисперсии s^2 выборки объема n равна

$$\varphi_{12}(t_1, t_2) = \int \dots \int \exp(it_1 \bar{x} + it_2 s^2) dF_n. \quad (15.74)$$

Необходимым и достаточным условием независимости $\bar{x}$ и s^2 является возможность представить функцию (15.74) в виде

произведения х. ф. $\bar{x}$ и х. ф. s^2 :

$$\varphi_{12}(t_1, t_2) = \varphi_1(t_1) \varphi_2(t_2). \quad (15.75)$$

Дифференцируя (15.75) по t_2 , получаем

$$\left[\frac{\partial \varphi_{12}}{\partial t_2} \right]_{t_2=0} = \varphi_1(t_1) \left[\frac{\partial \varphi_2}{\partial t_2} \right]_{t_2=0}. \quad (15.76)$$

Заметим, что

$$\varphi_1(t_1) = \left\{ \varphi\left(\frac{t_1}{n}\right) \right\}^n \quad (15.77)$$

и

$$\left[\frac{\partial \varphi_2}{\partial t_2} \right]_{t_2=0} = iE(s^2) = i \frac{n-1}{n} \sigma^2. \quad (15.78)$$

Согласно (15.74)

$$\frac{\partial \varphi_{12}}{\partial t_2} = i \int \dots \int s^2 \exp(it_1 \bar{x} + it_2 s^2) dF_n, \quad (15.79)$$

а для s^2 имеем представление

$$s^2 = \frac{1}{n} \sum_i x_i^2 - \bar{x}^2 = \frac{1}{n^2} \left\{ (n-1) \sum_i x_i^2 - \sum_{i \neq j} x_i x_j \right\}. \quad (15.80)$$

Подставляя (15.80) в (15.79), находим

$$\begin{aligned} \left[\frac{\partial \varphi_{12}}{\partial t_2} \right]_{t_2=0} &= \\ &= \frac{i}{n^2} \int \dots \int \left\{ (n-1) \sum_i x_i^2 - \sum_{i \neq j} x_i x_j \right\} \exp\left(it_1 \sum_i x_i/n\right) dF_n = \\ &= \frac{i(n-1)}{n} \left\{ \int x^2 \exp(it_1 x/n) dF_1 \left[\varphi\left(\frac{t_1}{n}\right) \right]^{n-1} - \right. \\ &\quad \left. - \left[\int x \exp(it_1 x/n) dF_1 \right]^2 \left[\varphi\left(\frac{t_1}{n}\right) \right]^{n-2} \right\} = \\ &= \frac{i(n-1)}{n} \left[\varphi\left(\frac{t_1}{n}\right) \right]^{n-2} \left\{ -\varphi\left(\frac{t_1}{n}\right) \frac{\partial^2 \varphi\left(\frac{t_1}{n}\right)}{\partial \left(\frac{t_1}{n}\right)^2} + \left[\frac{\partial \varphi\left(\frac{t_1}{n}\right)}{\partial \left(\frac{t_1}{n}\right)} \right]^2 \right\}. \quad (15.81) \end{aligned}$$

Заменив $\frac{t_1}{n}$ на t и подставляя (15.81), (15.77) и (15.78) в (15.76), получаем

$$-\varphi(t) \varphi''(t) + \{\varphi'(t)\}^2 = \{\varphi(t)\}^2 \sigma^2. \quad (15.82)$$

Соотношение (15.82) можно переписать в виде

$$\frac{d}{dt} \left\{ \frac{d}{dt} \log \varphi(t) \right\} = -\sigma^2. \quad (15.83)$$

Интегрируя (15.83), находим

$$\varphi'(t)/\varphi(t) = -\sigma^2 t + c. \quad (15.84)$$

В силу начальных условий

$$\varphi(0) = 1, \quad \varphi'(0) = i\mu$$

должно быть

$$c = i\mu.$$

Подставляя это значение c в (15.84), получаем

$$\frac{d}{dt} \log \varphi(t) = i\mu - \sigma^2 t. \quad (15.85)$$

После интегрирования (15.85) имеем

$$\log \varphi(t) = i\mu t - \frac{1}{2} \sigma^2 t^2,$$

поскольку появляющаяся при интегрировании константа равна нулю, так как $\log \varphi(0) = 0$. Таким образом, окончательно

$$\varphi(t) = \exp\left(i\mu t - \frac{1}{2} \sigma^2 t^2\right), \quad (15.86)$$

т. е. φ — х. ф. нормального распределения.

Мы доказали тем самым, что независимость среднего и дисперсии выборки свойственна только нормальному распределению среди всех распределений с конечной дисперсией *).

15.24 Результат предыдущего пункта непосредственно обобщается на многомерный случай. Если вектор выборочного среднего значения и ковариация $s_{l,m}$ l -й и m -й компонент элементов выборки независимы, то, рассуждая так же, как при выводе (15.82), будем иметь

$$-\varphi''_{l,m}/\varphi + \varphi'_l \varphi'_m / \varphi^2 = \sigma_{lm}, \quad (15.87)$$

где φ — х. ф. исходного распределения, и индексы указывают на номера тех переменных, по которым производится дифференцирование. Так как (15.87) справедливо для всех $l, m = 1, 2, \dots, p$, то мы имеем систему дифференциальных уравнений с частными производными; ее матричная форма, аналогичная (15.83), имеет вид

$$\left(\frac{\partial}{\partial t_l} \frac{\partial}{\partial t_m} \log \varphi \right) = -V. \quad (15.88)$$

*) Кавата и Сакамото (Kawata, Sakamoto, 1949, *Jour. Math. Soc. Japan* 1, 111) показали, что даже требование конечности дисперсии можно ослабить.

Дважды интегрируя (15.88) и используя начальные условия, получаем

$$\varphi(t) = \exp\left(it'\mu - \frac{1}{2}t'Vt\right),$$

т. е. φ — х. ф. многомерного нормального распределения.

15.25 Известны и другие интересные и важные характеристики нормального распределения. Например, в главе 11 при обсуждении геометрического метода отыскания выборочных распределений мы видели, что плотность распределения независимых наблюдений из генеральной совокупности с нормированным одномерным нормальным распределением постоянна на сфере $\sum x^2 = \text{const}$, поскольку она зависит только от $\sum x^2$. Возникает вопрос, нет ли других распределений, обладающих этим свойством. Следующее рассуждение, принадлежащее Бартлетту (1934), показывает, что других таких распределений нет.

Пусть плотность распределения выборки L выражается формулой

$$L = \prod_{i=1}^n f(x_i) \propto g\left(\sum x^2\right). \quad (15.89)$$

Тогда все точки сферы с уравнением

$$\sum_i x_i^2 = \text{const}$$

являются точками условного экстремума функции L , а следовательно, и $\log L$. Поэтому

$$\frac{\partial \log f(x_i)}{\partial x_i} + \lambda x_i = 0 \quad (\text{для всех } i), \quad (15.90)$$

где $\lambda/2$ — множитель Лагранжа. Интегрируя, получаем

$$\log f(x_i) = -\frac{1}{2} \lambda x_i^2 + k_i,$$

т. е.

$$f(x_i) \propto \exp\left(-\frac{1}{2} \lambda x_i^2\right). \quad (15.91)$$

Таким образом, (15.89) характеризует нормальное распределение.

15.26 Это рассуждение непосредственно обобщается на многомерный случай. Вместо одного вектора длины $(\sum x^2)^{1/2}$ в p -мерном случае нужно рассматривать p векторов. Если плотность зависит только от квадратов длин этих векторов (от сумм квадратов) и углов между ними, то распределение должно быть

нормальным. В самом деле, пусть $\mathbf{x}_i$ есть $(p \times 1)$ -вектор *), являющийся i -м наблюдением случайного p -мерного вектора, и $\mathbf{X}'$ есть $(p \times n)$ -матрица, образованная этими наблюдениями. Суммы квадратов и произведений являются элементами матрицы $\mathbf{X}'\mathbf{X}$. Согласно предположению имеем

$$L = \prod_i f(\mathbf{x}_i) \propto g(\mathbf{X}'\mathbf{X}). \quad (15.92)$$

Аналогично одномерному случаю, дифференцируя логарифм и вводя $(p \times p)$ -матрицу $\mathbf{\Lambda}$, очевидным образом связанную с множителями Лагранжа, получим

$$\frac{\partial \log f(\mathbf{x}_i)}{\partial \mathbf{x}_i} + \mathbf{\Lambda} \mathbf{x}_i = 0 \quad (\text{для всех } i).$$

Следовательно,

$$f(\mathbf{x}_i) \propto \exp\left(-\frac{1}{2} \mathbf{x}_i' \mathbf{\Lambda} \mathbf{x}_i\right),$$

так что (15.92) характеризует многомерное нормальное распределение.

15.27 Мы приведем еще одну характеризацию нормального распределения, которая проясняет, почему нормальное распределение занимает исключительно важное место в статистике. В примерах 11.2 и 11.3 мы видели, что независимые случайные величины с одинаковым нормированным нормальным распределением ортогональным преобразованием переводятся в независимые величины с тем же распределением. В этом пункте мы докажем, что если $\mathbf{x}$ — вектор, имеющий (не обязательно одинаково распределенные) независимые нормированные компоненты с конечными семинвариантами, и если нетривиальное **) преобразование

$$\mathbf{x} = \mathbf{C} \mathbf{y} \quad (15.93)$$

приводит к вектору $\mathbf{y}$ с независимыми нормированными компонентами, то каждая из компонент x_i (а следовательно, и y_i) нормальна и преобразование $\mathbf{C}$ ортогонально. Этот результат впервые был получен Лэнкастером (1954).

Ортогональность $\mathbf{C}$ получается сразу, поскольку в силу нормированности и независимости компонент

$$\sum_j c_{ij}^2 = 1 \quad (15.94)$$

*) Эти новые векторы, конечно, отличны по характеру от только что рассматривавшихся $(n \times 1)$ -векторов (наблюдений) одномерных случайных величин.

**) Имеется в виду, что преобразование $\mathbf{C}$ не является перестановкой. (Прим. перев.)

и

$$\sum_j c_{ij}c_{kj} = 0 \quad (i \neq k), \quad (15.95)$$

а (15.94) вместе с (15.95) и есть условие ортогональности. Следовательно, $C' = C^{-1}$ и

$$y = C'x. \quad (15.96)$$

Так как величины y_i независимы, то согласно (15.93)

$$\kappa_{si} = \sum_k c_{ik}^s \lambda_{sk} \quad (\text{для всех } s), \quad (15.97)$$

где κ_{si} есть s -й семиинвариант величины x_i и λ_{sk} есть s -й семиинвариант величины y_k . Обратно, согласно (15.96)

$$\lambda_{si} = \sum_j (c'_{ij})^s \kappa_{sj} \quad (\text{для всех } s). \quad (15.98)$$

Комбинируя (15.97) и (15.98), получаем

$$\kappa_{si} = \sum_j \kappa_{sj} \sum_k c_{ik}^s c_{jk}^s \quad (\text{для всех } s). \quad (15.99)$$

При $s=1, 2$ соотношение (15.99) сводится соответственно к $\kappa_{1i}=0, \kappa_{2i}=1$, как это и должно быть для нормированных случайных величин. Далее,

$$\sum_j \left| \sum_k c_{ik}^s c_{jk}^s \right| \leq \sum_j \sum_k |c_{ik}^s| |c_{jk}^s|. \quad (15.100)$$

В силу (15.94) из (15.100) при $s \geq 3$ вытекает, что

$$\sum_j \left| \sum_k c_{ik}^s c_{jk}^s \right| < 1. \quad (15.101)$$

Из (15.99) следует, очевидно, неравенство

$$|\kappa_{si}| \leq \sum_j |\kappa_{sj}| \left| \sum_k c_{ik}^s c_{jk}^s \right|. \quad (15.102)$$

В правой части (15.102) стоит взвешенная сумма модулей $|\kappa_{sj}|$, причем согласно (15.101) сумма весов меньше единицы. Если при фиксированном s взять то значение i , при котором $|\kappa_{si}|$ максимален, то из (15.102) будет следовать, что $|\kappa_{si}|=0$, так что все $\kappa_{si}=0$. Это верно при всех $s \geq 3$, т. е.

$$\kappa_{si}=0 \quad (\text{при всех } s \geq 3),$$

и поэтому все x_i нормальны. Из нормальности x_i вытекает и нормальность y_i .

Лукач (Lukacs, 1956, *Proc. Third Berkeley Sympos*, 2, 195) опубликовал общий обзор методов и результатов, касающихся характеристик нормального распределения. Лэнкастер (Lancaster, 1960, *J. Austral. Math. Soc.* 1, 368) дал обзор характери-

заций нормального распределения свойствами независимости; им показано, в частности, что эти свойства обеспечивают существование всех моментов, так что предполагать их наличие излишне *).

УПРАЖНЕНИЯ

15.1 Показать, что условное распределение, получающееся из многомерного нормального распределения при фиксировании некоторых координат, тоже является многомерным нормальным распределением и что вектор среднего значения условного распределения является функцией от фиксированных координат, а матрица, обратная к матрице его вторых моментов, совпадает с отвечающей ей частью соответствующей матрицы исходного распределения.

15.2 Трехмерное (не обязательно нормальное) распределение имеет коэффициенты корреляции ρ_{12} , ρ_{13} , ρ_{23} . Показать, что $1 + 2\rho_{12}\rho_{13}\rho_{23} \geq \rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2$.

15.3 Показать, что все коэффициенты корреляции p -мерного распределения могут быть равны одному и тому же числу ρ тогда и только тогда, когда $\rho \geq -1/(p-1)$.

15.4 Показать, что плотность многомерного нормального распределения удовлетворяет системе уравнений

$$\frac{\partial f}{\partial \rho_{ij}} = \frac{\partial^2 f}{\partial x_i \partial x_j},$$

где ρ_{ij} — коэффициент корреляции величин x_i и x_j .

15.5 Рассмотреть двумерное нормальное распределение с независимыми компонентами и показать, что

$$\int_{-x}^x (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} t^2\right) dt \leq \{1 - \exp(-2x^2/\pi)\}^{\frac{1}{2}}.$$

(Ошибка от замены неравенства знаком равенства всегда меньше 0,75%).

(Уильямс, 1946.)

15.6 Рассматривая площадь подходящего сферического треугольника, вывести формулу (15.32) для вероятности того, что все координаты в случае трехмерного нормального распределения с нулевым вектором средних и единичными дисперсиями положительны.

15.7 Производятся выстрелы по круговой цели радиуса единица, поставленной вертикально. Горизонтальные и вертикальные отклонения попаданий от центра цели имеют двумерное нормальное распределение, их средние равны нулю, дисперсии равны v и коэффициент корреляции равен ρ . Показать, что вероятность поражения цели равна

$$\frac{(1-\rho^2)^{\frac{1}{2}}}{\pi} \int_{1-\rho}^{1+\rho} \left[1 - \exp\left\{-\frac{u}{2v(1-\rho^2)}\right\} \right] \frac{du}{u\{\rho^2 - (1-u)^2\}^{\frac{1}{2}}}.$$

*) О характеристизации см. также: Ю. В. Линник (1960), *Разложения вероятностных законов*, Л., Изд-во ЛГУ; А. А. Зингер, Ю. В. Линник (1964), *Теория вероят. и ее примен.* IX, 692; Ю. В. Прохоров (1965), *Теория вероят. и ее примен.* X, 479. (Прим. ред.)

15.8 Случайные величины x и y имеют невырожденное двумерное нормальное распределение. Показать, что величины $x^2 + 2axy + y^2$ и $x^2 + 2bxy + y^2$ могут быть независимыми только тогда, когда a и b равны по абсолютной величине единице и имеют противоположные знаки.

15.9 Рассматривается p -мерное нормальное распределение, у которого все коэффициенты корреляции равны $1/2$. Используя вспомогательную последовательность $(p+1)$ независимых нормированных нормально распределенных величин, показать, что относительно рассматриваемого распределения все компоненты превосходят одновременно их математическое ожидание с вероятностью $1/(p+1)$.

(Фостер и Стюарт, 1954.)

15.10 Если в предыдущем упражнении вместо равенства всех коэффициентов корреляции $1/2$ потребовать, чтобы матрица рассеяния имела вид

$$V = \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & \dots & . & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \dots & . & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & \dots & . & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & \dots & . & 0 \\ . & . & . & . & \dots & . & . \\ . & . & . & . & \dots & . & . \\ 0 & . & . & . & \dots & 1 & -\frac{1}{2} \\ 0 & . & . & . & \dots & -\frac{1}{2} & 1 \end{pmatrix},$$

то вероятность, о которой там шла речь, будет равна $1/(p+1)!$.

15.11 Показать, что в случае многомерного нормального распределения с матрицей V^{-1} , равной матрице V предыдущего упражнения, все компоненты не превосходят их математических ожиданий с вероятностью $1/(p+1)$.

(Энис и Ллойд, 1953.)

15.12 Рассматривая разности $y_i = x_i - b_i x_0$, где x_i — независимые нормированные нормально распределенные случайные величины с плотностью f , показать, что при многомерном нормальном распределении с коэффициентами корреляции $\rho_{ij} = a_i a_j (i \neq j)$ все компоненты положительны с вероятностью, равной

$$\int_{-\infty}^{\infty} \left\{ \prod_i \int_{b_i x}^{\infty} f(t) dt \right\} f(x) dx,$$

где $b_i = a_i / (1 - a_i^2)^{1/2}$.

15.13 Если $\mathbf{x}$ имеет многомерное нормальное распределение со средним $\mathbf{0}$ и матрицей рассеяния V , то квадратичная и линейная формы $\mathbf{x}' A \mathbf{x}$ и $\mathbf{b}' \mathbf{x}$ независимы тогда и только тогда, когда $A V \mathbf{b} = \mathbf{0}$.

(Эйткин, 1950.)

15.14 Если в предыдущем упражнении предположить, что матрица V невырождена и имеет ранг p , то квадратичная форма $x'Ax$ имеет распределение χ^2 с s степенями свободы тогда и только тогда, когда матрица $(V^{-1} - A)$ имеет ранг $p - s$.

15.15 Если квадратичные формы $x'Ax$ и $x'Bx$ имеют распределение χ^2 с m и n ($n \geq m$) степенями свободы соответственно, то для того, чтобы одна из них была частью другой, необходимо и достаточно, чтобы их коэффициент корреляции был равен $(m/n)^{1/2}$.

(Лэнкастер, 1954.)

15.16 Модифицируя рассуждения 15.18, показать, что если квадратичная форма Q_1 имеет распределение χ^2 с r_1 степенями свободы, а квадратичная форма $(x'Ix - Q_1)$ неотрицательно определена, то они независимы и $(x'Ix - Q_1)$ имеет χ^2 -распределение с $(p - r_1)$ степенями свободы.

(Хогг и Крэйг (Hogg, Craig), 1958, *Ann. Math. Stat.* 29, 608.)

15.17 Пусть вектор x имеет многомерное нормальное распределение со средним 0 и матрицей рассеяния V и пусть $\sum_{i=1}^k x' A_i x = x' A x$, причем матрица AV идемпотентна. Показать, что любое из условий (а), (б), (в) из 15.16 и здесь влечет два остальных. В данном случае (б) равносильно также идемпотентности всех матриц $A_i V$, а (в) — выполнению соотношения $A_i V A_j = 0$ для всех $i \neq j$.

15.18 Случайные величины имеют двумерное нормальное распределение с нулевыми средними, единичными дисперсиями и коэффициентом корреляции ρ . Показать, что условное распределение x и y при условии $ax + by = 0$ есть одномерное нормальное распределение со средним нуль и дисперсией

$$(1 - \rho^2)(a^2 + b^2)/(a^2 + b^2 + 2\rho ab).$$

15.19 Случайные величины $x_1, x_2, \dots, x_n$ имеют многомерное нормальное распределение. Ковариацию величин x_i и x_j обозначим c_{ij} . Показать, что ковариация x_i и x_j при условии $l_1 x_1 + l_2 x_2 + \dots + l_n x_n = \text{const}$ равна

$$c_{ij} - \frac{\left(\sum_k l_k c_{ik}\right) \left(\sum_k l_k c_{jk}\right)}{\sum_k \sum_m l_k l_m c_{km}}.$$

15.20 Пусть случайные величины x и y имеют двумерное нормальное распределение с нулевыми средними. Образум новые случайные величины x_1 и y_1 , положив $x_1 = x$, $y_1 = y$, когда x и y обе положительны или обе отрицательны, и $x_1 = -x$, $y_1 = y$, когда x и y имеют разный знак. Показать, что каждая из величин x_1, y_1 распределена нормально, но их совместное распределение отлично от нормального. Обобщая этот результат, показать, что среди n случайных величин любые $p < n$ могут иметь совместное нормальное распределение, в то время как совместное распределение всех n величин отлично от нормального.

15.21 В примере 15.3 было установлено, что сумма $\sum_i \frac{(x_i - np_i)^2}{np_i}$

имеет распределение χ^2 с $(k - 1)$ степенями свободы. Получить этот результат, используя упражнение 15.19. (Рассмотреть условное распределение k

независимых распределенных по Пуассону случайных величин при условии $\sum_{i=1}^k x_i = n$.

15.22 Случайные величины $x_1, \dots, x_k$ независимы и имеют каждая (не обязательно одинаковое) гамма-распределение. Пользуясь характеристическими функциями, показать, что любая безразмерная функция от этих величин $h(x_1, \dots, x_k)$ не зависит от $S = \sum_{i=1}^k x_i$. В частности, это верно для

$$h(x_1, \dots, x_k) = \sum_{i=1}^n a_i x_i \left/ \sum_{i=1}^n x_i = R. \right.$$

(Питмэн (Pitman), (1937), *Proc. Camb. Phil. Soc.* 33, 212. Лаха (Laha), (1954), *Ann. Math. Statist.* 25, 784, доказал обратное утверждение: если независимые одинаково распределенные случайные величины x_i имеют конечную дисперсию, то из независимости R и S следует, что x_i имеют гамма-распределение.)

РАСПРЕДЕЛЕНИЯ, СВЯЗАННЫЕ С НОРМАЛЬНЫМ

16.1 В предшествующих главах мы не раз встречались с распределениями, связанными с нормальным. Важность этих распределений для статистической теории обусловлена тем, что они описывают выборочные распределения различных статистик с результатов наблюдений, взятых из нормальной совокупности. То специальное положение, которое в силу Центральной предельной теоремы в той или иной ее форме занимает нормальное распределение, объясняет и важную роль распределений, связанных с нормальными. Всякий раз, когда некоторая статистика (при соответствующих условиях регулярности) имеет асимптотически нормальное выборочное распределение, существуют статистики, асимптотическое распределение которых совпадает с некоторым распределением, связанным с нормальными. Мы уже встречались с этим обстоятельством в примере 15.3, когда речь шла о критерии согласия χ^2 .

В этой главе рассматриваются три важных распределения, связанных с одномерным нормальным распределением. Это χ^2 -распределение, t -распределение Стьюдента и F -распределение Фишера. Каждое из них детально изучается с подробным описанием специфических их свойств. В конце главы обсуждаются распределения, возникающие в связи с двумерным нормальным распределением.

Распределение χ^2

16.2 Мы уже видели (пример 11.6), что сумма z квадратов n независимых нормированных нормальных случайных величин имеет распределение

$$dF = \frac{1}{2^{\frac{1}{2}n} \Gamma\left(\frac{1}{2}n\right)} \exp\left(-\frac{1}{2}z\right) z^{\frac{1}{2}n-1} dz, \quad 0 \leq z < \infty. \quad (16.1)$$

Если при этом мы предположим, что эти случайные величины подчинены p линейным соотношениям ($p < n$), то распределение величины z будет даваться той же самой формулой (16.1), где вместо n надо подставить $n - p$. Далее, мы также видели (пример 11.7), что если сумма квадратов берется относительно среднего выборки, а не истинного среднего, то это накладывает одно линейное условие, т. е. распределение z в этом случае таково, что вместо n в (16.1) надо подставить $n - 1$.

Распределение (16.1) называется χ^2 -распределением с n степенями свободы. Если мы положим $z = \chi^2$ и $n = \nu$, то получим

$$dF = \frac{1}{\Gamma\left(\frac{1}{2}\nu\right)} \exp\left(-\frac{1}{2}\chi^2\right) \left(\frac{1}{2}\chi^2\right)^{\frac{1}{2}\nu-1} d\left(\frac{1}{2}\chi^2\right), \quad \nu > 0, \quad 0 \leq \chi^2 \leq \infty, \quad (16.2)$$

откуда сразу видно, что $\frac{1}{2}\chi^2$ имеет Γ -распределение с параметром $\frac{1}{2}\nu$. Распределение (16.2) является распределением Пирсона III типа. До сих пор мы рассматривали распределения типа (16.2) с целым положительным ν , теперь же мы можем воспользоваться тем фактом, что формула (16.2) задает распределение и при любых положительных ν , поскольку интеграл от плотности в (16.2) сходится.

Дифференцируя (16.2), мы найдем, что для $\nu > 2$, χ^2 -распределение имеет моду в точке $\nu - 2$. Для $\nu = 2$ плотность распределения J-образна с максимумом в нуле. При $0 < \nu < 2$ плотность также J-образна с бесконечной ординатой в нуле.

Свойства χ^2 -распределения

16.3 Характеристическую функцию χ^2 можно сразу получить из (16.1) (см. пример 3.6):

$$\varphi(t) = \frac{1}{(1 - 2it)^{\frac{1}{2}\nu}}, \quad (16.3)$$

откуда находим семиинварианты

$$\kappa_r = \nu 2^{r-1} (r-1)! \quad (16.4)$$

и центральные моменты

$$\begin{aligned} \mu_2 &= 2\nu, \\ \mu_3 &= 8\nu, \\ \mu_4 &= 48\nu + 12\nu^2, \\ \mu_5 &= 32\nu(5\nu + 12), \\ \mu_6 &= 40\nu(3\nu^2 + 52\nu + 96). \end{aligned} \quad (16.5)$$

Поскольку семинварианты κ_r линейны относительно v , то μ_r (который может содержать μ_2 только в степени, не превышающей $\left[\frac{r}{2}\right]$) по v должен иметь степень $\left[\frac{r}{2}\right]$, т. е. $\frac{r}{2}$, если r четно, и $\frac{1}{2}(r-1)$, если r нечетно.

Когда v стремится к бесконечности, χ^2 -распределение стремится к нормальному, поскольку после нормировки имеем

$$\varphi(t) = e^{-\frac{vit}{\sqrt{2v}}} \left(1 - \frac{2it}{\sqrt{2v}}\right)^{-\frac{1}{2}v}$$

и

$$\log \varphi(t) = \frac{-vit}{\sqrt{2v}} - \frac{v}{2} \left\{ \frac{-2it}{\sqrt{2v}} - \frac{1}{2} \left(\frac{2it}{\sqrt{2v}} \right)^2 + o\left(\frac{1}{v}\right) \right\} \rightarrow -\frac{1}{2} t^2.$$

Это стремление, однако, очень медленное и существуют лучшие аппроксимации, с которыми мы скоро познакомимся.

16.4 Функция распределения (16.2) есть неполная Γ -функция. Мы имеем

$$F(\xi) = \int_0^{\xi} \frac{1}{2^{\frac{1}{2}v} \Gamma\left(\frac{1}{2}v\right)} \exp\left(-\frac{1}{2}z\right) z^{\frac{1}{2}v-1} dz = \frac{\Gamma\left(\frac{1}{2}\xi\right) \Gamma\left(\frac{1}{2}v\right)}{\Gamma\left(\frac{1}{2}v\right)}$$

или в обозначениях таблиц Пирсона

$$= I\left(\frac{\frac{1}{2}\xi}{\sqrt{\frac{1}{2}v}}, \frac{1}{2}v-1\right).$$

Существует много различных таблиц распределений χ^2 . Укажем на те из них, которые наиболее употребительны и просты в работе*):

(а) Пирсон и Хартли (1950) дали таблицы функции распределения χ^2 для $v=1(1)20(2)70$ и $\chi^2=0,001(0,001)0,01(0,01)0,1(0,1)2,0(0,2)10(0,5)20(1)40(2)70$. Эти таблицы полностью воспроизведены в *Biometrika Tables*.

(б) Томпсон (1941) создала (обратные) таблицы процентных точек распределения (т. е. таблицы таких величин χ^2 , которые соответствуют значениям функции распределения, равным $1-P$) для $P=0,995, 0,990, 0,975, 0,950, 0,900, 0,750, 0,500, 0,250, 0,100, 0,050, 0,025, 0,010, 0,005$ и $v=1(1)30(10)100$. Эти таблицы при-

* Этим требованиям удовлетворяют также таблицы χ^2 -распределения в книгах: *Таблицы математической статистики и Сборник статистических таблиц.* (Прим. ред.)

ведены в *Biometrika Tables* с дополнительными значениями для $P=0,001$.

(в) Фишер и Иэйтс (1953) дали обратные таблицы для $P=0,99, 0,98, 0,95, 0,90, 0,10, 0,05, 0,02, 0,01, 0,001$ и $v=1(1)30$. Часть этих таблиц приведена в приложении (таблица 3).

(г) Таблицы, созданные Юлом, приведены в приложении (таблица 4а). В них даются значения P для $v=1, \chi^2=0(0,01)1(0,1)10$.

16.5 Наиболее употребительны следующие две аппроксимации распределения χ^2 , каждая из которых состоит в подборе подходящей функции от χ^2 такой, чтобы ее распределение было ближе к нормальному, чем само χ^2 -распределение.

(а) Фишеровская аппроксимация основана на том, что случайная величина $\sqrt{2\chi^2}$ распределена приблизительно нормально со средним $\sqrt{2v}-1$ и единичной дисперсией.

(б) Аппроксимация Уилсона и Хилфerti (1931) основана на том, что $(\chi^2/v)^{1/2}$ распределена приблизительно нормально со средним $1 - 2/(9v)$ и дисперсией $2/(9v)$.

Вторая из этих аппроксимаций более точна, но зато требует больше вычислительной работы.

16.6 Сравнение скорости приближения распределений статистик χ^2 и $\sqrt{2\chi^2}$ к нормальному можно провести следующим образом.

Для χ^2 из (16.5) имеем

$$\left. \begin{aligned} \gamma_1 &= \sqrt{\beta_1} = (8/v)^{1/2}, \\ \gamma_2 &= \beta_2 - 3 = \frac{12}{v}. \end{aligned} \right\} \quad (16.6)$$

Для моментов χ получаем

$$\mu'_r = \frac{1}{2^{\frac{1}{2}(v-2)} \Gamma\left(\frac{1}{2}v\right)} \int_0^\infty e^{-\frac{1}{2}\chi^2} \chi^{v+r-1} d\chi = \frac{2^{\frac{1}{2}r} \Gamma\left\{\frac{1}{2}(v+r)\right\}}{\Gamma\left(\frac{1}{2}v\right)}. \quad (16.7)$$

Тем самым

$$\mu'_1 = \sqrt{2} \frac{\Gamma\left\{\frac{1}{2}(v+1)\right\}}{\Gamma\left(\frac{1}{2}v\right)}.$$

Используя формулу Стирлинга

$$\log \Gamma(x+1) \sim \frac{1}{2} \log(2\pi) + \left(x + \frac{1}{2}\right) \log x - x + \frac{1}{12x} - \frac{1}{360x^3} + o\left(\frac{1}{x^3}\right),$$

после соответствующих преобразований мы найдем, что

$$\mu'_1 = \sqrt{v} \left(1 - \frac{1}{4v} + \frac{1}{32v^2} + \frac{5}{128v^3} + \dots \right),$$

откуда

$$(\mu'_1)^2 = v \left(1 - \frac{1}{2v} + \frac{1}{8v^2} + \frac{1}{16v^3} + \dots \right).$$

Аналогично

$$\mu'_2 = v,$$

$$\mu'_3 = (v+1)\mu'_1,$$

$$\mu'_4 = (v+2)v,$$

откуда для центральных моментов получаем формулы:

$$\mu_2 = \frac{1}{2} - \frac{1}{8v} + o\left(\frac{1}{v}\right), \quad \mu_3 = \frac{1}{4\sqrt{v}} + o\left(\frac{1}{\sqrt{v}}\right),$$

$$\mu_4 = \frac{3}{4} - \frac{3}{8v} + o\left(\frac{1}{v}\right).$$

Следовательно, моменты $\sqrt{2\chi^2}$ имеют вид

$$\mu_2 = 1 - \frac{1}{4v} + o\left(\frac{1}{v}\right),$$

$$\gamma_1 = \frac{1}{\sqrt{2v}} + o\left(\frac{1}{\sqrt{v}}\right), \quad \gamma_2 = \frac{3}{4v^2} + o\left(\frac{1}{v^2}\right). \quad (16.8)$$

Сравнение (16.8) с (16.6) показывает, что распределение $\sqrt{2\chi^2}$ стремится к нормальному со значительно большей скоростью, чем для χ^2 . Кроме того, выражение для μ'_1 случайной величины χ равно $\sqrt{v - \frac{1}{2}}$ с точностью до величин порядка $v^{-3/2}$.

Следовательно, $\sqrt{2\chi^2}$ распределена относительно среднего $\sqrt{2v-1}$ (с точностью до величин порядка $v^{-3/2}$) с дисперсией, которая с точностью до v^{-1} равна единице.

16.7 Для получения аппроксимации Уилсона — Хилфerti рассмотрим распределение χ^2 относительно его среднего значения v . Найдем распределение $y = \left(\frac{\chi^2}{v}\right)^h$, где h — пока неопределенная константа. Запишем $\xi = \chi^2 - v$. Тогда

$$y = \left(1 + \frac{\xi}{v}\right)^h = 1 + \frac{h\xi}{v} + \frac{h(h-1)}{2!} \frac{\xi^2}{v^2} + \dots \quad (16.9)$$

Беря математическое ожидание и используя результаты (16.5), после ряда преобразований получим

$$\begin{aligned} \mu'_1(y) = 1 + \frac{h(h-1)}{2!} \frac{\mu_2(\chi^2)}{v^2} + \dots = 1 + \frac{h(h-1)}{1} \frac{1}{v} + \\ + \frac{h(h-1)(h-2)(3h-1)}{6v^2} + \frac{h^2(h-1)^2(h-2)(h-3)}{6v^3} + o(v^{-4}). \end{aligned} \quad (16.10)$$

Если в (16.9) мы положим rh вместо h и найдем среднее значение для y^r , то получим

$$\mu'_r(y) = 1 + \frac{rh(rh-1)}{1} \frac{1}{v} + \dots \quad (16.11)$$

Выберем h так, чтобы третий член в (16.10) обращался в нуль. Для этого надо взять $h = \frac{1}{3}$. Тогда из (16.10) и (16.11) находим

$$\begin{aligned} \mu'_1(y) &= 1 - \frac{2}{9v} + \frac{80}{3^7v^3} + o(v^{-4}), \\ \mu'_2(y) &= 1 - \frac{2}{9v} + \frac{4}{3^4v^2} + \frac{56}{3^7v^3} + o(v^{-4}), \\ \mu'_3(y) &= 1, \\ \mu'_4(y) &= 1 + \frac{4}{9v} - \frac{4}{3^3v^2} + \frac{80}{3^7v^3} + o(v^{-4}), \end{aligned}$$

откуда

$$\left. \begin{aligned} \mu_2(y) &= \frac{2}{9v} - \frac{104}{3^7v^3} + o(v^{-4}), \\ \mu_3(y) &= \frac{32}{3^6v^3} + o(v^{-4}), \\ \mu_4(y) &= \frac{4}{3^3v^2} - \frac{16}{3^6v^3} + o(v^{-4}), \end{aligned} \right\} \quad (16.12)$$

и

$$\gamma_1 = \frac{2^{7/2}}{3^3v^{3/2}} + o(v^{-3/2}), \quad \gamma_2 = -\frac{4}{9v} + o(v^{-1}). \quad (16.13)$$

Сравнение коэффициентов асимметрии в (16.13), (16.6) и (16.8) показывает, что распределение $(\chi^2/v)^{1/3}$ стремится к симметричному более быстро, чем χ^2 и $\sqrt{2\chi^2}$. С точностью до членов порядка v^{-2} дисперсия согласно (16.12) равна $2/9v$ и среднее равно $1 - \frac{2}{9v}$. Этот результат можно переформулировать, сказав, что случайная величина

$$\left\{ \left(\frac{\chi^2}{v} \right)^{1/3} + \frac{2}{9v} - 1 \right\} \left(\frac{9v}{2} \right)^{1/2} \quad (16.14)$$

распределена приближенно нормально с нулевым средним и единичной дисперсией.

16.8 Следующая таблица, заимствованная у Гарвуда (1936), дает представление об этих двух аппроксимациях на конкретных числовых примерах. В каждом случае вычислялись значения

$$\frac{1}{2}\chi^2, \text{ соответствующие значениям } P = \int_{\chi^2}^{\infty} dF = 0,01, 0,05, 0,95$$

и 0,99. В таблице приняты следующие обозначения: m_T — точное значение $\frac{1}{2}\chi^2$ при заданных P и v ; m_F — аппроксимация Фишера; m_W — аппроксимация Уилсона и Хилферти.

Таблица 16.1
Сравнение аппроксимаций интеграла χ^2
(Гарвуд (1936))

	v	m_T	m_F	$m_T - m_F$	m_W	$m_T - m_W$
$P = 0,99$	40	11,082	10,764	0,318	11,070	0,012
	60	18,742	18,414	0,328	18,732	0,010
	80	26,770	26,436	0,334	26,761	0,009
	100	35,032	34,094	0,338	35,025	0,007
$P = 0,95$	40	13,255	13,116	0,139	13,254	0,001
	60	21,594	21,455	0,139	21,594	0,000
	80	30,196	30,056	0,140	30,196	0,000
	100	38,965	38,825	0,140	38,965	0,000
$P = 0,05$	42	29,062	28,919	0,143	29,060	0,002
	62	40,691	40,598	0,143	40,689	0,002
	82	52,069	51,926	0,143	52,068	0,001
	102	63,287	63,144	0,143	63,286	0,001
$P = 0,01$	42	33,103	32,700	0,403	33,113	—0,010
	62	45,401	45,003	0,398	45,409	—0,008
	82	57,347	56,953	0,394	57,355	—0,008
	102	69,067	68,676	0,391	69,074	—0,007

Очевидно, что аппроксимация m_W очень хороша, тогда как m_F только несколько хуже.

Слущкий (1950) табулировал*) распределение χ^2 с пятью десятичными знаками, используя нормальную аппроксимацию для $\{\sqrt{2\chi^2} - \sqrt{2v}\}$, где $\sqrt{2/v}$ изменяется от 0 до 0,25, т. е. для значений $32 \leq v \leq \infty$.

16.9 Третий метод аппроксимации может быть получен на основании методов 6.25 и 6.26 и пример 6.4 был, по существу, посвящен аппроксимации распределения χ^2 . Взяв в уравнении (6.48)

$$l_1 = l_2 = 0, \quad l_3 = k_3 = \sqrt{\frac{8}{v}}, \quad l_4 = k_4 = \frac{12}{v}, \quad l_5 = \frac{48\sqrt{2}}{v^{3/2}}, \quad l_6 = \frac{480}{v^2},$$

с помощью (6.54) или (6.56) убеждаемся, что рассмотренное преобразование случайной величины χ^2 имеет распределение, близкое к нормальному.

Стьюдентское t -распределение

16.10 В примере 11.8 мы видели, что в нормальном случае отношение разности между средним выборки и истинным средним к оценке ее стандартной ошибки

$$t = (\bar{x} - \mu) \sqrt{n(n-1)} / \sqrt{\sum (x - \bar{x})^2} = \frac{\bar{x} - \mu}{\sqrt{\frac{s^2}{n}}},$$

где $s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$ подчиняется распределению, элемент которого выражается формулой

$$dF = dt / \left\{ v^{\frac{1}{2}} B\left(\frac{1}{2}, \frac{1}{2}v\right) \left(1 + \frac{t^2}{v}\right)^{\frac{1}{2}(v+1)} \right\}, \quad v > 0, \quad -\infty \leq t \leq \infty. \quad (16.15)$$

Как и раньше, мы используем v для обозначения числа степеней свободы, в данном случае равного $n-1$. Происхождение этого термина нуждается в некотором объяснении.

Мы видели (пример 11.7), что дисперсия нормальной выборки распределена подобно сумме $(n-1)$ квадратов независимых случайных величин. И вообще, если имеется k линейных соотношений, связывающих исходные случайные величины, то сумма их квадратов распределена как сумма $(n-k)$ квадратов

*) Принцип построения таблиц χ^2 -распределения, предложенный Е. Е. Слущким, послужил основой при создании соответствующего раздела в сборнике *Таблицы математической статистики*. Этот принцип позволил табулировать χ^2 -распределение во всей его естественной области определения. (Прим. ред.)

независимых нормированных нормальных величин. Каждое линейное соотношение, связывающее величины $x_1, \dots, x_n$, снижает размерность области значений вектора $(x_1, \dots, x_n)$ на единицу. Естественно поэтому говорить, что случайная величина типа χ^2 имеет ν степеней свободы, если ее распределение совпадает с распределением суммы ν квадратов независимых нормированных случайных величин.

Статистика t зависит от параметра ν весьма простым и удобным для табулирования образом. Кроме того, ν является числом степеней свободы статистики s^2 , входящей в знаменатель t . Поэтому, распространяя предыдущую терминологию, ν можно называть числом степеней свободы статистики t . Распределение (16.15) является распределением Пирсона VII типа. Оно, очевидно, симметрично относительно начала, унимодально и простирается до бесконечности в обе стороны. Обычно его называют t -распределением Стьюдента. (Стьюдент — псевдоним английского статистика У. С. Госсета, первооткрывателя этого распределения.)

Из того факта, что выборочная стандартная ошибка при увеличении объема выборки сходится по вероятности к истинной ошибке, ясно, что распределение (16.15) должно стремиться с ростом ν к нормированному нормальному распределению, так как среднее нормальной выборки распределено нормально (пример 11.12). Ср. с примером 4.8.

Свойства t -распределения Стьюдента

16.11 Характеристическая функция этого распределения изучалась в примере 3.13, где в соответствии с новыми обозначениями надо положить $m = \frac{1}{2}(\nu + 1)$. Моменты μ_r этого распределения существуют только для $r < \nu$, причем нечетные моменты равны нулю, а четные моменты выражаются формулой (пример 3.3)

$$\mu_{2r} = \nu^r \frac{\Gamma\left(r + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\nu - r\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\nu\right)}, \quad 2r < \nu. \quad (16.16)$$

Функция распределения статистики t может быть получена по таблицам бета-функции, если в (16.15) сделать преобразование

$$\xi = \left(1 + \frac{t^2}{\nu}\right)^{-1}.$$

Имеем

$$F(\tau) = \int_{-\infty}^{\tau} dF = \frac{1}{2} \int_0^{\tau} \frac{\Gamma\left\{\frac{1}{2}(v+1)\right\}}{\sqrt{v\pi} \Gamma\left(\frac{v}{2}\right) \left(1 + \frac{t^2}{v}\right)^{\frac{1}{2}(v+1)}},$$

откуда

$$\begin{aligned} 2F - 1 &= \frac{2}{B\left(\frac{v}{2}, \frac{1}{2}\right)} \int_0^{\tau} \frac{1}{\left(1 + \frac{t^2}{v}\right)^{\frac{1}{2}(v+1)}} \frac{dt}{\sqrt{v}} = \\ &= \frac{1}{B\left(\frac{v}{2}, \frac{1}{2}\right)} \int_{\xi(\tau)}^1 \xi^{\frac{1}{2}v-1} (1-\xi)^{-\frac{1}{2}} d\xi = 1 - I_{\xi}\left(\frac{v}{2}, \frac{1}{2}\right). \end{aligned}$$

Следовательно,

$$F = 1 - \frac{1}{2} I_{\xi}\left(\frac{v}{2}, \frac{1}{2}\right). \quad (16.17)$$

Значения аргумента ξ , для которого I принимает значения 0,50, 0,25, 0,10, 0,05, 0,025, 0,01, 0,005 и $v=1(1)30, 40, 60, 120, \infty$, даны с пятью значащими цифрами (Томпсон и др. (1941)) и могут, следовательно, использоваться для отыскания значений t , отвечающих этим уровням вероятностей. Эти таблицы*) воспроизведены в *Biometrika Tables*.

16.12 Однако за исключением специальных случаев использование В-функции вовсе не обязательно, поскольку имеются как асимптотическое разложение функции распределения самой статистики t , так и таблицы, основанные на этом разложении.

Мы имеем

$$-\log\left(1 + \frac{t^2}{v}\right) = -\frac{t^2}{v} + \frac{t^4}{2v^2} - \dots + \frac{(-t^2)^j}{jv^j} + \dots,$$

и, следовательно,

$$\begin{aligned} -\frac{1}{2}(v+1)\log\left(1 + \frac{t^2}{v}\right) &= -\frac{1}{2}t^2 + \dots - \\ &\quad - \frac{j(-t^2)^{j+1} + (j+1)(-t^2)^j}{2j(j+1)v^j} + \dots \quad (16.18) \end{aligned}$$

*) Эти таблицы полностью воспроизведены в сборнике *Таблицы математической статистики* с добавлением двух новых уровней $I=0,001$ и $0,0025$. (Прим. ред.)

Далее, из разложения $\log \Gamma(1+x)$ находим

$$\log \left\{ \frac{\Gamma\left\{\frac{1}{2}(v+1)\right\}}{\Gamma\left(\frac{1}{2}v\right)} \sqrt{\frac{2}{v}} \right\} = -\frac{1}{4v} + \frac{1}{24v^3} - \frac{1}{20v^5} \dots \quad (16.19)$$

Когда $v \rightarrow \infty$, распределение t сходится к распределению нормированной нормальной случайной величины, что, впрочем, мы уже видели раньше. Записывая

$$y = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right), \quad (16.20)$$

мы можем воспользоваться (16.18), (16.19) и (16.20) для разложения логарифма плотности (16.15) по убывающим степеням v :

$$\log y + \frac{1}{4v}(t^4 - 2t^2 - 1) - \frac{1}{12v^2}(2t^6 - 3t^4) + \frac{1}{24v^3}(3t^8 - 4t^6 + 1) - \\ - \frac{1}{40v^4}(4t^{10} - 5t^8) + \frac{1}{60v^5}(5t^{12} - 6t^{10} - 3) - \dots \quad (16.21)$$

Беря экспоненту от (16.21) и интегрируя от t до ∞ , найдем

$$1 - F = y \left\{ \frac{1}{4v} t(t^2 + 1) + \frac{1}{96v^2}(3t^6 - 7t^4 - 5t^2 - 3)t + \right. \\ \left. + \frac{1}{384v^3}(t^{10} - 11t^8 + 14t^6 + 6t^4 - 3t^2 - 15)t + \frac{1}{92160v^4}(15t^{14} - \right. \\ \left. - 375t^{12} + 2225t^{10} - 2141t^8 - 939t^6 - 213t^4 + 915t^2 + 945)t + \dots \right\}. \quad (16.22)$$

Выражение (16.22) дано Фишером. Впрочем, оно было использовано самим Стьюдентом для вычисления функции распределения статистики t . Для значений $v \geq 20$ первые четыре члена в (16.22) дают значения F с максимальной ошибкой 0,000005.

16.13 Существует ряд таблиц t -распределения. Наиболее распространены следующие из них:

(а) Уже упоминавшиеся таблицы Стьюдента — Фишера для функции распределения. В них даются значения ф.р. для $t=0(0,1)6$ и $v=1(1)20$.

(б) Таблицы функции распределения, содержащиеся в *Biometrika Tables*, дают значения ф.р. для $t=0(0,1)8$ и $v=1(1)24; 30; 40; 60; 120; \infty$.

(в) Фишер и Иэйтс (1953) создали (обратные) таблицы для квантилей распределения, т. е. для значений t , соответствующих таким значениям функции распределения, которые удовлетворяют соотношениям $2(1-F)=0,9(0,1)0,1; 0,05; 0,02; 0,01; 0,001$ для $v=1(1)30; 40; 60; 120; \infty$. Эти таблицы приведены в приложении (таблица 5).

(г) Уже упоминавшиеся выше в 16.11 обратные таблицы Томпсон.

(д) Болдуин (1946) дал обратные таблицы для $2(1 - F) = 0,05; 0,01$ и $v = 1(1)30(2)100$.

(е) Федериги (Federighi, 1959, J. Amer. Statist. Ass. 54, 683) дал обратные таблицы для 20 значений $1 - F$ в интервале между 0,25 и 0,0000001 при $v = 1(1)30(5)60(10)100; 200; 500; 1000; 2000; 10\ 000; \infty$.

(ж) Смирнов (1960, *Таблицы функций распределения и плотностей распределения Стьюдента* (под ред. Н. В. Смирнова), М., Изд-во АН СССР) табулировал ф.р. и ф.п. с шестью десятичными знаками для $t = 0(0,01)3(0,02)4,5(0,05)6,5$ и $v = 1(1)12$, а также для $t = 0(0,01)2,5(0,02)3,5(0,05)6,5$ и $v = 13(1)25$. Кроме того, там же даны таблицы ф.р. для $t = 0(0,01)2,5(0,02)3,5(0,05)5$ и $v = 25(1)35$. Имеются вспомогательные таблицы для больших значений t и v . Обратные таблицы *) содержат квантили распределения Стьюдента, соответствующие вероятностям $2(1 - F) = 0,4; 0,25; 0,10; 0,05; 0,01; 0,005; 0,0025; 0,001; 0,0005$ и количествам степеней свободы $v = 1(1)30(10)100, 120, 150(50)500$ и т. д.

16.14 Заканчивая изложение материала, посвященного t -распределению, отметим, что во многих приложениях интерес представляет не само t -распределение, а распределение для t^2 . Конечно, переход от t к t^2 в (16.15) не представляет трудностей. Оказывается, однако, что распределение для t^2 есть просто специальный случай более общих распределений, связанных с нормальным, так называемых F - или z -распределений Фишера, играющих, как и распределения χ^2 и t , важную роль в статистической теории. К их исследованию мы сейчас и переходим.

F - и z -распределения Фишера

16.15 При обсуждении распределения отношений мы видели (пример 11.20), что в случае независимых нормальных выборок с одинаковыми дисперсиями, взятых из двух генеральных совокупностей, выборочное распределение для отношения

$$F = \frac{\sum_{i=1}^{n_1} (x_{1i} - \bar{x}_1)^2 / (n_1 - 1)}{\sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2 / (n_2 - 1)} \quad (16.23)$$

*) В таблице квантилей имеются ошибки. Исправленная таблица квантилей t -распределения опубликована в сборнике *Таблицы математической статистики*, с помощью которого можно находить значения ф.р. с пятью десятичными знаками и квантили с четырьмя десятичными знаками при всех степенях свободы. (Прим. ред.)

выражается формулой

$$dG = \frac{v_1^{\frac{1}{2}} v_2^{\frac{1}{2}} F^{\frac{1}{2}(v_1-2)} dF}{B\left(\frac{1}{2} v_1, \frac{1}{2} v_2\right) (v_1 F + v_2)^{\frac{1}{2}(v_1+v_2)}}, \quad v_1, v_2 > 0, \quad 0 \leq F \leq \infty, \quad (16.24)$$

где $v_1 = n_1 - 1$, $v_2 = n_2 - 1$, как и раньше, — числа «степеней свободы».

Сравнивая (16.24) с (16.15), мы замечаем, что если положить $v_1 = 1$ и $v_2 = v$, то (16.24) становится идентичным с элементом распределения t^2 , получаемым немедленно из (16.15). Как можно видеть в примере 11.20, при выводе распределения (16.24) был использован лишь тот факт, что $\sum_i (x_{1i} - \bar{x}_1)^2 / \sum_j (x_{2j} - \bar{x}_2)^2$ является отношением двух независимых величин χ^2 с v_1 и v_2 степенями свободы для числителя и знаменателя соответственно. Отсюда следует, что формула (16.24) справедлива для любого отношения такого типа. Например, возвращаясь к (16.23), предположим, что нормальные совокупности имеют различные дисперсии, скажем σ_1^2 и σ_2^2 . Тогда сразу получим, что $\frac{\sigma_2^2}{\sigma_1^2} F$, где F задается формулой (16.23), имеет распределение (16.24), ибо для каждой выборки величина

$$\sum_i (x_{pi} - \bar{x}_p)^2 / \sigma_p^2, \quad p = 1, 2,$$

подчиняется χ^2 -распределению, поскольку стандартное отклонение является масштабным множителем.

16.16 При рассмотрении примера 11.20 мы ввели величину

$$z = \frac{1}{2} \log F, \quad (16.25)$$

распределение которой немедленно получается из (16.24):

$$dH = \frac{2v_1^{\frac{1}{2}} v_2^{\frac{1}{2}}}{B\left(\frac{1}{2} v_1, \frac{1}{2} v_2\right)} \frac{\exp(v_1 z)}{\{v_1 \exp(2z) + v_2\} \frac{1}{2}(v_1 + v_2)}, \quad (16.26)$$

$$v_1, v_2 > 0, \quad -\infty \leq z \leq \infty.$$

Именно в такой форме Р. А. Фишером и было получено первоначально распределение для отношения дисперсий. В современной

практике предпочитают использовать более простую статистику F , хотя распределение статистики z было табулировано раньше и первоначальные исследования относились к z -статистике, а не к F .

Свойства F - и z -распределений

16.17 Функция распределения статистики F сосредоточена на полупрямой $(0, \infty)$, и при $v_1 \leq 2$ плотность этого распределения имеет J -образную форму. Для случая $v_1 > 2$ — это одновершинное асимметричное распределение с модальным значением

$$\tilde{F} = \frac{v_1 - 2}{v_1} \frac{v_2}{v_2 + 2}. \quad (16.27)$$

Непосредственно из (16.24) легко вывести, что среднее и дисперсия F выражаются формулами

$$\left. \begin{aligned} \mu'_1 &= v_2 / (v_2 - 2), \quad v_2 > 2; \\ \mu'_2 &= 2v_2^2(v_1 + v_2 - 2) / \{v_1(v_2 - 2)^2(v_2 - 4)\}, \quad v_2 > 4. \end{aligned} \right\} \quad (16.28)$$

Условия, налагаемые на v_2 , обеспечивают существование этих моментов.

Сравнение (16.27) и (16.28) показывает, что для $v_1, v_2 > 2$ F -распределение всегда имеет модальное значение, меньшее $F=1$, и среднее значение, большее $F=1$. Отсюда следует, что при этих условиях распределение имеет «положительную» асимметрию, в других же случаях это очевидно.

16.18 Характеристическая функция статистики z пропорциональна

$$\int_{-\infty}^{\infty} \frac{e^{(\theta + v_1)z} dz}{(v_1 e^{2z} + v_2)^{\frac{1}{2}(v_1 + v_2)}},$$

где $\theta = it$, и поэтому

$$\varphi(t) = \left(\frac{v_2}{v_1}\right)^{\frac{1}{2}\theta} \frac{\Gamma\left\{\frac{1}{2}(v_2 - 0)\right\} \Gamma\left\{\frac{1}{2}(v_1 + \theta)\right\}}{\Gamma\left(\frac{1}{2}v_1\right) \Gamma\left(\frac{1}{2}v_2\right)}. \quad (16.29)$$

Поэтому, беря логарифмы и используя разложение Стирлинга

$$\log \Gamma(1+x) = \frac{1}{2} \log 2\pi + \left(x + \frac{1}{2}\right) \log x - x + \frac{1}{12x} - \dots,$$

из (16.29) найдем для больших v_1 и v_2 , что

$$\log \varphi(t) = \frac{\theta}{2} \left(\frac{1}{v_2} - \frac{1}{v_1}\right) + \frac{\theta^2}{4} \left(\frac{1}{v_1} + \frac{1}{v_2}\right) - \dots \quad (16.30)$$

Для больших v_1 и v_2 статистика z распределена приблизительно нормально с моментами

$$\mu'_1 = \frac{1}{2} \left(\frac{1}{v_2} - \frac{1}{v_1} \right), \quad \mu_2 = \frac{1}{2} \left(\frac{1}{v_1} + \frac{1}{v_2} \right). \quad (16.31)$$

16.19 Значение функции распределения для F (или z) можно получить из таблиц неполной В-функции, используя формулу (16.24) и преобразование

$$\xi = (1 + v_1 F / v_2)^{-1}.$$

Мы можем тогда поступить точно так же, как и в 16.11 с соотношением (16.17) (нужно, конечно, помнить, что в (16.17) в левой части обозначение F относится к значениям функции распределения, а не к самим значениям случайной величины). Однако существуют специальные таблицы функций распределения для F , что в обычных применениях делает указанную процедуру вовсе не обязательной.

В силу того, что имеется два количества «степеней свободы», таблицы для F имеют три входа. Поэтому удобнее использовать менее громоздкие обратные таблицы квантилей распределения. Такие таблицы создали:

(а) Фишер и Иэйтс (1953); в таблицах содержатся значения F (и z), отвечающие значениям $(1 - P)$ функции распределения для $P = 0,20; 0,10; 0,05; 0,01; 0,001$; $v_1 = 1(1)6; 8; 12; 24; \infty$ и $v_2 = 1(1)30; 40; 60; 120; \infty$.

(б) Меррингтон и Томпсон (1943); эти таблицы*) также помещены в *Biometrika Tables*. В них приведены значения аргумента F , отвечающие таким значениям $(1 - P)$ функции распределения, для которых $P = 0,50; 0,25; 0,10; 0,05; 0,025; 0,01; 0,005; 0,001$; $v_1 = 1(1)10; 12; 15; 20; 24; 30; 40; 60; 120; \infty$ и $v_2 = 1(1)30; 40; 60; 120; \infty$.

16.20 Как это легко видно из определения F как отношения, мы можем поменять числитель и знаменатель местами и получить распределение той же самой формы, но с заменой v_1 на v_2 и v_2 на v_1 . Это можно доказать, сделав преобразование $F = 1/y$ в (16.24). В терминах функции распределения $G_{v_1, v_2}(F)$ для статистики F это означает, что

$$G_{v_1, v_2}(F) = 1 - G_{v_2, v_1} \left(\frac{1}{F} \right). \quad (16.32)$$

*) Эти таблицы полностью воспроизведены в сборнике *Таблицы математической статистики* с исправлением ошибок и с добавлением еще одного уровня $P = 0,0005$. (Прим. ред.)

В силу взаимно однозначного соответствия между F и z и того, что $\log \frac{1}{F} = -\log F$, (16.32) эквивалентно соотношению

$$H_{v_1, v_2}(z) = 1 - H_{v_2, v_1}(-z). \quad (16.33)$$

Соотношения (16.32) и (16.33) делают необязательным табулирование этих распределений для всех значений z . В самом деле, если мы интересуемся значением функции распределения для $F < 1$ (т. е. для отрицательных значений z), то мы можем обратить отношение (т. е. поменять знак у z) и использовать указанные соотношения для получения тех значений функции распределения, которыми мы интересуемся.

16.21 Существуют различные аппроксимации для случаев, когда v_1 и v_2 недостаточно велики для того, чтобы использование нормального приближения было оправдано.

(а) Аппроксимация Корниша и Фишера (1937). Их метод — тот же, что и в 6.25, и связан с разложением в ряды Грама — Шарлье. Последовательным дифференцированием $\log \Gamma(1+x)$ мы можем найти производные $\varphi(t)$ и получить семиинварианты для z . Полагая $r_1 = \frac{1}{v_1}$, $r_2 = \frac{1}{v_2}$, найдем

$$\left. \begin{aligned} \kappa_1 &= -\frac{1}{2}(r_1 - r_2) - \frac{1}{6}(r_1^2 - r_2^2), \\ \kappa_2 &= \frac{1}{2}(r_1 + r_2) + \frac{1}{2}(r_1^2 + r_2^2) + \frac{1}{3}(r_1^3 + r_2^3), \\ \kappa_3 &= -\frac{1}{2}(r_1^2 - r_2^2) - (r_1^3 - r_2^3), \\ \kappa_4 &= r_1^3 + r_2^3 + 3(r_1^4 + r_2^4), \\ \kappa_5 &= -3(r_1^4 - r_2^4), \\ \kappa_6 &= 12(r_1^5 + r_2^5), \end{aligned} \right\} \quad (16.34)$$

откуда, полагая $\sigma = r_1 + r_2$ и $\delta = r_1 - r_2$, находим значения l , определенные в 6.25 ($m=0$, дисперсия $= \frac{1}{2} \sigma$):

$$\begin{aligned} l_1 &= -\sqrt{\frac{2}{\sigma}} \left(\frac{1}{2} \delta + \frac{1}{6} \delta \sigma \right), \\ l_2 &= \frac{1}{2} \left(\sigma + \frac{\delta^2}{\sigma} \right) + \frac{1}{6} (\sigma^2 + 3\delta^2) \text{ и т. д.} \end{aligned}$$

После ряда преобразований найдем значение z , соответствующее вероятности α (которая в свою очередь соответствует

нормальному отклонению ξ):

$$z = \xi \sqrt{\frac{\sigma}{2}} - \frac{1}{6} \delta (\xi^2 + 2) + \sqrt{\frac{\sigma}{2}} \left\{ \frac{\sigma}{24} (\xi^3 + 3\xi) + \right. \\ \left. + \frac{1}{72} \frac{\delta^2}{\sigma} (\xi^3 + 11\xi) \right\} - \frac{\delta\sigma}{120} (\xi^4 + 9\xi^2 + 8) + \frac{\delta^3}{3240\sigma} (3\xi^4 + 7\xi^2 - 16) + \\ + \sqrt{\frac{\sigma}{2}} \left\{ \frac{\sigma^2}{1920} (\xi^5 + 20\xi^3 + 15\xi) + \frac{\delta^4}{2880} (\xi^5 + 44\xi^3 + 183\xi) + \right. \\ \left. + \frac{\delta^4}{155520\sigma^2} (9\xi^5 - 284\xi^3 - 1513\xi) \right\}. \quad (16.35)$$

(6) Аппроксимация Фишера и ее обобщение, данное Кокрэном (1940). Обозначая v_1 и v_2 буквой n , из (16.35) с точностью до величин порядка $n^{-3/2}$ получим

$$z = \xi \sqrt{\frac{\sigma}{2}} - \frac{1}{6} \delta (\xi^2 + 2) + \sqrt{\frac{\sigma}{2}} \left\{ \frac{\sigma}{24} (\xi^3 + 3\xi) + \frac{1}{72} \frac{\delta^2}{\sigma} (\xi^3 + 11\xi) \right\}.$$

Положим $h = 2/\sigma$. Тогда

$$z = \frac{\xi}{\sqrt{h}} - \frac{1}{6} \delta (\xi^2 + 2) + \frac{1}{\sqrt{h}} \left\{ \frac{\xi^3 + 3\xi}{12h} + \frac{\xi^3 + 11\xi}{144} h \delta^2 \right\}. \quad (16.36)$$

Но

$$\frac{\xi}{\sqrt{h} - \lambda} = \frac{\xi}{\sqrt{h}} + \frac{\lambda \xi}{2h\sqrt{h}} + o(n^{-2}).$$

Следовательно, если мы обозначим

$$z = \frac{\xi}{\sqrt{h} - \lambda} - \frac{1}{6} \delta (\xi^2 + 2), \quad (16.37)$$

то разность между правыми частями (16.37) и (16.36) при условии, что $\lambda = \frac{\xi^3 + 3}{6}$, будет равна

$$\frac{(\xi^3 + 11\xi) \delta^2 \sqrt{h}}{144}.$$

Эта разность мала в силу большого знаменателя и множителя $\delta^2 = \left(\frac{1}{v_1} - \frac{1}{v_2} \right)^2$, который мал, если v_1 и v_2 не очень различны между собой. Поэтому приближенно можно считать, что статистика z распределена как величина, определяемая соотношением (16.37). Значения λ для различных значений ф. р. (1—P) таковы:

100%	40%	30%	20%	10%	5%	1%	0,1%
λ	0,51	0,55	0,62	0,77	0,95	1,40	2,09

Для типичных значений P правая часть (16.37) равна соответственно

$$20\%: \frac{0,8416}{\sqrt{h-\lambda}} - 0,4514\delta, \quad (16.38)$$

$$5\%: \frac{1,6449}{\sqrt{h-\lambda}} - 0,7843\delta, \quad (16.39)$$

$$1\%: \frac{2,3263}{\sqrt{h-\lambda}} - 1,235\delta, \quad (16.40)$$

$$0,1\%: \frac{3,0902}{\sqrt{h-\lambda}} - 1,925\delta. \quad (16.41)$$

Точность аппроксимации при $v_1=24$, $v_2=60$ ясна из следующего сравнения:

100P %	Значения z из (16.37)	Точные значения
20	0,1337	0,1338
1	0,3748	0,3746
0,1	0,4966	0,4955

(в) Аппроксимация Полсона (1942). Аппроксимация Уилсона и Хилферти для χ^2 (см. 16.7) показывает, что отношение $\left(\frac{\chi^2}{v}\right)^{1/3}$ распределено примерно нормально со средним $1 - \frac{2}{9v}$ и дисперсией $\frac{2}{9v}$. Отношение $\frac{s_1^2}{s_2^2}$ само есть отношение двух независимых величин, распределенных как χ^2/v , соответственно с v_1 и v_2 степенями свободы. Далее, в силу теоремы Гири (упражнение 11.11) отношение $\frac{m_1 - m_2 p}{(\sigma_1^2 + \sigma_2^2 p^2)^{1/2}}$ имеет нормированное нормальное распределение, где p^2 — отношение s_1^2/s_2^2 . Поэтому случайную величину

$$u = \frac{\left(1 - \frac{2}{9v_2}\right) \left(\frac{s_1}{s_2}\right)^{2/3} - \left(1 - \frac{2}{9v_1}\right)}{\left\{ \frac{2}{9v_2} \left(\frac{s_1}{s_2}\right)^{4/3} + \frac{2}{9v_1} \right\}^{1/2}} \quad (16.42)$$

можно рассматривать как имеющую (приблизленно) нормированное нормальное распределение. Эта аппроксимация является

исключительно хорошей *). Следующая таблица дает точные и приближенные значения p^2 при $v_1=6$, $v_2=12$.

100P %	$\left(\frac{s_1}{s_2}\right)^2 = p^2$ из (16.42)	Точные значения
20	1,72	1,72
5	3,00	3,00
1	4,85	4,82
0,1	8,58	8,38

16.22. Мы закончили изучение трех фундаментальных выборочных распределений, возникающих в связи с рассмотрением одномерного нормального распределения. Но прежде чем переходить к двумерному нормальному распределению, приведем сводку основных свойств этих распределений, их соотношений между собой и с исходным нормальным распределением.

(1) Сумма квадратов n нормированных нормальных случайных величин в случае, когда за начало отсчета берется истинное (или выборочное) среднее, имеет χ^2 -распределение с n (или $n-1$) степенями свободы.

(2) χ^2 -распределение аппроксимируется нормальным (с ростом числа степеней свободы).

(3) Отношение квадрата нормальной случайной величины с нулевым средним к независимой от нее оценке ее дисперсии (предполагается, что эта оценка распределена как χ^2/v) имеет t^2 -распределение Стьюдента, а квадратный корень этого отношения подчиняется t -распределению Стьюдента.

(4) Стьюдентовское t -распределение стремится к нормированному нормальному (с ростом числа степеней свободы).

(5) Отношение двух независимых случайных величин, каждая из которых распределена как χ^2/v , имеет F -распределение Фишера, а половина его логарифма имеет z -распределение. Распределение t^2 является специальным случаем F -распределения,

*) Не менее точная и, пожалуй, более простая аппроксимация указана в сборнике *Таблицы математической статистики* для случая $v_1 < v_2$:

$$p^2 \approx \frac{v_2}{v_1} \frac{2y}{v_1 + 2v_2 - 2 - y - \frac{[v_1^2 - 4 + (v_1 - 2)y - 2y^2]}{6(v_1 + 2v_2 - 2)}},$$

где y есть 100P%-ная точка χ^2 -распределения с v_1 степенями свободы. По этой формуле в данном примере получаются соответственно значения $p^2 \approx 1,72$; 3,00; 4,82; 8,32. (Прим. ред.)

когда случайная величина в числителе имеет только одну степень свободы, см. выше (3) и (1).

(6) Когда число степеней свободы и числителя и знаменателя увеличивается, F -распределение стремится к нормальному.

(7) Если только число степеней свободы знаменателя, v_2 , стремится к бесконечности, то распределение $v_1 F$ стремится к χ^2 -распределению с v_1 степенями свободы.

Наличие указанных связей между распределениями часто позволяет использовать какие-то одни таблицы для многих различных целей.

Двумерное нормальное распределение

16.23 Мы видели (пример 15.1), что элемент нормального распределения двух случайных величин x и y может быть записан в виде

$$dF = \frac{dx dy}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right\}, \quad (16.43)$$

где σ_1^2, σ_2^2 — дисперсии, μ_{11} — их ковариация и $\rho = \mu_{11}/(\sigma_1\sigma_2)$ — так называемый коэффициент корреляций между x и y . В (16.43) мы считали средние значения у случайных величин равными нулю. Это предположение упрощает многие выкладки без потери общности получаемых результатов.

Переписывая показатель экспоненты в (16.43) как

$$-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x}{\sigma_1} - \rho\frac{y}{\sigma_2}\right)^2 + \frac{y^2}{\sigma_2^2}(1-\rho^2)\right\} \quad (16.44)$$

и

$$-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{y}{\sigma_2} - \rho\frac{x}{\sigma_1}\right)^2 + \frac{x^2}{\sigma_1^2}(1-\rho^2)\right\}, \quad (16.45)$$

мы видим из (16.44), что условное распределение x при заданном значении y является нормальным со средним и дисперсией

$$\left. \begin{aligned} M(x|y) &= \rho y \sigma_1 / \sigma_2, \\ D(x|y) &= \sigma_1^2 (1 - \rho^2). \end{aligned} \right\} \quad (16.46)$$

С помощью (16.45) аналогичный результат можно получить и для y при заданном x , если в (16.46) поменять местами x, y и индексы 1 и 2.

Из (16.46) следует, что условное среднее значение каждой случайной величины при заданном значении другой есть линейная функция от этого значения и что условная дисперсия постоянна и не зависит от этого значения. Этот факт выражают

словами: «Регрессия x на y (и y на x) линейна и гомоскедастична».

В последующих главах (второй том) мы будем изучать теорию корреляции и регрессии в деталях. В оставшихся пунктах этой главы мы остановимся на выводах распределений некоторых статистик, определяемых по выборке из двумерной нормальной совокупности.

16.24 Совместное распределение n независимых выборочных значений $(x_1, y_1), \dots, (x_n, y_n)$ из двумерной нормальной совокупности со средними μ_1 и μ_2 задается формулой

$$dF = \frac{1}{(2\pi)^n \sigma_1^n \sigma_2^n (1-\rho^2)^{\frac{1}{2}n}} \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \sum \left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \sum \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1 \sigma_2} + \sum \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right\} \right] dx_1 dy_1 \dots dx_n dy_n. \quad (16.47)$$

Показатель экспоненты в (16.47) может быть целиком выражен через пять параметров распределения $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ и выборочные статистики

$$\left. \begin{aligned} \bar{x} &= \frac{1}{n} \sum x, \quad \bar{y} = \frac{1}{n} \sum y, \\ s_1^2 &= \frac{1}{n} \sum (x - \bar{x})^2, \quad s_2^2 = \frac{1}{n} \sum (y - \bar{y})^2, \\ r &= \frac{1}{n} \sum (x - \bar{x})(y - \bar{y}) / (s_1 s_2). \end{aligned} \right\} \quad (16.48)$$

Например, опуская множитель $-\frac{1}{2(1-\rho^2)}$ при первом члене в экспоненте (16.47), получим

$$\sum \left(\frac{x-\mu_1}{\sigma_1} \right)^2 = \sum \left\{ \frac{(x-\bar{x}) + (\bar{x}-\mu_1)}{\sigma_1} \right\}^2 = \frac{n}{\sigma_1^2} \{s_1^2 + (\bar{x}-\mu)^2\}. \quad (16.49)$$

При этом смешанные произведения исчезают, поскольку $\sum (x - \bar{x}) = 0$. Другие два слагаемых в экспоненте могут быть преобразованы аналогичным образом. В целом получим выражение

$$\begin{aligned} n \left\{ \left(\frac{\bar{x}-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{\bar{x}-\mu_1}{\sigma_1} \right) \left(\frac{\bar{y}-\mu_2}{\sigma_2} \right) + \left(\frac{\bar{y}-\mu_2}{\sigma_2} \right)^2 \right\} + \\ + n \left\{ \frac{s_1^2}{\sigma_1^2} - \frac{2\rho r s_1 s_2}{\sigma_1 \sigma_2} + \frac{s_2^2}{\sigma_2^2} \right\}. \quad (16.50) \end{aligned}$$

Сейчас мы найдем совместное распределение всех пяти статистик, определенных формулами (16.48). Чтобы это сделать, мы должны подходящим образом преобразовать (16.47). С помощью (16.50) мы можем выразить коэффициент при дифференциалах в (16.47) через пять статистик и пять параметров. Рассмотрим теперь дифференциал в (16.47).

Обобщая геометрический подход из главы 11, мы рассмотрим выборочное пространство n переменных для x и другое такое же пространство для y , пристроенное к первому. Выборочные точки изменяются в каждом из этих пространств, но не независимым образом. В самом деле, если P представляет точку $(x_1, \dots, x_n)$ в x -пространстве, а Q — точку $(y_1, \dots, y_n)$ в y -пространстве и если O_1 и O_2 есть точки $(\bar{x}, \dots, \bar{x})$, $(\bar{y}, \dots, \bar{y})$, тогда

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{ns_1s_2} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\left\{ \sum (x - \bar{x})^2 \sum (y - \bar{y})^2 \right\}^{1/2}},$$

и поэтому r есть косинус угла θ между PO_1 и QO_2 ; так что если P и r фиксированы, то Q меняется в таком конусе в пространстве y , который получен вращением O_2Q так, чтобы его угол с Q_1P оставался постоянным. Как было показано в примере 11.7, дифференциальный элемент в x -пространстве пропорционален $s_1^{n-2} ds_1 d\bar{x}$. Для заданных $r, \bar{y}$ и s_2 точка Q принадлежит $(n-2)$ -мерной поверхности, лежащей на гиперсфере размерности $(n-1)$ с радиусом $s_2 \sqrt{n}$ и с центром в точке y . Но в силу того, что угол θ фиксирован, точка Q в действительности принадлежит $(n-3)$ -мерной поверхности. Эта $(n-3)$ -мерная поверхность имеет радиус $s_2 \sqrt{n} \sin \theta$ и «толщину» $s_2 \sqrt{n} d\theta = \frac{s_2 \sqrt{n} dr}{(1-r^2)^{1/2}}$. Поэтому ее «объем» пропорционален

$$\{s_2 \sqrt{n} (1-r^2)^{1/2}\}^{n-3} \frac{s_2 \sqrt{n} dr}{(1-r^2)^{1/2}},$$

т. е. пропорционален

$$s_2^{n-2} (1-r^2)^{\frac{1}{2}(n-4)}.$$

Тем самым для элемента объема можно записать формулу

$$\begin{aligned} dv &\propto s_1^{n-2} ds_1 d\bar{x} s_2^{n-2} (1-r^2)^{\frac{1}{2}(n-4)} ds_2 d\bar{y} dr \propto \\ &\propto s_1^{n-2} s_2^{n-2} ds_1 ds_2 (1-r^2)^{\frac{1}{2}(n-4)} dr d\bar{x} d\bar{y}. \quad (16.51) \end{aligned}$$

Из (16.47), (16.50) и (16.51) следует, что элемент совместного распределения пяти случайных величин пропорционален

выражению

$$\exp\left(-\frac{n}{2(1-\rho_2)}\left[\left\{\frac{(\bar{x}-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(\bar{x}-\mu_1)(\bar{y}-\mu_2)}{\sigma_1\sigma_2}+\frac{(\bar{y}-\mu_2)^2}{\sigma_2^2}\right\}+\right.\right. \\ \left.\left.+\left\{\frac{s_1^2}{\sigma_1^2}-2\rho r\frac{s_1s_2}{\sigma_1\sigma_2}+\frac{s_2^2}{\sigma_2^2}\right\}\right]\right)dv. \quad (16.52)$$

Этот фундаментальный результат был установлен Р. А. Фишером (1915).

16.25 Отметим одно очень важное свойство (16.52). Это распределение может быть факторизовано на две части: одну, содержащую только $\bar{x}$ и $\bar{y}$, и другую, содержащую s_1 , s_2 и r :

$$dF \propto \exp\left[-\frac{n}{2(1-\rho^2)}\left\{\frac{(\bar{x}-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(\bar{x}-\mu_1)(\bar{y}-\mu_2)}{\sigma_1\sigma_2}+\frac{(\bar{y}-\mu_2)^2}{\sigma_2^2}\right\}\right]d\bar{x}d\bar{y} \quad (16.53)$$

и

$$dF \propto \exp\left[-\frac{n}{2(1-\rho^2)}\left\{\frac{s_1^2}{\sigma_1^2}-\frac{2\rho rs_1s_2}{\sigma_1\sigma_2}+\frac{s_2^2}{\sigma_2^2}\right\}\right]\times \\ \times s_1^{n-2}s_2^{n-2}(1-r^2)^{\frac{1}{2}(n-4)}ds_1ds_2dr. \quad (16.54)$$

Тем самым для нормального случая средние значения выборки не зависят от дисперсий и ковариации выборки. Это свойство является характеристическим для многомерного нормального распределения, см. **15.24**.

Из (16.53) также видно, что сами средние значения нормально распределены с $M(\bar{x})=\mu_1$, $M(\bar{y})=\mu_2$, $D\bar{x}=\frac{\sigma_1^2}{n}$, $D\bar{y}=\frac{\sigma_2^2}{n}$ (все эти результаты нам уже известны) и

$$\text{cov}(\bar{x}, \bar{y}) = \frac{\sigma_1\sigma_2\rho}{n}, \quad (16.55)$$

так что коэффициент корреляции между $\bar{x}$ и $\bar{y}$ равен ρ , т. е. он остается таким же, как и у исходного распределения. Более общий результат приведен в упражнении 13.2.

16.26 Используем (16.54) для получения распределения коэффициента корреляции выборки интегрированием по s_1 и s_2 в пределах от 0 до ∞ . Вычислим прежде всего нормирующий множитель у распределения (16.54) из условия нормировки $\int dF = 1$.

Сделаем преобразование

$$\left. \begin{aligned} a &= \frac{s_1^2}{\sigma_1^2} \frac{n}{2(1-\rho^2)}, \\ b &= \frac{rs_1s_2}{\sigma_1\sigma_2} \frac{n}{2(1-\rho^2)}, \\ c &= \frac{s_2^2}{\sigma_2^2} \frac{n}{2(1-\rho^2)}. \end{aligned} \right\} \quad (16.56)$$

Его якобиан равен

$$\begin{aligned} \frac{\partial(a, b, c)}{\partial(s_1, r, s_2)} &= \begin{vmatrix} \frac{2s_1}{\sigma_1^2} \frac{n}{2(1-\rho^2)} & 0 & 0 \\ \frac{rs_2}{\sigma_1\sigma_2} \frac{n}{2(1-\rho^2)} & \frac{s_1s_2}{\sigma_1\sigma_2} \frac{n}{2(1-\rho^2)} & \frac{rs_1}{\sigma_1\sigma_2} \frac{n}{2(1-\rho^2)} \\ 0 & 0 & \frac{2s_2}{\sigma_2^2} \frac{n}{2(1-\rho^2)} \end{vmatrix} = \\ &= \frac{s_1^2 s_2^2 n^3}{2\sigma_1^3 \sigma_2^3 (1-\rho^2)^3} = \frac{2acn}{\sigma_1\sigma_2(1-\rho^2)}. \end{aligned}$$

Из (16.56) имеем также

$$r^2 = \frac{b^2}{ac}.$$

Тогда интеграл от (16.54) можно представить в следующем виде:

$$\begin{aligned} \int \exp[-a + 2\rho b - c] &\left\{ \frac{2a\sigma_1^2(1-\rho^2)}{n} \right\}^{\frac{1}{2}(n-2)} \left\{ \frac{2c\sigma_2^2(1-\rho^2)}{n} \right\}^{\frac{1}{2}(n-2)} \times \\ &\times \left(1 - \frac{b^2}{ac}\right)^{\frac{1}{2}(n-4)} \frac{\sigma_1\sigma_2(1-\rho^2)}{2acn} da db dc = \\ &= \frac{2^{n-3}\sigma_1^{n-1}\sigma_2^{n-1}(1-\rho^2)^{n-1}}{n^{n-1}} \int \exp[-a + 2\rho b - c] \times \\ &\times (ac - b^2)^{\frac{1}{2}(n-4)} da db dc, \quad (16.57) \end{aligned}$$

где пределы изменения a и c есть 0 и ∞ и $-\sqrt{ac} \leq b \leq \sqrt{ac}$. Этот интеграл можно вычислить с помощью Γ -функции,

Полагая

$\xi = a - \frac{b^2}{c}$, найдем ($0 \leq \xi \leq \infty$, $-\infty \leq b \leq \infty$, $0 \leq c \leq \infty$)

$$\begin{aligned} & \int \exp(-\xi) \exp\left(2\rho b - c - \frac{b^2}{c}\right) \xi^{\frac{1}{2}(n-4)} c^{\frac{1}{2}(n-4)} d\xi db dc = \\ &= \Gamma\left\{\frac{1}{2}(n-2)\right\} \int \exp\left[-\frac{1}{c} \{(b-\rho c)^2 + (1-\rho^2)c^2\}\right] c^{\frac{1}{2}(n-4)} db dc = \\ &= \Gamma\left\{\frac{1}{2}(n-2)\right\} \sqrt{\pi} \int \exp\{-(1-\rho^2)c\} c^{\frac{1}{2}(n-3)} dc = \\ &= \sqrt{\pi} \frac{\Gamma\left\{\frac{1}{2}(n-2)\right\} \Gamma\left\{\frac{1}{2}(n-1)\right\}}{(1-\rho^2)^{\frac{1}{2}(n-1)}} = \frac{\pi \Gamma(n-2)}{2^{n-3} (1-\rho^2)^{\frac{1}{2}(n-1)}}. \quad (16.58) \end{aligned}$$

Из (16.57) и (16.58) находим совместное распределение s_1 , s_2 и r :

$$\begin{aligned} dF &= \frac{n^{n-1}}{\pi \sigma_1^{n-1} \sigma_2^{n-1} (1-\rho^2)^{\frac{1}{2}(n-1)} \Gamma(n-2)} \times \\ &\times \exp\left[-\frac{n}{2(1-\rho^2)} \left\{ \frac{s_1^2}{\sigma_1^2} - \frac{2\rho r s_1 s_2}{\sigma_1 \sigma_2} + \frac{s_2^2}{\sigma_2^2} \right\}\right] \times \\ &\times s_1^{n-2} s_2^{n-2} (1-r^2)^{\frac{1}{2}(n-4)} ds_1 ds_2 dr. \quad (16.59) \end{aligned}$$

16.27 Положим теперь

$$\zeta = \frac{s_1 s_2}{\sigma_1 \sigma_2}, \quad \beta = \log \frac{\sigma_2 s_1}{\sigma_1 s_2}, \quad r = r.$$

Тогда якобиан преобразования равен

$$\frac{\partial(\zeta, \beta, r)}{\partial(s_1, s_2, r)} = \begin{vmatrix} \frac{s_2}{\sigma_1 \sigma_2} & \frac{s_1}{\sigma_1 \sigma_2} & 0 \\ \frac{1}{s_1} & -\frac{1}{s_2} & 0 \\ 0 & 0 & 1 \end{vmatrix} = -\frac{2}{\sigma_1 \sigma_2}.$$

Экспонента в (16.59) превратится в

$$\exp\left[-\frac{n}{2(1-\rho^2)} \{\zeta e^\beta - 2\rho r \zeta + \zeta e^{-\beta}\}\right],$$

и после простых преобразований получим

$$\begin{aligned} dF &= \frac{n^{n-1}}{\pi (1-\rho^2)^{\frac{1}{2}(n-1)} \Gamma(n-2)} \exp\left[-\frac{n}{(1-\rho^2)} \zeta (\operatorname{ch} \beta - \rho r)\right] \times \\ &\times \zeta^{n-2} d\zeta d\beta (1-r^2)^{\frac{1}{2}(n-4)} dr. \end{aligned}$$

Интегрируя по ζ в пределах от 0 до ∞ , найдем

$$dF = \frac{(1-\rho^2)^{\frac{1}{2}(n-1)} \Gamma(n-1)}{\pi \Gamma(n-2)} \frac{(1-r^2)^{\frac{1}{2}(n-4)}}{(\operatorname{ch} \beta - \rho r)^{n-1}} d\beta dr; \quad (16.60)$$

(16.60) является четной функцией от β , поэтому интеграл по β от $-\infty$ до ∞ равен удвоенному интегралу в пределах от 0 до ∞ . Положим $-\rho r = \cos \theta$, тогда, поскольку

$$\int_0^\infty \frac{d\beta}{\operatorname{ch} \beta + \cos \theta} = \frac{\theta}{\sin \theta},$$

то

$$\begin{aligned} dF &= \frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi \Gamma(n-2)} (1-r^2)^{\frac{1}{2}(n-4)} \frac{d^{n-2}}{d(-\cos \theta)^{n-2}} \left(\frac{\theta}{\sin \theta} \right) dr = \\ &= \frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi \Gamma(n-2)} (1-r^2)^{\frac{1}{2}(n-4)} \frac{d^{n-2}}{d(r\rho)^{n-2}} \left\{ \frac{\arccos(-\rho r)}{\sqrt{1-\rho^2 r^2}} \right\} dr. \end{aligned} \quad (16.61)$$

Эти результаты были получены Фишером (1915).

16.28 В частном случае $\rho=0$ (16.61) сводится, что, впрочем, можно увидеть непосредственно из (16.52), к элементу распределения

$$dF = \frac{1}{B\left\{\frac{1}{2}(n-2), \frac{1}{2}\right\}} (1-r^2)^{\frac{1}{2}(n-4)} dr. \quad (16.62)$$

Предположение о том, что именно таким должно быть распределение для случая $\rho=0$, было высказано Стьюдентом в 1908 году. Значения соответствующей функции распределения можно получить, используя таблицы неполной В-функции. Кроме того, полагая

$$t = \{(n-2)r^2/(1-r^2)\}^{1/2}, \quad (16.63)$$

можно (16.62) свести к t -распределению Стьюдента (16.15) с $\nu = n-2$ степенями свободы.

16.29 Общее распределение (16.61) было изучено довольно подробно. Мы укажем здесь только основные его свойства. Остановимся сначала на форме кривой плотности этого распределения. При $n=2$ распределение (16.61) не определено в силу наличия множителя $\Gamma(n-2)$. Это и понятно, поскольку для $n=2$ выборочный коэффициент корреляции r может быть или $+1$ или -1 . Этот случай, когда мы имеем дело с разрывным

распределением, можно рассматривать как предельный случай U-образного распределения.

При $n=3$ мы видим, что функция плотности

$$y = y_0 \left\{ \frac{1}{\sin^2 \theta} - \frac{\theta \cos \theta}{\sin^3 \theta} \right\} \frac{1}{(1-r^2)^{\frac{1}{2}}}$$

имеет U-образную форму. Для $n=4$

$$y = \frac{y_0}{\sin^3 \theta} (\theta - 3 \operatorname{ctg} \theta + 3\theta \operatorname{ctg}^2 \theta).$$

Если $\rho=0$, то это распределение является равномерным с плотностью $y \equiv \frac{1}{2}$. В других случаях кривая плотности J-образна.

Для $n>4$ кривая плотности унимодальна и ее асимметрия растет при возрастании $|\rho|$, как это следует из того факта, что r , будучи косинусом угла, удовлетворяет условию $r^2 \leq 1$. Для любого ρ при $n \rightarrow \infty$ распределение r стремится к нормальному, хотя и очень медленно. Некоторые интересные фотографии кривых этих распределений даны в *Co-operative Study* (1917).

16.30 Плотность распределения (16.61) можно, как это было отмечено Хотеллингом (1953), представить в виде гипергеометрического ряда. Рассмотрим интеграл

$$I_{n-1} = \int_0^\infty \frac{d\beta}{(\operatorname{ch} \beta - \rho r)^{n-1}}. \quad (16.64)$$

Полагая $\operatorname{ch} \beta = (1 - \rho rz)/(1 - z)$, преобразуем (16.64) к виду

$$I_{n-1} = \frac{1}{\sqrt{2}(1-\rho r)^{n-3/2}} \int_0^1 z^{-1/2} (1-z)^{n-2} \left\{ 1 - \frac{1}{2}(1+\rho r)z \right\}^{-1/2} dz.$$

Разложим теперь выражение в фигурных скобках под знаком интеграла в равномерно сходящийся ряд и проинтегрируем почленно. Тогда получим

$$I_{n-1} = \frac{B\left(\frac{1}{2}, n-1\right)}{\sqrt{2}(1-\rho r)^{n-\frac{3}{2}}} F\left(\frac{1}{2}, \frac{1}{2}, n-\frac{1}{2}, \frac{1}{2}(1+\rho r)\right), \quad (16.65)$$

где приняты обычные гипергеометрические обозначения.

Итак, (16.60), (16.64) и (16.65) дают для (16.61) представление

$$dF = \frac{(n-2) dr}{(n-1) \sqrt{2} B\left(\frac{1}{2}, n-\frac{1}{2}\right)} (1-\rho^2)^{\frac{1}{2}(n-1)} (1-r^2)^{\frac{1}{2}(n-4)} \times \\ \times (1-\rho r)^{\frac{3}{2}-n} F\left(\frac{1}{2}, \frac{1}{2}, n-\frac{1}{2}, \frac{1}{2}(1+\rho r)\right). \quad (16.66)$$

Ряд (16.66) быстро сходится даже для небольших n . Для больших же значений n часто достаточно ограничиться только первым членом. Ошибка от использования для аппроксимации лишь k членов ряда значительно меньше, чем $2/(1-\rho r)$ раз взятый k -й член.

16.31 Почленное интегрирование равномерно сходящегося ряда (16.66) дает нам выражение для функции распределения r . При этом удобно $1+\rho r$ представить в виде $2-(1-\rho r)$ и воспользоваться в (16.66) биномиальным разложением по степеням $(1-\rho r)$. Для $r > \rho$ Хотеллинг (1953) получил следующий результат:

$$1-F(r) = \frac{n-2}{(n-1) \sqrt{2} B\left(\frac{1}{2}, n-\frac{1}{2}\right)} \left\{ M_0 + \frac{2M_0-M_1}{4(2n-1)} + \right. \\ \left. + \frac{9(4M_0-4M_1+M_2)}{32(2n-1)(2n+1)} + \dots \right\}, \quad (16.67)$$

где

$$M_k = \int_r^1 (1-\rho^2)^{\frac{1}{2}(n-1)} (1-x^2)^{\frac{1}{2}(n-4)} (1-\rho x)^{\frac{3}{2}+k-n} dx, \\ k=0, 1, 2, \dots,$$

и асимптотическую формулу ($r > \rho$)

$$1-F(r) \sim \frac{n-2}{(n-1)^2 \sqrt{2} B\left(\frac{1}{2}, n-\frac{1}{2}\right)} \times \\ \times \frac{(1-\rho^2)^{\frac{1}{2}(n-1)} (1-r^2)^{\frac{1}{2}(n-2)}}{(r-\rho)(1-\rho r)^{n-\frac{5}{2}}} \{1 + O(n^{-1})\}. \quad (16.68)$$

Функции плотности и распределения коэффициента корреляции не выражаются в элементарных функциях. Однако ф. п. и ф. р. были табулированы Дэвид (1938) для значений $n=3(1)25, 50, 100, 200$ и 400 ; $\rho=0,0(0,1)0,9$ и для $r=-1,00(0,05)+1,00$ (местами — с более мелким шагом).

16.32 Моменты этого распределения также выражаются в терминах гипергеометрических функций. Возвратимся к (16.59)

и положим

$$\alpha_1^2 = \frac{\sigma_1^2(1-\rho^2)}{n}, \quad \alpha_2^2 = \frac{\sigma_2^2(1-\rho^2)}{n}.$$

После несложных преобразований вместо (16.59) получим

$$dF = \frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi\Gamma(n-2)} \exp\left\{-\frac{1}{2}\frac{s_1^2}{\alpha_1^2} - \frac{1}{2}\frac{s_2^2}{\alpha_2^2} + \rho r \frac{s_1 s_2}{\alpha_1 \alpha_2}\right\} \times \\ \times \left(\frac{s_1}{\alpha_1}\right)^{n-2} \left(\frac{s_2}{\alpha_2}\right)^{n-2} (1-r^2)^{\frac{1}{2}(n-4)} \left(\frac{ds_1}{\alpha_1}\right) \left(\frac{ds_2}{\alpha_2}\right) dr. \quad (16.69)$$

Положим $u_1 = \frac{s_1}{\alpha_1}$, $u_2 = \frac{s_2}{\alpha_2}$ и разложим функцию $\exp\left(\rho r \frac{s_1 s_2}{\alpha_1 \alpha_2}\right)$, входящую в (16.69). Получим

$$dF = \frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi\Gamma(n-2)} \exp\left(-\frac{1}{2}u_1^2 - \frac{1}{2}u_2^2\right) u_1^{n-2} u_2^{n-2} (1-r^2)^{\frac{1}{2}(n-4)} \times \\ \times \sum_{j=0}^{\infty} \frac{(\rho r u_1 u_2)^j}{j!} du_1 du_2 dr. \quad (16.70)$$

Интегрируя в (16.70) по u_2 от 0 до ∞ , найдем элемент распределения u_1 и r

$$dF = \frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi\Gamma(n-2)} \exp\left(-\frac{1}{2}u_1^2\right) u_1^{n-2} (1-r^2)^{\frac{1}{2}(n-4)} \times \\ \times \sum \frac{(\rho r u_1)^j}{j!} \Gamma\left(\frac{n-1+j}{2}\right) 2^{\frac{1}{2}(n+j-3)} du_1 dr. \quad (16.71)$$

Умножая правую часть в (16.71) на r и интегрируя от -1 до 1 , получаем

$$\frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi\Gamma(n-2)} \exp\left(-\frac{1}{2}u_1^2\right) u_1^{n-2} \sum_{j=0}^{\infty} \frac{(\rho u_1)^{2j+1}}{(2j+1)!} \Gamma\left(\frac{n+2j}{2}\right) \times \\ \times B\left(\frac{n-2}{2}, \frac{2j+3}{2}\right) \cdot 2^{\frac{1}{2}(n+2j-2)} du_1.$$

И, наконец, интегрируя по u_1 , находим

$$\mu'_1(r) = \frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\pi\Gamma(n-2)} \sum_{j=0}^{\infty} \frac{\rho^{2j+1}}{(2j+1)!} \Gamma\left(\frac{n+2j}{2}\right) B\left(\frac{n-2}{2}, \frac{2j+3}{2}\right) \times \\ \times \Gamma\left(\frac{n+2j}{2}\right) \cdot 2^{n+2j-2}. \quad (16.72)$$

Выражая В-функции в терминах Γ -функций и учитывая, что

$$\Gamma(x) \Gamma\left(x + \frac{1}{2}\right) = \frac{\pi^{1/2}}{2^{2x-1}} \Gamma(2x),$$

найдем

$$\begin{aligned} \mu'_1(r) &= \frac{\rho(1-\rho^2)^{\frac{1}{2}(n-1)} \Gamma^2\left(\frac{1}{2}n\right)}{\Gamma\left\{\frac{1}{2}(n-1)\right\} \Gamma\left\{\frac{1}{2}(n+1)\right\}} \left\{ 1 + \frac{\frac{1}{2}n \cdot \frac{1}{2}n}{\frac{1}{2}(n+1)} \frac{\rho^2}{1!} + \right. \\ &\quad \left. + \frac{\frac{1}{2}n\left(\frac{1}{2}n+1\right) \cdot \frac{1}{2}n\left(\frac{1}{2}n+1\right)}{\frac{1}{2}(n+1) \cdot \frac{1}{2}(n+3)} \frac{\rho^4}{2!} + \dots \right\} = \\ &= \frac{\rho(1-\rho^2)^{\frac{1}{2}(n-1)} \Gamma^2\left(\frac{1}{2}n\right)}{\Gamma\left\{\frac{1}{2}(n-1)\right\} \Gamma\left\{\frac{1}{2}(n+1)\right\}} F\left(\frac{1}{2}n, \frac{1}{2}n, \frac{1}{2}(n+1), \rho^2\right), \end{aligned}$$

и поскольку

$$F(\alpha, \beta, \gamma, x) = (1-x)^{\gamma-\alpha-\beta} F(\gamma-\alpha, \gamma-\beta, \gamma, x),$$

то

$$\mu'_1(r) = \frac{\rho \Gamma^2\left(\frac{1}{2}n\right)}{\Gamma\left\{\frac{1}{2}(n-1)\right\} \Gamma\left\{\frac{1}{2}(n+1)\right\}} F\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}(n+1), \rho^2\right). \quad (16.73)$$

Аналогичным путем (ср. Хотеллинг, 1953) можно установить, что

$$\left. \begin{aligned} \mu'_2(r) &= 1 - (1-\rho^2) \frac{n-2}{n-1} F\left(1, 1, \frac{1}{2}(n+1), \rho^2\right) \\ \text{и} \\ \gamma_1 &= -\frac{6\rho}{n^{1/2}} + o(n^{-1/2}), \quad \gamma_2 = \frac{6(12\rho^2-1)}{n} + o(n^{-1}). \end{aligned} \right\} \quad (16.74)$$

Поэтому асимметрия возрастает с ростом $|\rho|$ и убывает только как $n^{-1/2}$.

Фишеровское преобразование коэффициента корреляции r

16.33 Фишером (1921) было найдено замечательное преобразование выборочного коэффициента корреляции r , распределение которого приближается к нормальному намного быстрее, чем распределение самого r , и дисперсия преобразованной величины

почти не зависит от ρ . Полагая

$$\left. \begin{aligned} r &= \operatorname{th} z, \quad z = \frac{1}{2} \log \frac{1+r}{1-r}, \\ \rho &= \operatorname{th} \zeta, \quad \zeta = \frac{1}{2} \log \frac{1+\rho}{1-\rho}, \end{aligned} \right\} \quad (16.75)$$

мы можем разложить плотность случайной величины r по степеням $x = z - \zeta$ и обратным степеням n . Фишер получил следующее разложение:

$$\begin{aligned} f = \frac{n-2}{\sqrt{2\pi(n-1)}} e^{-\frac{1}{2}(n-1)x^2} & \left\{ 1 + \frac{1}{2} \rho x + \left(\frac{2+\rho}{8(n-1)} + \frac{4-\rho^2}{8} x^2 + \right. \right. \\ & + \frac{n-1}{12} x^4) + \rho x \left(\frac{4-\rho^2}{16(n-1)} + \frac{4+3\rho^2}{48} x^2 + \frac{n-1}{24} x^4 \right) + \\ & + \left(\frac{4+12\rho^2+9\rho^4}{128(n-1)^2} + \frac{8-2\rho^2+3\rho^4}{64(n-1)} x^2 + \frac{8+4\rho^2-5\rho^4}{128} x^4 + \right. \\ & \left. \left. + \frac{28-15\rho^2}{1440} x^6(n-1) \right) + \frac{(n-1)^2}{288} x^8 + \dots \right\}. \quad (16.76) \end{aligned}$$

Вычисляя момент относительно $x=0$, находим *)

$$\left. \begin{aligned} \mu'_1 &= \frac{\rho}{2(n-1)} \left\{ 1 + \frac{5+\rho^2}{4(n-1)} + \frac{11+2\rho^2+3\rho^4}{8(n-1)^2} + \dots \right\}, \\ \mu'_2 &= \frac{1}{n-1} \left\{ 1 + \frac{4-\rho^2}{2(n-1)} + \frac{22-6\rho^2-3\rho^4}{6(n-1)^2} + \dots \right\}, \\ \mu'_3 &= \frac{\rho^3}{(n-1)^3} + \dots, \\ \mu'_4 &= \frac{1}{(n-1)^2} \left\{ 3 + \frac{14-3\rho^2}{n-1} + \frac{184-48\rho^2-21\rho^4}{4(n-1)^2} + \dots \right\}, \end{aligned} \right\} \quad (16.77)$$

и, следовательно,

$$\left. \begin{aligned} \gamma_1 &= \frac{\rho^3}{(n-1)^{3/2}} + \dots, \\ \gamma_2 &= \frac{2}{n-1} + \frac{4+2\rho^2-3\rho^4}{(n-1)^2} + \dots \end{aligned} \right\} \quad (16.78)$$

Поэтому дисперсия случайной величины $z - \zeta$ почти не зависит от ρ , а γ_1 убывает как $n^{-3/2}$. Тем самым мы можем считать распределение $z - \zeta$ приближенно нормальным со средним и дисперсией, указанными в (16.77). В качестве грубой

*) Уравнения (16.77) в том виде, в котором они содержались в статье Фишера, имели некоторые ошибки, исправленные Гейеном (1950). Однако на основной результат (16.80) они влияния не оказывают.

аппроксимации можно взять

$$\begin{aligned}\mu'_1(z - \zeta) &= \frac{\rho}{2(n-1)}, \\ D(z - \zeta) &= \frac{1}{n-1} + \frac{4-\rho^2}{2(n-1)^2}.\end{aligned}\quad (16.79)$$

При малых ρ

$$D(z - \zeta) \approx \frac{1}{n-1} + \frac{2}{(n-1)^2} \approx \frac{1}{n-3}. \quad (16.80)$$

Когда n умеренно, мы можем брать еще более грубую аппроксимацию, предполагая $z - \zeta$ нормально распределенной с нулевым средним и дисперсией $\frac{1}{n-3}$. Некоторое сравнение различных аппроксимаций дано во введении к таблицам Дэвид (1938). Оказывается, что для $n > 50$ формулы (16.79) и (16.80) адекватны. Аппроксимация, задаваемая формулами (16.77), удовлетворительна уже при $n=11$.

Хотеллинг (1953) изучал возможность улучшения z -преобразования. Он нашел, что случайная величина

$$z^* = z - (3z + r)/(4n) \quad (16.81)$$

имеет моменты

$$\left. \begin{aligned}\mu'_1 &= \left(\zeta - \frac{3\zeta + \rho}{4n} \right) + \frac{\rho}{2(n-1)} + \frac{3\rho}{8(n-1)^2} + O(n^{-3}), \\ \mu_2 &= \frac{1}{n-1} + O(n^{-3})\end{aligned} \right\} \quad (16.82)$$

и что ее распределение более близко к нормальному, чем распределение z . Изучались также и другие улучшения.

16.34 Мы заканчиваем обсуждение вопросов, связанных с распределением r -статистики в нормальном случае. Но прежде чем расстаться с двумерным нормальным распределением, мы рассмотрим одну статистику, тесно связанную с r . По аналогии с коэффициентом корреляции ρ генеральной совокупности выборочный коэффициент корреляции r определялся как отношение выборочной ковариации к произведению выборочных стандартных отклонений. Рассмотрим теперь статистики (см. (16.48))

$$\left. \begin{aligned}b_1 &= \frac{1}{n} \sum (x - \bar{x})(y - \bar{y})/s_y^2, \\ b_2 &= \frac{1}{n} \sum (x - \bar{x})(y - \bar{y})/s_x^2,\end{aligned} \right\} \quad (16.83)$$

аналогами которых в генеральной совокупности являются отношения $\beta_1 = \mu_{11}/\sigma_2^2$ и $\beta_2 = \mu_{11}/\sigma_1^2$ соответственно. Статистики b_1 и b_2 называются выборочными коэффициентами регрессии x на y

и y на x соответственно. Их аналоги β_1 и β_2 в генеральной совокупности называются коэффициентами регрессии.

Обращаясь к (16.48), мы видим, что

$$M(x|y) = \beta_1 y, \quad M(y|x) = \beta_2 x.$$

Далее мы рассматриваем распределение статистики b_2 . Результаты для b_1 получаются соответствующей перестановкой индексов.

Распределение коэффициентов регрессии в нормальном случае

16.35 Обращаясь к уравнению (16.59) и полагая $b_2 = \frac{rs_2}{s_1}$, мы получим для элемента совместного распределения s_1 , s_2 и b_2 формулу

$$dF \propto \exp \left[-\frac{n}{2(1-\rho^2)} \left\{ \frac{s_1^2}{\sigma_1^2} - \frac{2\rho s_1^2 b_2}{\sigma_1 \sigma_2} + \frac{s_2^2}{\sigma_2^2} \right\} \right] s_1^{n-1} s_2^{n-3} \times \\ \times \left(1 - \frac{s_1^2 b_2^2}{s_2^2} \right)^{\frac{1}{2}(n-4)} ds_1 ds_2 db_2. \quad (16.84)$$

Интегрирование (16.84) по s_2 дает элемент распределения s_1 и b_2 :

$$dF \propto \exp \left[-\frac{n}{2(1-\rho^2)} \left\{ \frac{s_1^2}{\sigma_1^2} - \frac{2\rho s_1^2 b_2}{\sigma_1 \sigma_2} + \frac{s_1^2 b_2^2}{\sigma_2^2} \right\} \right] s_1^{n-1} ds_1 db_2. \quad (16.85)$$

Дальнейшее интегрирование (16.85) по s_1 приводит к элементу распределения b_2 :

$$dF \propto \frac{db_2}{\left(1 - \frac{2\rho\sigma_1}{\sigma_2} b_2 + \frac{\sigma_1^2}{\sigma_2^2} b_2^2 \right)^{\frac{1}{2}n}},$$

или, с учетом нормировки,

$$dF = \frac{\Gamma\left(\frac{1}{2}n\right) \sigma_2^{n-1} (1-\rho^2)^{\frac{1}{2}(n-1)}}{\sqrt{\pi} \Gamma\left(\frac{n-1}{2}\right) \sigma_1^{n-1}} \frac{db_2}{\left\{ \frac{\sigma_2^2}{\sigma_1^2} (1-\rho^2) + \left(b_2 - \frac{\rho\sigma_2}{\sigma_1} \right)^2 \right\}^{\frac{1}{2}n}}. \quad (16.86)$$

Распределение (16.86) является пирсоновским распределением VII типа, симметричным относительно точки $\beta_2 = \frac{\rho\sigma_2}{\sigma_1}$, т. е.

относительно коэффициента регрессии генеральной совокупности. Это распределение стремится к нормальному довольно быстро, и поэтому при малых значениях n использование стандартных ошибок для регрессии более оправдано, чем для коэффициента корреляции. Для малых выборок, однако, распределение (16.86) используется редко, поскольку оно зависит от неизвестных величин σ_1 , σ_2 и ρ , т. е. от истинных дисперсий и коэффициента корреляции.

16.36 Можно найти статистики, отличные от b_1 и b_2 , которые могут использоваться при построении критериев для регрессии. Рассмотрим

$$u = \frac{(b_2 - \beta_2) s_1}{(s_2^2 - b_2^2 s_1^2)^{1/2}} \quad (16.87)$$

и обратимся теперь к распределению величин a , b , c , заданных формулами (16.56), а именно (см. (16.57)):

$$dF \propto \exp[-a + 2\rho b - c](ac - b^2)^{\frac{1}{2}(n-4)} da db dc. \quad (16.88)$$

Так как согласно (16.87)

$$u = \frac{b - \rho a}{\sqrt{ac - b^2}},$$

то, заменяя в (16.88) c на u , по этой формуле после простых преобразований найдем

$$dF \propto \frac{\exp\left[-a\left(1 + \frac{\rho^2}{u^2}\right)\right] da du}{au^{n-1}} \exp\left[\frac{u^2 + 1}{u^2}\left(2\rho b - \frac{b^2}{a}\right)\right] (b - \rho a)^{n-2} db. \quad (16.89)$$

Интегрирование второго множителя в (16.89) по b дает выражение, пропорциональное

$$\exp\left\{\frac{\rho^2(u^2 + 1)}{u^2}\right\} \left(\frac{au^2}{u^2 + 1}\right)^{\frac{1}{2}n-1}, \quad (16.90)$$

и, следовательно, для элемента распределения a и u из (16.89) и (16.90) получим формулу

$$dF \propto \frac{a^{\frac{1}{2}(n-1)} \exp(-a + \rho^2 a) da}{a} \frac{du}{(1 + u^2)^{\frac{1}{2}(n-1)}}. \quad (16.91)$$

Отсюда следует, что a и u независимы и что элемент распределения u имеет вид

$$dF \propto \frac{du}{(1 + u^2)^{\frac{1}{2}(n-1)}}. \quad (16.92)$$

Это распределение не содержит каких бы то ни было параметров исходного распределения. Случайная величина

$$t = u \sqrt{n-2} = \frac{(b_2 - \beta_2) s_1 \sqrt{n-2}}{(s_2^2 - b_2^2 s_1^2)^{1/2}} \quad (16.93)$$

подчиняется распределению Стьюдента (16.15) с $\nu = n - 2$ степенями свободы.

16.37 В последующих главах (второй том) будут обсуждаться приложения распределений r и b , выведенных формально в предшествующих пунктах, и будут даны также обобщения наших результатов на случай многомерных нормальных распределений.

УПРАЖНЕНИЯ

16.1 Показать, что моменты распределения дисперсионного отношения (16.24) выражаются формулой

$$\mu'_r = \left(\frac{v_2}{v_1}\right)^r \frac{\Gamma\left(\frac{1}{2}v_1 + r\right) \Gamma\left(\frac{1}{2}v_2 - r\right)}{\Gamma\left(\frac{1}{2}v_1\right) \Gamma\left(\frac{1}{2}v_2\right)},$$

если $2r < v_2$, и не существуют, когда $2r \geq v_2$.

16.2 Пусть $u_1, \dots, u_n$ независимы и распределены как χ^2 . Показать, воспользовавшись упражнением 11.15, что $\sum_{i=1}^n a_i u_i$, где $a_i = \pm 1$, распределена или

(а) как χ^2 , или (б) как величина, распределение которой выражается через функцию Бесселя, или как разность величин в (а) и (б).

16.3 Показать, что для χ^2 -распределения с ν степенями свободы центральные моменты удовлетворяют соотношению

$$\mu_{j+1} = 2j(\mu_j + \nu \mu_{j-1}), \quad j \geq 1,$$

и, таким образом, проверить формулу (16.5).

16.4 n независимых случайных величин $x_1, \dots, x_n$ равномерно распределены на отрезке $[0, 1]$. Показать, что величина $-2 \log(x_1 x_2 \dots x_n)$ распределена как χ^2 с $2n$ степенями свободы.

16.5 Пусть $X_i (i=1, \dots, n)$ — случайные величины с нулевыми средними, и пусть $\varepsilon_i (i=1, \dots, n)$ — нормально распределенные случайные величины с нулевым средним и дисперсией σ^2 . Обозначим

$$y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

и предположим, что в качестве оценок для β_0 и β_1 взяты величины b_0 и b_1 соответственно такие, что

$$b_0 = \bar{y}, \quad b_1 = \sum (y_i - \bar{y}) X_i / \sum X_i^2$$

Если $y'_i = y_i - (b_0 + b_1 X_i)$, то

$$\sum (y_i - \bar{y})^2 = \sum (y'_i)^2 + \frac{\left\{ \sum X_i (y_i - \bar{y}) \right\}^2}{\sum X_i^2}.$$

Отсюда вывести, что сумма $\sum (y'_i)^2$ распределена как $\sigma^2 \chi^2$ с $(n-2)$ степенями свободы и не зависит от $b_1^2 \sum X_i^2$, которая, в свою очередь, распределена как $\sigma^2 \chi^2$ с одной степенью свободы.

Если x и y — независимые нормально распределенные величины, то показать, что отношение

$$(n-2)r^2/(1-r^2)$$

распределено как студентовское t^2 с $(n-2)$ степенями свободы.

16.6 Две независимые машины непрерывно производят проволоку. Пусть λ_1 и λ_2 — средние значения количества дефектов на единицу длины для первой и второй машин соответственно. Показать, что если t_i — интервал (по времени) между $(i-1)$ -м и i -м дефектами во время непрерывной инспекции обеих машин, то он распределен по закону $(\lambda_1 + \lambda_2) \exp\{-(\lambda_1 + \lambda_2)t_i\}$, и, таким

образом, если $T = \sum_{i=1}^n t_i$, то $2(\lambda_1 + \lambda_2)T$ имеет χ^2 -распределение с $2n$ степенями свободы.

16.7 Показать, что функция распределения (16.2) при нечетных v представима в виде

$$1 - F(\chi^2) = \left(\frac{2}{\pi}\right)^{1/2} \left\{ \int_{\chi}^{\infty} \exp\left(-\frac{1}{2}y^2\right) dy + \exp\left(-\frac{1}{2}\chi^2\right) \sum_{s=1}^{\frac{1}{2}(v-1)} \frac{\chi^{2s-1} \cdot 2^{s-1} (s-1)!}{(2s-1)!} \right\},$$

а для четных v — в виде

$$1 - F(\chi^2) = \exp\left(-\frac{1}{2}\chi^2\right) \sum_{s=0}^{\frac{1}{2}(v-2)} \left(\frac{1}{2}\chi^2\right)^s / s!.$$

Отсюда установить, что в последнем случае функция распределения χ^2 может быть использована для вычисления функции пуассоновского распределения.

16.8 Методами 16.7 показать, что величина

$$\left[\left(\frac{13\chi^2 - v}{12v} \right)^{3/13} + \frac{5}{18v} \left(1 + \frac{7}{48v} \right) - 1 \right] \frac{6v\sqrt{2}}{5} \left(1 - \frac{1}{18v} \right)$$

приближенно нормальна с нулевым средним и единичной дисперсией.

(Холдейн, 1937.)

16.9 Непосредственно из (16.15) показать, что t -распределение Стьюдента стремится к нормальному при $v \rightarrow \infty$.

16.10 Полагая $q = (1 + t^2/v)^{-1}$, показать, что функция распределения $F(t)$ студентовской t -статистики при четном v представляется в форме

$$2F(t) - 1 = (1 - q)^{1/2} \left\{ 1 + \sum_{s=1}^{\frac{1}{2}(v-2)} \frac{q^s}{2^{2s-1} s B(s, s)} \right\},$$

где в правой части выражение в фигурных скобках является частичной суммой биномиального разложения $(1 - q)^{-1/2}$.

(Фишер, 1935.)

16.11 Проверить для F -распределения справедливость формул (16.27) и (16.28).

16.12 Доказать, что х. ф. z -распределения имеет вид (16.29). Изучая члены низших порядков в (16.30), доказать, что с ростом числа степеней свободы распределение стремится к нормальному.

16.13 Используя методы главы 10, показать, что для выборок большого объема из двумерной нормальной совокупности дисперсия выборочного коэффициента регрессии приближенно выражается формулой

$$D\left(\frac{m_{11}}{m_{20}}\right) = \frac{\sigma_2^2}{\sigma_1^2} \frac{(1 - \rho^2)}{n}.$$

16.14 Дифференцируя (16.64) по (ρr) , установить, что

$$\{1 - (\rho r)^2\} I_2 = 1 + \rho r I_1$$

и что

$$n \{1 - (\rho r)^2\} I_{n+1} - (2n - 1) I_n - (n - 1) I_{n-1} = 0.$$

Отсюда вывести, что плотность распределения f статистики r удовлетворяет рекуррентной формуле

$$(n - 1) \{1 - (\rho r)^2\} f_{n+1} = (2n - 1) \rho r (1 - \rho^2)^{\frac{1}{2}} (1 - r^2)^{\frac{1}{2}} f_n + \\ + (n - 1)^2 (1 - \rho^2) (1 - r^2) f_{n-1} / (n - 2).$$

(Co-operative Study, 1917; Хотеллинг, 1953.)

16.15 Доказать, что плотность распределения статистики r удовлетворяет уравнению

$$r \frac{\partial f_n}{\partial r} + \frac{(n - 3) r^2 f_n}{1 - r^2} = \rho \frac{\partial f_n}{\partial \rho} + \frac{n \rho^2 f_n}{1 - \rho^2}.$$

(Хотеллинг, 1953.)

16.16 Из (16.60) вывести, что элемент совместного распределения статистик $v = \frac{s_1}{\sigma_1} / \frac{s_2}{\sigma_2}$ и r выражается формулой

$$dF \propto \frac{v^{n-2} (1 - r^2)^{\frac{1}{2}(n-4)}}{(1 - 2\rho r v + v^2)^{n-1}} dr dv.$$

Интегрируя по r от -1 до $+1$, с помощью замены

$$r = \frac{u(\lambda + \mu) - (1 - u)(\lambda - \mu)}{u(\lambda + \mu) + (1 - u)(\lambda - \mu)}, \quad \lambda = 1 + v^2, \quad \mu = 2\rho v,$$

показать, что элемент распределения v имеет вид

$$dF = \frac{2(1-\rho^2)^{\frac{1}{2}(n-1)}}{B\left(\frac{n-1}{2}, \frac{n-1}{2}\right)} \frac{v^{n-2}}{(1+v^2)^{n-1}} \left\{1 - \frac{4\rho^2 v^2}{(1+v^2)^2}\right\}^{-\frac{1}{2}n}.$$

(Бозе, 1935; Финни, 1938.)

16.17 Пусть x и y — независимые случайные величины с непрерывными распределениями. Показать, что для выборки объема n из генеральной совокупности с двумерным распределением, отвечающим этим случайным величинам, выборочный коэффициент корреляции r имеет нулевое среднее и дисперсию $(n-1)^{-1}$.

(Питмэн, 1937.)

16.18 Пусть t — статистика, которая при увеличении объема выборки сходится по вероятности к некоторому параметру генеральной совокупности θ , и дисперсия статистики t есть некоторая функция от θ , скажем $f(\theta)/n + o(n^{-1})$. Показать, что с точностью до величин порядка n^{-1} случайная величина

$$z = \int_0^t \{f(\theta)\}^{-1/2} d\theta$$

имеет дисперсию $n^{-1} + o(n^{-1})$.

Отсюда следует, что $z = \operatorname{argth} r$ имеет дисперсию $n^{-1} + o(n^{-1})$, так как для больших выборок

$$Dr = (1 - \rho^2)^2/n.$$

16.19 Замечая, что согласно (16.77)

$$Dz = \frac{1}{n} \left(1 + \frac{4 - \rho^2}{2n}\right) + o(n^{-2}),$$

показать, что статистика

$$z^* = \int_0^r \frac{dr}{(1-r^2) \left(1 + \frac{4-r^2}{2n}\right)^{1/2}} = z - \frac{3z+r}{4n} + o(n^{-1})$$

имеет дисперсию

$$Dz^* = \frac{1}{n-1} + o(n^{-2}).$$

(Хотеллинг, 1953, Кендалл, 1953.)

16.20 Пусть $x_1, \dots, x_n$ — случайная выборка из нормальной генеральной совокупности. Образует разности $\bar{x}_i - \bar{x}$ и выберем случайным образом одну из этих разностей, скажем $x_h - \bar{x}$. Найти распределение для $(x_h - \bar{x})/s$, где $s = \sum (x - \bar{x})^2/(n-1)$, и показать, что, если x_j, x_h выбраны случайно, то отношение $(x_j - x_h)/s$ распределено так же, как и

$$\frac{t \{2(n-1)\}}{\{t^2 + (n-2)\}^{1/2}},$$

где t является t -статистикой Стьюдента.

(Дэвид и др., 1954.)

16.21 Показать, что в случае $\rho=0$ точное распределение для z -статистики Фишера (см. (16.75)) совпадает с z -распределением дисперсионного отношения (16.26) при $\nu_1=\nu_2=n-2$. Используя х. ф. (16.29) и соотношение

$$\frac{\Gamma''(p)}{\Gamma(p)} - \left\{ \frac{\Gamma'(p)}{\Gamma(p)} \right\}^2 = \frac{\pi^2}{6} - \sum_{r=1}^{p-1} \frac{1}{r^2},$$

справедливое при целых p , показать, что для четных $n \geq 4$ дисперсия z в точности равна

$$\mu_2 = \frac{\pi^2}{3} - \frac{1}{2} \sum_{r=1}^{\frac{1}{2}n-2} \frac{1}{r^2}.$$

При $n=6$ отсюда следует, что $\mu_2=0,32$, в то время как согласно приближенной формуле (16.80) $\mu_2=0,33$.

16.22 В случае $\rho=0$ установить, используя (16.59), что распределение rs_i/σ_i нормально со средним нуль и дисперсией $1/n$ ($i=1, 2$).

ПРИЛОЖЕНИЕ

Таблица 1

Функция плотности нормального распределения

$$y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad \text{с первыми и вторыми разностями}$$

x	y	$\Delta^1 (-)$	Δ^2	x	y	$\Delta^1 (-)$	Δ^2
0,0	0,39894	199	-392	2,5	0,01753	395	+79
0,1	0,39695	591	-374	2,6	0,01358	316	+66
0,2	0,39104	965	-347	2,7	0,01042	250	+53
0,3	0,38139	1312	-308	2,8	0,00792	197	+45
0,4	0,36827	1620	-265	2,9	0,00595	152	+36
0,5	0,35207	1885	-212	3,0	0,00443	116	+27
0,6	0,33322	2097	-159	3,1	0,00327	89	+23
0,7	0,31225	2256	-104	3,2	0,00238	66	+17
0,8	0,28969	2360	-52	3,3	0,00172	49	+13
0,9	0,26609	2412	-0	3,4	0,00123	36	+10
1,0	0,24197	2412	+46	3,5	0,00087	26	+7
1,1	0,21785	2366	+84	3,6	0,00061	19	+6
1,2	0,19419	2282	+118	3,7	0,00042	13	+4
1,3	0,17137	2164	+143	3,8	0,00029	9	+2
1,4	0,14973	2021	+161	3,9	0,00020	7	+3
1,5	0,12952	1860	+173	4,0	0,00013	4	—
1,6	0,11092	1687	+177	4,1	0,00009	3	—
1,7	0,09405	1510	+177	4,2	0,00006	2	—
1,8	0,07895	1333	+170	4,3	0,00004	2	—
1,9	0,06562	1163	+162	4,4	0,00002	—	—
2,0	0,05399	1001	+150	4,5	0,00002	—	—
2,1	0,04398	851	+137	4,6	0,00001	—	—
2,2	0,03547	714	+120	4,7	0,00001	—	—
2,3	0,02833	594	+108	4,8	0,00000	—	—
2,4	0,02239	486	+91				

Таблица 2

Функция нормального распределения

В таблице даны значения $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}x^2} dx$, т. е. значения площади

под кривой $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, лежащей левее точки x ; например, для $x = 1,86 (= 1,5 + 0,36) \quad F(1,86) = 0,9686$

x	0,0+	0,5+	1,0+	1,5+	2,0+	2,5+	3,0+	3,5+
0,00	5000	6915	8413	9332	9772	98379	98665	9877
0,01	5040	6950	8438	9345	9778	98396	98669	9878

x	0,0+	0,5+	1,0+	1,5+	2,0+	2,5+	3,0+	3,5+
0,02	5080	6985	8461	9357	9783	9 ² 413	9 ² 874	9 ³ 78
0,03	5120	7019	8485	9370	9788	9 ² 430	9 ² 878	9 ³ 79
0,04	5160	7054	8508	9382	9793	9 ² 446	9 ² 882	9 ³ 80
0,05	5199	7088	8531	9394	9798	9 ² 461	9 ² 886	9 ³ 81
0,06	5239	7123	8554	9406	9803	9 ² 477	9 ² 889	9 ³ 81
0,07	5279	7157	8577	9418	9808	9 ² 492	9 ² 893	9 ³ 82
0,08	5319	7190	8599	9429	9812	9 ² 506	9 ² 897	9 ³ 83
0,09	5359	7224	8621	9441	9817	9 ² 520	9 ² 900	9 ³ 83
0,10	5398	7257	8643	9452	9821	9 ² 534	9 ³ 03	9 ³ 84
0,11	5438	7291	8665	9463	9826	9 ² 547	9 ³ 06	9 ³ 85
0,12	5478	7324	8686	9474	9830	9 ² 560	9 ³ 10	9 ³ 85
0,13	5517	7357	8708	9484	9834	9 ² 573	9 ³ 13	9 ³ 86
0,14	5557	7389	8729	9495	9838	9 ² 585	9 ³ 16	9 ³ 86
0,15	5596	7422	8749	9505	9842	9 ² 598	9 ³ 18	9 ³ 87
0,16	5636	7454	8770	9515	9846	9 ² 609	9 ³ 21	9 ³ 87
0,17	5675	7486	8790	9525	9850	9 ² 621	9 ³ 24	9 ³ 88
0,18	5714	7517	8810	9535	9854	9 ² 632	9 ³ 26	9 ³ 88
0,19	5753	7549	8830	9545	9857	9 ² 643	9 ³ 29	9 ³ 89
0,20	5793	7580	8849	9554	9861	9 ² 653	9 ³ 31	9 ³ 89
0,21	5832	7611	8869	9564	9864	9 ² 664	9 ³ 34	9 ³ 90
0,22	5871	7642	8888	9573	9868	9 ² 674	9 ³ 36	9 ³ 90
0,23	5910	7673	8907	9582	9871	9 ² 683	9 ³ 38	9 ⁴ 04
0,24	5948	7704	8925	9591	9875	9 ² 693	9 ³ 40	9 ⁴ 08
0,25	5987	7738	8944	9599	9878	9 ² 702	9 ³ 42	9 ⁴ 12
0,26	6026	7764	8962	9608	9881	9 ² 711	9 ³ 44	9 ⁴ 15
0,27	6064	7794	8980	9616	9884	9 ² 720	9 ³ 46	9 ⁴ 18
0,28	6103	7823	8997	9625	9887	9 ² 728	9 ³ 48	9 ⁴ 22
0,29	6141	7852	9015	9633	9890	9 ² 736	9 ³ 50	9 ⁴ 25
0,30	6179	7881	9032	9641	9893	9 ² 744	9 ³ 52	9 ⁴ 28
0,31	6217	7910	9049	9649	9896	9 ² 752	9 ³ 53	9 ⁴ 31
0,32	6255	7939	9066	9656	9898	9 ² 760	9 ³ 55	9 ⁴ 33
0,33	6293	7967	9082	9664	9901	9 ² 767	9 ³ 57	9 ⁴ 36
0,34	6331	7995	9099	9671	9904	9 ² 774	9 ³ 58	9 ⁴ 39
0,35	6368	8023	9115	9678	9906	9 ² 781	9 ³ 60	9 ⁴ 41
0,36	6406	8051	9131	9686	9909	9 ² 788	9 ³ 61	9 ⁴ 43
0,37	6443	8078	9147	9693	9911	9 ² 795	9 ³ 62	9 ⁴ 46
0,38	6480	8106	9162	9699	9913	9 ² 801	9 ³ 64	9 ⁴ 48
0,39	6517	8133	9177	9706	9916	9 ² 807	9 ³ 65	9 ⁴ 50
0,40	6554	8159	9192	9713	9918	9 ² 813	9 ³ 66	9 ⁴ 52
0,41	6591	8186	9207	9719	9920	9 ² 819	9 ³ 68	9 ⁴ 54
0,42	6628	8212	9222	9726	9922	9 ² 825	9 ³ 69	9 ⁴ 56
0,43	6664	8238	9236	9732	9925	9 ² 831	9 ³ 70	9 ⁴ 58
0,44	6700	8264	9251	9738	9927	9 ² 836	9 ³ 71	9 ⁴ 59
0,45	6736	8289	9265	9744	9929	9 ² 841	9 ³ 72	9 ⁴ 61
0,46	6772	8315	9279	9750	9931	9 ² 846	9 ³ 73	9 ⁴ 63
0,47	6808	8340	9292	9756	9932	9 ² 851	9 ³ 74	9 ⁴ 64
0,48	6844	8365	9306	9761	9934	9 ² 856	9 ³ 75	9 ⁴ 66
0,49	6879	8389	9319	9767	9936	9 ² 861	9 ³ 76	9 ⁴ 67

Примечание. Знаки целой части (т. е. 0) в таблице опущены. Повторяющиеся девятки заменены степенями; например, 9³71 означает 0,99971.

$P = 1 - F$	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
$\nu = 1$	0,0157	0,03628	0,07393	0,1158	0,1642	0,2148	0,2745	0,3445	0,4145	0,4945	0,5745	0,6545	0,7345
2	0,0201	0,0404	0,103	0,211	0,321	0,431	0,541	0,651	0,761	0,871	0,981	1,091	1,201
3	0,115	0,185	0,352	0,584	0,816	1,048	1,280	1,512	1,744	1,976	2,208	2,440	2,672
4	0,297	0,429	0,711	1,064	1,417	1,770	2,123	2,476	2,829	3,182	3,535	3,888	4,241
5	0,554	0,752	1,145	1,600	2,055	2,510	2,965	3,420	3,875	4,330	4,785	5,240	5,695
6	0,872	1,134	1,635	2,204	2,833	3,462	4,091	4,720	5,349	5,978	6,607	7,236	7,865
7	1,239	1,564	2,167	2,833	3,499	4,165	4,831	5,497	6,163	6,829	7,495	8,161	8,827
8	1,646	2,032	2,733	3,490	4,247	5,004	5,761	6,518	7,275	8,032	8,789	9,546	10,303
9	2,088	2,532	3,325	4,168	5,009	5,850	6,691	7,532	8,373	9,214	10,055	10,896	11,737
10	2,358	3,059	3,940	4,865	5,789	6,714	7,638	8,563	9,487	10,412	11,336	12,261	13,185
11	3,053	3,609	4,575	5,578	6,581	7,584	8,587	9,590	10,593	11,596	12,599	13,602	14,605
12	3,571	4,178	5,226	6,304	7,382	8,460	9,538	10,616	11,694	12,772	13,850	14,928	16,006
13	4,107	4,765	5,892	7,042	8,160	9,278	10,396	11,514	12,632	13,750	14,868	15,986	17,104
14	4,660	5,368	6,571	7,790	8,967	10,144	11,321	12,498	13,675	14,852	16,029	17,206	18,383
15	5,229	5,985	7,261	8,547	9,773	11,000	12,226	13,452	14,678	15,904	17,130	18,356	19,582
16	5,812	6,614	7,962	9,312	10,562	11,812	13,062	14,312	15,562	16,812	18,062	19,312	20,562
17	6,408	7,255	8,672	10,085	11,297	12,509	13,721	14,933	16,145	17,357	18,569	19,781	20,993
18	7,015	7,906	9,390	10,865	12,037	13,209	14,381	15,553	16,725	17,897	19,069	20,241	21,413
19	7,633	8,567	10,117	11,651	12,773	13,895	15,017	16,139	17,261	18,383	19,505	20,627	21,749
20	8,260	9,237	10,851	12,443	13,565	14,687	15,809	16,931	18,053	19,175	20,297	21,419	22,541
21	8,897	9,915	11,591	13,240	14,362	15,484	16,606	17,728	18,850	19,972	21,094	22,216	23,338
22	9,542	10,600	12,338	14,041	15,163	16,285	17,407	18,529	19,651	20,773	21,895	23,017	24,139
23	10,196	11,293	13,091	14,848	15,970	17,092	18,214	19,336	20,458	21,580	22,702	23,824	24,946
24	10,856	11,992	13,848	15,659	16,781	17,903	19,025	20,147	21,269	22,391	23,513	24,635	25,757
25	11,524	12,697	14,611	16,473	17,595	18,717	19,839	20,961	22,083	23,205	24,327	25,449	26,571
26	12,198	13,409	15,379	17,292	18,414	19,536	20,658	21,780	22,902	24,024	25,146	26,268	27,390
27	12,879	14,125	16,151	18,114	19,233	20,355	21,477	22,599	23,721	24,843	25,965	27,087	28,209
28	13,565	14,847	16,928	18,939	20,058	21,180	22,302	23,424	24,546	25,668	26,790	27,912	29,034
29	14,256	15,574	17,708	19,768	20,868	21,990	23,112	24,234	25,356	26,478	27,600	28,722	29,844
30	14,953	16,306	18,493	20,599	21,703	22,825	23,947	25,069	26,191	27,313	28,435	29,557	30,679

Примечание. Для значений ν , больших 30, величину $\sqrt{2\chi^2}$ можно считать распределенной нормально со средним $\sqrt{2\nu - 1}$ и единичной дисперсией.

Т а б л и ц а 4а

Функция χ^2 -распределения с одной степенью свободы
для $0 \leq \chi^2 \leq 1$ с шагом 0,01

χ^2	P	Δ	χ^2	P	Δ
0	1,00000	7966	0,50	0,47950	436
0,01	0,92034	3280	0,51	0,47514	430
0,02	0,88754	2505	0,52	0,47084	423
0,03	0,86249	2101	0,53	0,46661	418
0,04	0,84148	1842	0,54	0,46243	411
0,05	0,82306	1656	0,55	0,45832	406
0,06	0,80650	1516	0,56	0,45426	400
0,07	0,79134	1404	0,57	0,45026	395
0,08	0,77730	1312	0,58	0,44631	389
0,09	0,76418	1235	0,59	0,44242	384
0,10	0,75183	1169	0,60	0,43858	379
0,11	0,74014	1111	0,61	0,43479	374
0,12	0,72903	1060	0,62	0,43105	369
0,13	0,71843	1015	0,63	0,42736	365
0,14	0,70828	974	0,64	0,42371	360
0,15	0,69854	938	0,65	0,42011	355
0,16	0,68916	905	0,66	0,41656	351
0,17	0,68011	874	0,67	0,41305	346
0,18	0,67137	845	0,68	0,40959	343
0,19	0,66292	820	0,69	0,40616	338
0,20	0,65472	795	0,70	0,40278	334
0,21	0,64677	773	0,71	0,39944	330
0,22	0,63904	752	0,72	0,39614	326
0,23	0,63152	731	0,73	0,39288	322
0,24	0,62421	713	0,74	0,38966	318
0,25	0,61708	696	0,75	0,38648	315
0,26	0,61012	679	0,76	0,38333	311
0,27	0,60333	663	0,77	0,38022	308
0,28	0,59670	648	0,78	0,37714	304
0,29	0,59022	634	0,79	0,37410	301
0,30	0,58388	620	0,80	0,37109	297
0,31	0,57768	607	0,81	0,36812	294
0,32	0,57161	595	0,82	0,36518	291
0,33	0,56566	583	0,83	0,36227	287
0,34	0,55983	572	0,84	0,35940	285
0,35	0,55411	560	0,85	0,35655	281
0,36	0,54851	551	0,86	0,35374	278
0,37	0,54300	540	0,87	0,35096	276
0,38	0,53760	530	0,88	0,34820	272
0,39	0,53230	521	0,89	0,34548	270
0,40	0,52709	512	0,90	0,34278	267
0,41	0,52197	503	0,91	0,34011	264
0,42	0,51694	495	0,92	0,33747	261
0,43	0,51199	487	0,93	0,33486	258
0,44	0,50712	479	0,94	0,33228	256
0,45	0,50233	471	0,95	0,32972	253
0,46	0,49762	463	0,96	0,32719	251
0,47	0,49299	457	0,97	0,32468	248
0,48	0,48842	449	0,98	0,32220	246
0,49	0,48393	443	0,99	0,31974	243
0,50	0,47950	436	1,00	0,31731	241

Таблица 46

Функция χ^2 -распределения с одной степенью свободы
для $1 \leq \chi^2 \leq 10$ с шагом 0,1.

χ^2	P	Δ	χ^2	P	Δ
1,0	0,31731	2304	5,5	0,01902	106
1,1	0,29427	2095	5,6	0,01796	99
1,2	0,27332	1911	5,7	0,01697	94
1,3	0,25421	1749	5,8	0,01603	89
1,4	0,23672	1605	5,9	0,01514	83
1,5	0,22067	1477	6,0	0,01431	79
1,6	0,20590	1361	6,1	0,01352	74
1,7	0,19229	1258	6,2	0,01278	71
1,8	0,17971	1163	6,3	0,01207	66
1,9	0,16808	1078	6,4	0,01141	62
2,0	0,15730	1000	6,5	0,01079	59
2,1	0,14730	929	6,6	0,01020	56
2,2	0,13801	864	6,7	0,00964	52
2,3	0,12937	803	6,8	0,00912	50
2,4	0,12134	749	6,9	0,00862	47
2,5	0,11385	699	7,0	0,00815	44
2,6	0,10686	651	7,1	0,00771	42
2,7	0,10035	609	7,2	0,00729	39
2,8	0,09426	568	7,3	0,00690	38
2,9	0,08858	532	7,4	0,00652	35
3,0	0,08326	497	7,5	0,00617	33
3,1	0,07829	465	7,6	0,00584	32
3,2	0,07364	436	7,7	0,00552	30
3,3	0,06928	408	7,8	0,00522	28
3,4	0,06520	383	7,9	0,00494	26
3,5	0,06137	359	8,0	0,00468	25
3,6	0,05778	337	8,1	0,00443	24
3,7	0,05441	316	8,2	0,00419	23
3,8	0,05125	296	8,3	0,00393	21
3,9	0,04829	279	8,4	0,00375	20
4,0	0,04550	262	8,5	0,00355	19
4,1	0,04288	246	8,6	0,00336	18
4,2	0,04042	231	8,7	0,00318	17
4,3	0,03811	217	8,8	0,00301	16
4,4	0,03594	205	8,9	0,00285	15
4,5	0,03389	192	9,0	0,00270	14
4,6	0,03197	181	9,1	0,00256	14
4,7	0,03016	170	9,2	0,00242	13
4,8	0,02846	160	9,3	0,00229	12
4,9	0,02686	151	9,4	0,00217	12
5,0	0,02535	142	9,5	0,00205	10
5,1	0,02393	134	9,6	0,00195	11
5,2	0,02259	126	9,7	0,00184	10
5,3	0,02133	119	9,8	0,00174	9
5,4	0,02014	112	9,9	0,00165	8
5,5	0,01902	106	10,0	0,00157	8

Квантили t -распределения
(Займствованы из таблиц сэра Рональда Фишера и Ф. Изйтса: *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd Ltd., Edinburgh.)

$P=2(1-F)$	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
$v = 1$	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,599
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,893	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,795	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073

ПРИЛОЖЕНИЕ

16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,336	2,576	3,291

5 %-ные точки z-распределения

(точки в которых $\Phi. p. = 0,95$). (Занимствованы из таблицы VI книги сэра Рональда Фишера *Statistical Methods for Research Workers*, Oliver and Boyd Ltd., Edinburgh.)

ПРИЛОЖЕНИЕ

	Значения v_1														
	1	2	3	4	5	6	8	12	24	∞					
1	2,5421	2,6479	2,6870	2,7071	2,7194	2,7276	2,7380	2,7484	2,7588	2,7693					
2	1,4592	1,4722	1,4765	1,4787	1,4800	1,4808	1,4819	1,4830	1,4840	1,4851					
3	1,1577	1,1284	1,1137	1,1051	1,0994	1,0953	1,0899	1,0842	1,0781	1,0716					
4	1,0212	0,9690	0,9429	0,9272	0,9168	0,9093	0,8993	0,8885	0,8767	0,8639					
5	0,9441	0,8777	0,8441	0,8236	0,8097	0,7997	0,7862	0,7714	0,7550	0,7368					
6	0,8948	0,8188	0,7798	0,7558	0,7394	0,7274	0,7112	0,6931	0,6729	0,6499					
7	0,8606	0,7777	0,7347	0,7080	0,6896	0,6761	0,6576	0,6369	0,6134	0,5862					
8	0,8355	0,7475	0,7014	0,6725	0,6525	0,6378	0,6175	0,5945	0,5682	0,5371					
9	0,8163	0,7242	0,6757	0,6450	0,6238	0,6080	0,5862	0,5613	0,5324	0,4979					
10	0,8012	0,7058	0,6553	0,6232	0,6009	0,5843	0,5611	0,5346	0,5035	0,4657					
11	0,7889	0,6909	0,6387	0,6055	0,5822	0,5648	0,5406	0,5126	0,4795	0,4387					
12	0,7788	0,6786	0,6250	0,5907	0,5666	0,5487	0,5234	0,4941	0,4592	0,4156					
13	0,7703	0,6682	0,6134	0,5783	0,5535	0,5350	0,5089	0,4785	0,4419	0,3957					
14	0,7630	0,6594	0,6036	0,5677	0,5423	0,5233	0,4964	0,4649	0,4269	0,3782					
15	0,7568	0,6518	0,5950	0,5585	0,5326	0,5131	0,4855	0,4532	0,4138	0,3628					

16	0,7514	0,6451	0,5876	0,5505	0,5241	0,5042	0,4760	0,4428	0,4022	0,3490
17	0,7466	0,6393	0,5811	0,5434	0,5166	0,4964	0,4676	0,4337	0,3919	0,3366
18	0,7424	0,6341	0,5753	0,5371	0,5099	0,4894	0,4602	0,4255	0,3827	0,3253
19	0,7386	0,6295	0,5701	0,5315	0,5040	0,4832	0,4535	0,4182	0,3743	0,3151
20	0,7352	0,6254	0,5654	0,5265	0,4986	0,4776	0,4474	0,4116	0,3668	0,3057
21	0,7322	0,6216	0,5612	0,5219	0,4938	0,4725	0,4420	0,4055	0,3599	0,2971
22	0,7294	0,6182	0,5574	0,5178	0,4894	0,4679	0,4370	0,4001	0,3536	0,2892
23	0,7269	0,6151	0,5540	0,5140	0,4854	0,4636	0,4325	0,3950	0,3478	0,2818
24	0,7246	0,6123	0,5508	0,5106	0,4817	0,4598	0,4283	0,3904	0,3425	0,2749
25	0,7225	0,6097	0,5478	0,5074	0,4783	0,4562	0,4244	0,3862	0,3376	0,2685
26	0,7205	0,6073	0,5451	0,5045	0,4752	0,4529	0,4209	0,3823	0,3330	0,2625
27	0,7187	0,6051	0,5427	0,5017	0,4723	0,4499	0,4176	0,3786	0,3287	0,2569
28	0,7171	0,6030	0,5403	0,4992	0,4696	0,4471	0,4146	0,3752	0,3248	0,2516
29	0,7155	0,6011	0,5382	0,4969	0,4671	0,4444	0,4117	0,3720	0,3211	0,2466
30	0,7141	0,5994	0,5362	0,4947	0,4648	0,4420	0,4090	0,3691	0,3176	0,2419
60	0,6933	0,5738	0,5073	0,4632	0,4311	0,4064	0,3702	0,3255	0,2654	0,1644
∞	0,6729	0,5486	0,4787	0,4319	0,3974	0,3706	0,3309	0,2804	0,2085	0

Т а б л и ц а 7

5%-ные точки *F*-распределения (точки, в которых $\Phi. p. = 0,95$)
 (Займствованы из таблиц сэра Рональда Фишера и Ф. Ийтса: *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd Ltd., Edinburgh.)

$v_2 \backslash v_1$	1	2	3	4	5	6	8	12	24	∞
1	161,40	199,50	215,70	224,60	230,20	234,00	238,90	243,90	249,00	254,30
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,95	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Примечание. Нижняя 5%-ная точка представляет собой обратную величину того табличного значения, которое получится если v_1 и v_2 поменять местами, т. е. в качестве v_1 всегда следует выбирать количество степеней свободы, соответствующее наибольшей выборочной дисперсии.

1%-ные точки z-распределения (точки, в которых $\Phi. p. = 0,99$)(Займствованы из таблицы VI книги сэра Рональда Фишера *Statistical Methods for Research Workers*, Oliver and Boyd Ltd., Edinburgh.)

	Значения v_1											∞
	1	2	3	4	5	6	8	12	24			
1	4,1535	4,2585	4,2974	4,3175	4,3297	4,3379	4,3482	4,3585	4,3689			4,3794
2	2,2950	2,2976	2,2984	2,2988	2,2991	2,2992	2,2994	2,2997	2,2999			2,3001
3	1,7649	1,7140	1,6915	1,6786	1,6703	1,6645	1,6569	1,6489	1,6404			1,6314
4	1,5270	1,4452	1,4075	1,3856	1,3711	1,3609	1,3473	1,3327	1,3170			1,3000
5	1,3943	1,2929	1,2449	1,2164	1,1974	1,1838	1,1656	1,1457	1,1239			1,0997
6	1,3103	1,1955	1,1401	1,1068	1,0843	1,0680	1,0460	1,0218	0,9948			0,9643
7	1,2526	1,1281	1,0672	1,0300	1,0048	0,9864	0,9614	0,9335	0,9020			0,8658
8	1,2106	1,0787	1,0135	0,9734	0,9459	0,9259	0,8983	0,8673	0,8319			0,7904
9	1,1786	1,0411	0,9724	0,9299	0,9006	0,8791	0,8494	0,8157	0,7769			0,7305
10	1,1535	1,0114	0,9399	0,8954	0,8646	0,8419	0,8104	0,7744	0,7324			0,6816
11	1,1333	0,9874	0,9136	0,8674	0,8354	0,8116	0,7785	0,7405	0,6958			0,6408
12	1,1166	0,9677	0,8919	0,8443	0,8111	0,7864	0,7520	0,7122	0,6649			0,6061
13	1,1027	0,9511	0,8737	0,8248	0,7907	0,7652	0,7295	0,6882	0,6386			0,5761
14	1,0909	0,9370	0,8581	0,8082	0,7732	0,7471	0,7103	0,6675	0,6159			0,5500
15	1,0807	0,9249	0,8448	0,7939	0,7582	0,7314	0,6937	0,6496	0,5961			0,5269

ПРИЛОЖЕНИЕ

Значения χ_1

	1	2	3	4	5	6	8	12	24	∞
16	1,0719	0,9144	0,8331	0,7814	0,7450	0,7177	0,6791	0,6339	0,5786	0,5064
17	1,0641	0,9051	0,8229	0,7705	0,7335	0,7057	0,6663	0,6199	0,5630	0,4879
18	1,0572	0,8970	0,8138	0,7607	0,7232	0,6950	0,6549	0,6075	0,5491	0,4712
19	1,0511	0,8897	0,8057	0,7521	0,7140	0,6854	0,6447	0,5964	0,5366	0,4560
20	1,0457	0,8831	0,7985	0,7443	0,7058	0,6768	0,6355	0,5864	0,5253	0,4421
21	1,0408	0,8772	0,7920	0,7372	0,6984	0,6690	0,6272	0,5773	0,5150	0,4294
22	1,0363	0,8719	0,7860	0,7309	0,6916	0,6620	0,6196	0,5691	0,5056	0,4176
23	1,0322	0,8670	0,7806	0,7251	0,6855	0,6555	0,6127	0,5615	0,4969	0,4068
24	1,0285	0,8626	0,7757	0,7197	0,6799	0,6496	0,6064	0,5545	0,4890	0,3967
25	1,0251	0,8585	0,7712	0,7148	0,6747	0,6442	0,6006	0,5481	0,4816	0,3872
26	1,0220	0,8548	0,7670	0,7103	0,6699	0,6392	0,5952	0,5422	0,4748	0,3784
27	1,0191	0,8513	0,7631	0,7062	0,6655	0,6346	0,5902	0,5367	0,4685	0,3701
28	1,0164	0,8481	0,7595	0,7023	0,6614	0,6303	0,5856	0,5316	0,4626	0,3624
29	1,0139	0,8451	0,7562	0,6987	0,6576	0,6263	0,5813	0,5269	0,4570	0,3550
30	1,0116	0,8423	0,7531	0,6954	0,6540	0,6226	0,5773	0,5224	0,4519	0,3481
60	0,9784	0,8025	0,7086	0,6472	0,6028	0,5687	0,5189	0,4574	0,3746	0,2352
∞	0,9462	0,7636	0,6651	0,5999	0,5522	0,5152	0,4604	0,3908	0,2913	0

Таблица 9

1%-ные точки *F*-распределения (точки, в которых ф. р. = 0,99)(Займствованы из книги сэра Рональда Фишера и Ф. Иейтса: *Statistical Tables for Biological, Medical and Agricultural Research*, Oliver and Boyd Ltd., Edinburgh.)

$v_1 \backslash v_2$	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,42	99,46	99,50
3	34,12	30,81	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,12
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,27	9,89	9,47	9,02
6	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,67	5,28	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,11	4,73	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91
11	9,65	7,20	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36
13	9,07	6,70	5,74	5,20	4,86	4,62	4,30	3,96	3,59	3,16
14	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,43	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,18	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,45	3,08	2,65
18	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,00	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,86	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,51	3,17	2,80	2,36
22	7,94	5,72	4,82	4,31	3,99	3,76	3,45	3,12	2,75	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,41	3,07	2,70	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,03	2,66	2,21
25	7,77	5,57	4,68	4,18	3,86	3,63	3,32	2,99	2,62	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,23	2,90	2,52	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,20	2,87	2,49	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,47	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	2,99	2,66	2,29	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,66	2,34	1,95	1,38
∞	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,00

Примечание. Нижняя 1%-ная точка представляет собой обратную величину того табличного значения, которое получится, если v_1 и v_2 поменять местами, т. е. в качестве v_1 всегда следует выбирать количество степеней свободы, соответствующее наибольшей выборочной дисперсии.

Симметрические функции. Формулы представлений расширенных симметрических функций в терминах сумм степеней и обратные формулы
(Займствованы из таблиц Дэвид и Кендалла (1949).)

Вес 1

(1) = [1]

Вес 2

	[2]	[1 ²]
(2)	1	-1
(1) ²	1	1

Вес 3

	[3]	[21]	[1 ³]
(3)	1	-1	2
(2)(1)	1	1	-3
(1) ³	1	3	1

Вес 4

	[4]	[31]	[2 ²]	[21 ²]	[1 ⁴]
(4)	1	-1	-1	2	-6
(3)(1)	1	1	.	-2	8
(2) ²	1	.	1	-1	3
(2)(1) ²	1	2	1	1	-6
(1) ⁴	1	4	3	6	1

Вес 5

	[5]	[41]	[32]	[31 ²]	[2 ² 1]	[21 ³]	[1 ⁵]
(5)	1	-1	-1	2	2	-6	24
(4)(1)	1	1	.	-2	-1	6	-30
(3)(2)	1	.	1	-1	-2	5	-20
(3)(1) ²	1	2	1	1	.	-3	20
(2) ² (1)	1	1	2	.	1	-3	15
(2)(1) ³	1	3	4	3	3	1	-10
(1) ⁵	1	5	10	10	15	10	1

Вес 6

	[6]	[51]	[42]	[41 ²]	[3 ²]	[321]	[31 ³]	[2 ³]	[2 ² 1 ²]	[21 ⁴]	[1 ⁶]
(6)	1	−1	−1	2	−1	2	−6	2	−6	24	−120
(5)(1)	1	1	.	−2	.	−1	6	4	−24	144	
(4)(2)	1	.	1	−1	.	−1	3	−3	5	−18	90
(4)(1) ²	1	2	1	1	.	.	−3	.	−1	12	−90
(3) ²	1	.	.	.	1	−1	2	.	2	−8	40
(3)(2)(1)	1	1	1	1	1	1	−3	.	−4	20	−120
(3)(1) ³	1	3	3	3	1	3	1	.	.	−4	40
(2) ³	1	.	3	.	.	.	.	1	−1	3	−15
(2) ² (1) ²	1	2	3	1	2	4	.	1	1	−6	45
(2)(1) ⁴	1	4	7	6	4	16	4	3	6	1	−15
(1) ⁶	1	6	15	15	10	60	20	15	45	15	1

Примечание. — Для представления функций [] в терминах () используются только числа, стоящие в соответствующем столбце таблицы над главной диагональю (включительно). Например, $[41^2] = 2(6) - 2(5)(1) - (4)(2) + (4)(1)^2$. Аналогично, для представления () в терминах функций [] следует использовать только числа, стоящие в соответствующей строке слева от главной диагонали (включительно). Например, $(4)(1)^2 = [6] + 2[51] + [42] + [41^2]$.

Таблица 11

Многомерные k -статистики. Таблицы представлений l -статистик в терминах расширенных симметрических функций и обратные таблицы (Займствованы из статьи Уишарта (1952).)

1-й порядок

$$l_1 = [1]/n$$

2-й порядок

	l_{11}	l_2
$[1^2]/n^{[2]}$	1	-1
$[2]/n$	1	1

3-й порядок

	l_{111}	l_{21}	l_3
$[1^3]/n^{[3]}$	1	-1	2
$[21]/n^{[2]}$	1	1	-3
$[3]/n$	1	3	1

4-й порядок

	l_{1111}	l_{211}	l_{22}	l_{31}	l_4
$[1^4]/n^{[4]}$	1	-1	1	2	-6
$[21^2]/n^{[3]}$	1	1	-2	-3	12
$[2^2]/n^{[2]}$	1	2	1	.	-3
$[31]/n^{[2]}$	1	3	.	1	-4
$[4]/n$	1	6	3	4	1

5-й порядок

	l_{11111}	l_{2111}	l_{221}	l_{311}	l_{32}	l_{41}	l_5
$[1^5]/n^{[5]}$	1	-1	1	2	-2	-6	24
$[21^3]/n^{[4]}$	1	1	-2	-3	5	12	-60
$[2^21]/n^{[3]}$	1	2	1	.	-3	-3	30
$[31^2]/n^{[3]}$	1	3	.	1	-1	-4	20
$[32]/n^{[2]}$	1	4	3	1	1	.	-10
$[41]/n^{[2]}$	1	6	3	4	.	1	-5
$[5]/n$	1	10	15	10	10	5	1

6-й порядок

	l_{111111}	l_{21111}	$l_{221.}$	l_{3111}	l_{222}	l_{321}	l_{411}	l_{33}	l_{42}	l_{51}	l_6
$[1^6]/n^{[6]}$	1	-1	1	2	-1	-2	-6	4	6	24	-120
$[21^4]/n^{[5]}$	1	1	-2	-3	3	5	12	-12	-18	-60	360
$[2^21^2]/n^{[4]}$	1	2	1	.	-3	-3	-3	9	15	30	-270
$[31^3]/n^{[4]}$	1	3	.	1	.	-1	-4	4	4	20	-120
$[2^3]/n^{[3]}$	1	3	3	.	1	.	.	.	-3	.	30
$[321]/n^{[3]}$	1	4	3	1	.	1	.	-6	-4	-10	120
$[41^2]/n^{[3]}$	1	6	3	4	.	.	1	.	-1	-5	30
$[3^2]/n^{[2]}$	1	6	9	2	.	6	.	1	.	.	-10
$[42]/n^{[2]}$	1	7	9	4	3	4	1	.	1	.	-15
$[51]/n^{[2]}$	1	10	15	10	.	10	5	.	.	1	-6
$[6]/n$	1	15	45	20	15	60	15	10	15	6	1

Примечание. Для представления функций l в терминах расширенных симметрических функций используются лишь числа, стоящие в соответствующем столбце таблицы над главной диагональю (включительно); например, $l_{222} = -[1^6]/n^{[6]} + 3[21^4]/n^{[5]} - 3[2^21^2]/n^{[4]} + [2^3]/n^{[3]}$.

Для представления расширенных симметрических функций в терминах функций l используются лишь числа, стоящие в соответствующей строке таблицы под главной диагональю (включительно); например, $[2^21^2]/n^{[4]} = l_{111111} + 2l_{21111} + l_{2211}$.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- Абдель Ати (Abdel Aty S. H.) (1954), Tables of generalised k -statistics, *Biom.* **41**, 253.
- Бартлетт (Bartlett M. S.) (1934), The vector representation of a sample, *Proc. Camb. Phil. Soc.* **30**, 327.
- Бёрр (Burr I. W.) (1942), Cumulative frequency functions, *Ann. Math. Statist.* **13**, 215.
- Бозе (Bose S. S.) (1935), On the distribution of the ratio of variances of two samples drawn from a given normal bivariate correlated population, *Sankhya* **2**, 65.
- Болдуин (Baldwin E. M.) (1946), Percentage points of the t distribution, *Biom.* **33**, 362.
- Бэйкер (Baker G. A.) (1930), Distribution of the means of samples of n drawn at random from a population represented by the Gram-Charlier Series, *Ann. Math. Statist.* **1**, 199.
- Волд (Wold H.) (1934a), Sulle correzione di Sheppard, *Giorn. Ist. Ital. Att.* **4**, 304.
- Волд (Wold H.) (1934b), Sheppard's correction formulae in several variables, *Skand. Akt.* **17**, 248.
- Гамбел (Gumbel E. J.) (1934), Les valeurs extrêmes des distributions statistiques, *Ann. Inst. H. Poincaré*, **5**, 115.
- Гамбел (Gumbel E. J.) (1947), The distribution of the range, *Ann. Math. Statist.* **18**, 384.
- Гамбел (Gumbel E. J.) (1949), Probability tables for the range, *Biom.* **36**, 142.
- Гамбел (Gumbel E. J.) (1954), Statistical Theory of Extreme Values and some Practical Applications. (National Bureau of Standards, A. M. S. 33.)
- Гамбургер (Hamburger H.) (1920, 1921), Über eine Erweiterung des Stieltjesschen Momentproblems, *Math. Ann.* **81**, 235; **82**, 120 and 168.
- Гарвуд (Garwood F.) (1936), Fiducial limits for the Poisson distribution, *Biom.* **28**, 437.
- Гейен (Gayen A. K.) (1950), The variance ratio in random samples of any size drawn from non-normal universes, *Biom.* **37**, 236.
- Гири (Geary R. C.) (1930), The frequency distribution of the quotient of two normal variables, *J. Roy. Statist. Soc.* **93**, 442.
- Гири (Geary R. C.) (1936), The distribution of «Student's» ratio for non-normal samples, *Supp. J. Roy. Statist. Soc.* **3**, 178.
- Годвин (Godwin H. J.) (1945), On the distribution of the estimate of mean deviation obtained from samples from a normal population, *Biom.* **33**, 254.
- Годвин (Godwin H. J.) (1949), Some low moments of order statistics, *Ann. Math. Statist.* **20**, 279.
- Годвин (Godwin H. J.) (1955), Generalisations of Tchebycheff's inequality, *J. Amer. Statist. Ass.* **50**, 923.

- Годвин и Хартли (Godwin H. J. and Hartley H. O.) (1945), Probability integral and percentage points of the mean deviation in normal samples, *Biom.* **33**, 257.
- Гринвуд и Юл (Greenwood M. and Yule G. U.) (1920), An inquiry into the nature of frequency-distributions of multiple happenings, etc., *J. Roy. Statist. Soc.* **83**, 255.
- Гулдберг (Guldberg S.) (1935), Recurrence formulae for the semi-invariants of some discontinuous frequency distributions of n variables, *Skand. Akt.* **18**, 270.
- Дайсон (Dyson F. J.) (1943), A note on kurtosis, *J. Roy. Statist. Soc.* **106**, 360.
- Даниэлс (Daniels H. E.) (1954), Saddlepoint approximations in statistics, *Ann. Math. Statist.* **25**, 631.
- Джеймс (James G. S.) (1952), Notes on a theorem of Cochran, *Proc. Camb. Phil. Soc.* **48**, 443.
- Джексон (Jacson D.) (1921), Note on the median of a set of numbers, *Bull. Amer. Math. Soc.* **27**, 160.
- Дженкинсон (Jenkinson A. F.) (1955), The frequency-distribution of the annual maximum (or minimum) values of meteorological elements, *Quart. Jour. Roy. Met. Soc.* **81**, 158.
- Джини (Gini C.) (1912), Variabilità e Mutabilità, contributo allo studio delle distribuzioni e relazioni statistiche, *Studi Economico-Giuridici della R. Università di Cagliari*.
- Джонсон (Johnson N. L.) (1949a), Systems of frequency curves generated by methods of translation, *Biom.* **36**, 149.
- Джонсон (Johnson N. L.) (1949b), Bivariate distributions based on simple translation systems, *Biom.* **36**, 297.
- Джонсон (Johnson N. L.) (1957), A note on the mean deviation of the binomial distribution, *Biom.* **44**, 532.
- Дэвид, Хартли и Пирсон (David H. A., Hartley H. O. and Pearson E. S.) (1954), Distribution of the ratio, in a single normal sample, of range to standard deviation, *Biom.* **41**, 482.
- Дэвид Ф. (David F. N.) (1938), *Tables of the Correlation Coefficient*, Cambridge University Press.
- Дэвид Ф. (David F. N.) (1949a), Note on the application of Fisher's k -statistics, *Biom.* **36**, 383.
- Дэвид Ф. (David F. N.) (1949b), Moments of the z and F distribution, *Biom.* **36**, 394.
- Дэвид Ф. и Джонсон (David F. N. and Johnson N. L.) (1951), The effect of non-normality on the power function of the F -test, *Biom.* **38**, 43.
- Дэвид Ф. и Джонсон (David F. N. and Johnson N. L.) (1954), Statistical treatment of censored data. Part I, *Fundamental Formulae*, *Biom.* **41**, 225.
- Дэвид Ф. и Кендалл (David F. N. and Kendall M. G.) (1949, 1951, 1953, 1955), *Tables of symmetric functions*, *Biom.* **36**, 431; **38**, 435; **40**, 427; **42**, 223.
- Дэс (Das S. C.) (1956), The numerical evaluation of a class of integral. II, *Proc. Camb. Phil. Soc.* **52**, 442.
- Зя Уд-дин (Zia Ud-Din M.) (1954), Expression of the k -statistics, k_9 and k_{10} , in terms of power sums and sample moments, *Ann. Math. Statist.* **25**, 800.
- Ирвин (Irwin J. O.) (1927), On the frequency-distribution of the means of samples from a population having any law of frequency with finite moments etc, *Biom.* **19**, 225 and **21**, 431.
- Ирвин (Irwin J. O.) (1937), The frequency-distribution of the difference between two independent variates following the same Poisson distribution, *J. Roy. Statist. Soc.* **100**, 415.

- Ирвин и Кендалл (Irwin J. O. and Kendall M. G.) (1944), Sampling moments of moments for a finite population, *Ann. Eug. Lond.* **12**, 138.
- Иэйтс (Yates F.) (1935), Some examples of biased sampling, *Ann. Eug. Lond.* **6**, 202.
- Йоргенсен (Jørgensen N. R.) (1916), *Undersøgelser over Frequensflader og Korrelation* Busck, Copenhagen.
- Каллбэк (Kullback S.) (1934), An application of characteristic functions to the distribution problem of statistics, *Ann. Math. Statist.* **5**, 264.
- Каплан (Kaplan E. L.) (1952), Tensor notation and the sampling cumulants of k -statistics, *Biom.* **39**, 319.
- Каплански (Kaplansky I.) (1945), A common error concerning kurtosis, *J. Amer. Statist. Ass.* **40**, 259.
- Карлеман (Carleman T.) (1925), *Les fonctions quasi-analytiques*, Gauthier — Villars, Paris.
- Карлтон (Carlton G. A.) (1946), Estimating the parameters of a rectangular distribution, *Ann. Math. Statist.* **17**, 355.
- Кендалл Д. и Рао (Kendall D. G. and Rao K. S.) (1950), On the generalized second limit theorem in the theory of probabilities, *Biom.* **37**, 224.
- Кендалл (Kendall M. G.) (1938), The conditions under which Sheppard's corrections are valid, *J. Roy. Statist. Soc.* **101**, 592.
- Кендалл (Kendall M. G.) (1940a, b, c), Some properties of k -statistics, *Ann. Eug. Lond.* **10**, 106; Proof of Fischer's rules for ascertaining the sampling semiinvariants of k -statistics, *Ibid.* **10**, 215; The derivation of multivariate sampling formulae from univariate formulae by symbolic operation, *Ibid.* **10**, 392.
- Кендалл (Kendall M. G.) (1941), Relations connected with the tetrachoric series and its generalisation, *Biom.* **32**, 196.
- Кендалл (Kendall M. G.) (1942), On semiinvariant statistics, *Ann. Eug. Lond.* **11**, 300.
- Кендалл (Kendall M. G.) (1949a), Reconciliation of probability, *Biom.* **36**, 101.
- Кендалл (Kendall M. G.) (1949b), Rank and product-moment correlation, *Biom.* **36**, 177.
- Кендалл (Kendall M. G.) (1952), Moment-statistics in samples from a finite population, *Biom.* **39**, 14.
- Кендалл (Kendall M. G.) (1953), Discussion of Hotelling (1953), *J. Roy. Statist. Soc. B*, **15**, 225.
- Кендалл (Kendall M. G.) (1954), Two problems in sets of measurements, *Biom.* **41**, 560.
- Кендалл и Смит (Kendall M. G. and Smith B. B.) (1938—1939), Randomness and random sampling numbers, *J. Roy. Statist. Soc.* **101**, 147 and *Supp. J. Roy. Statist. Soc.* **6**, 51.
- Кокрэн (Cochran W. G.) (1934), The distribution of quadratic forms in a normal system, with applications to the analysis of covariance, *Proc. Camb. Phil. Soc.* **30**, 178.
- Кокрэн (Cochran W. G.) (1940), Note on an approximative formula for significance levels of z , *Ann. Math. Statist.* **11**, 93.
- Кокс (Cox D. R.) (1948), Asymptotic distribution of the range, *Biom.* **35**, 310.
- Кокс (Cox D. R.) (1954), Mean and coefficient of variation of range in small samples from non-normal populations, *Biom.* **41**, 469; correction, **42**, 277.
- Кооператив Стади (Co-operative Study) (1917), On the distribution of the correlation coefficient in small samples, *Biom.* **11**, 328.
- Корниш и Фишер (Cornish E. A. and Fisher R. A.) (1937), Moments and cumulants in the specification of distributions, *Rev. Inst. Int. Statist.* **5**, 307.

- Крамер (Cramér H.) (1926), On some classes series used in mathematical statistics, Skandinaviske Matematikercongres, Copenhagen.
- Крамер (Cramér H.) (1928), On the composition of elementary errors, Skand. Akt. 11, 13 and 141.
- Крамер (Cramér H.) (1937), Random Variables and Probability Distributions, Cambridge University Press. (Есть русский перевод: Г. Крамер, Случайные величины и распределения вероятностей, М., ИЛ, 1947.)
- Крамер и Вольд (Cramér H. and Wold H.) (1936), Some theorems on distribution functions, J. Lond. Math. Soc. 11, 290.
- Крейг (Craig A. T.) (1943), Note on the independence of certain quadratic forms, Ann. Math. Statist. 14, 195.
- Крейг (Craig C. C.) (1936), Sheppard's corrections for a discrete variable, Ann. Math. Statist. 7, 55.
- Кук (Cook M. B.) (1951), Bivariate k -statistics and cumulants of their joint sampling distribution, Biom. 38, 179.
- Ломницки (Lomnicki Z. A.) (1952), The standard error of Gini's mean difference, Ann. Math. Statist. 23, 635.
- Лукач (Lukacs E.) (1942), A characterisation of the normal distribution, Ann. Math. Statist. 13, 91.
- Лукач (Lukacs E.) (1952), An essential property of the Fourier transforms of distribution functions, Proc. Amer. Math. Soc. 3, 508.
- Лукач и Сас (Lukacs E. and Szász O.) (1952), On analytic characteristic functions, Pacific Jour. Maths. 2, 615.
- Лэнкастер (Lancaster H. O.) (1954), Traces and cumulants of quadratic forms in normal variables, J. Roy. Statist. Soc. B, 16, 247.
- Ляпунов (Liapounoff A.) (1901), Nouvelle forme du théorème sur la limite de probabilité, Mem. Acad. Sci. St. Pétr. (8), 12, No. 5.
- Мак Кей (McKay A. T.) (1935), Sampling distribution of the difference between the extreme observation and the mean, Biom. 27, 466.
- Мартин (Martin E. S.) (1934), On the correction for the moment coefficients of frequency-distributions when the start of the frequency is one of the characteristics to be determined, Biom. 26, 12.
- Марцинкевич (Marcinkiewicz J.) (1938), Sur une propriété de la loi de Gauss, Math. Zeitschr. 44, 612.
- Матерн (Matérn B.) (1949), Independence of non-negative quadratic forms in normally correlated variables, Ann. Math. Statist. 20, 119.
- Меррингтон и Томпсон (Merrington M. and Thompson C. M.) (1943), Tables of the percentage points of the inverted beta (F) distribution, Biom. 33, 73.
- Мизес (Von Mises R.) (1936), La distribution de la plus grande de n valeurs, Revue de l'Union Interbalkanique 1, 1.
- Миллер (Miller J. C. P.) (1954), Tables of binomial coefficients, Roy. Soc. Math. Tables, vol. 3.
- Молина (Molina E. C.) (1942), Tables of Poisson's Exponential Limit, Van Nostrand Co., Inc., New York.
- Мориуги (Moriguti S.) (1951), Extremal properties of extreme value distributions, Ann. Math. Statist. 22, 523.
- Морэн (Moran P. A. P.) (1956), The numerical evaluation of a class of integrals, Proc. Camb. Phil. Soc. 52, 230.
- Мэллоуз (Mallows C. L.) (1956), Generalisations of Tchebycheff's Inequalities, J. Roy. Statist. Soc. B, 18, 139.
- Мэрти (Murty V. N.) (1955), Distribution of the quotient of maximum values in samples from a rectangular distribution, J. Amer. Statist. Ass. 50, 1136.
- Нэйр (Nair U. S.) (1936), The standard error of Gini's mean difference, Biom. 28, 428.

- Пирсон (Pearson E. S.) (1926), A further note on the distribution of range in samples taken from a normal population, *Biom.* 18, 173.
- Пирсон (Pearson E. S.) (1930), A further development of tests for normality, *Biom.* 22, 239.
- Пирсон (Pearson E. S.) (1932), The percentage limits for the distribution of range in samples from a normal population, *Biom.* 24, 404.
- Пирсон и Хартли (Pearson E. S. and Hartley H. O.) (1942), The probability integral of the range in samples of n observations from a normal population, *Biom.* 32, 301.
- Пирсон и Хартли (Pearson E. S. and Hartley H. O.) (1950), Tables of the χ^2 integral and of the cumulative Poisson distribution, *Biom.* 37, 313.
- Пирсон К. (Pearson K.) (1900), On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling, *Phil. Mag.* (5), 50, 157.
- Пирсон К. (Pearson K.) (1919), On generalised Tchebycheff theorems in the mathematical theory of statistics, *Biom.* 12, 284.
- Пирсон К. (Pearson K.) (1924a), On the moments of the hypergeometrical series, *Biom.* 16, 157.
- Пирсон К. (Pearson K.) (1924b), On a certain double hypergeometrical series and its representation by continuous frequency surfaces, *Biom.* 16, 172.
- Пирсон К. (Pearson K.) (1931), Appendix to a paper by Professor Tokishige Hojo. On the standard error of the median to a third approximation, etc., *Biom.* 23, 361.
- Пирсон К., Стуффер и Дэвид (Pearson K., Stouffer S. A. and David F. N.) (1932), Further applications in statistics of the $T_m(x)$ Bessel function, *Biom.* 24, 293.
- Питман (Pitman E. J. G.) (1937), Significance tests which may be applied to samples from any population. II. The correlation coefficient test, *Supp. J. Roy. Statist. Soc.* 4, 225.
- Плэкетт (Plackett R. L.) (1947), Limits of the ratio of mean range to standard deviation, *Biom.* 34, 120.
- Плэкетт (Plackett R. L.) (1954), A reduction formula for normal multivariate integrals, *Biom.* 41, 351.
- Пойа (Polya G.) (1945), Remarks on characteristic functions, *Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press.
- Полсон (Paulson E.) (1942), An approximate normalisation of the analysis of variance distribution, *Ann. Math. Statist.* 13, 233.
- Пэрмэн и Пирсон К. (Pairman E. and Pearson K.) (1919), On the corrections for moment coefficients of limited range frequency-distributions when there are finite or infinite ordinates and any slopes at the terminals of the range, *Biom.* 12, 231.
- Райдер (Rider P. R.) (1955), The distribution of the product of maximum values in samples from a rectangular distribution, *J. Amer. Statist. Ass.* 50, 1142.
- Рафф (Raff M. S.) (1956), On approximating the point binomial, *J. Amer. Statist. Ass.* 51, 293.
- Романовский (Romanovsky V.) (1925), On the moments of the hypergeometrical series, *Biom.* 17, 57.
- Рубин (Ruben H.) (1954), On the moments of order statistics in samples from normal populations, *Biom.* 41, 200.
- Сархан и Гринберг (Sarhan A. E. and Greenberg B. G.) (1956), Estimation of location and scale parameters by order statistics from singly and doubly censored samples, *Ann. Math. Statist.* 27, 427.

- Сичел (Sichel H. S.) (1949), The method of frequency moments and its application to Type VII distributions, *Biom.* **36**, 404.
- Слуцкий Е. Е. (1950), Таблицы для вычисления неполной Γ -функции и функции вероятностей χ^2 . М., Изд-во АН СССР.
- Стильтес (Stieltjes J.) (1918), *Recherches sur les fractions continues, Œuvres*, Groningen.
- Стьюдент («Student») (1908), On the probable error of a mean, *Biom.* **6**, 1.
- Стьюдент («Student») (1919), An explanation of deviations from Poisson's law in practice, *Biom.* **12**, 211.
- Сэмпфорд (Sampford M.) (1953), Some inequalities on Mill's ratio and related functions, *Ann. Math. Statist.* **24**, 130.
- Сэндон (Sandon F.) (1924), Note on the simplification of the calculation of abruptness coefficients to correct crude moments, *Biom.* **16**, 193.
- Суй и Лоли (Hsu C. T. and Lawley D. N.) (1939), The derivation of the fifth and sixth moments of b_2 in samples from a normal population, *Biom.* **31**, 238.
- Тейкроу (Teichroew D.) (1956), Tables of expected values of order statistics and products of order statistics, *Ann. Math. Statist.* **27**, 410.
- Тиле (Thiele T. N.) (1903), *Theory of Observations*. Воспроизведено в *Ann. Math. Statist.* **2**, 165 с английского издания 1903 г.
- Типпетт (Tippett L. H. C.) (1925), On the extreme individuals and the range of samples taken from a normal population, *Biom.* **17**, 364.
- Томпсон (Thompson C. M.) (1941), Tables of percentage points of the χ^2 -distribution, *Biom.* **32**, 187.
- Томпсон, Пирсон, Комри и Хартли (Thompson C. M., Pearson E. S., Comrie L. J. and Hartley H. O.) (1941), Tables of percentage points of the incomplete beta-function, *Biom.* **32**, 151.
- Тьюки (Tukey J. W.) (1950), Some sampling simplified, *J. Amer. Statist. Ass.* **45**, 501.
- Уайз (Wise M. E.) (1954), A quickly convergent expansion for cumulative hypergeometric probabilities, direct and inverse, *Biom.* **41**, 317.
- Уикселл (Wicksell S. D.) (1917), On logarithmic correlation with an application to the distribution of ages at first marriage, *Medd. Lunds Astr. Obs.*, No. 84.
- Уильямс (Williams J. D.) (1946), An approximation to the probability integral, *Ann. Math. Statist.* **17**, 363.
- Уилсон и Хилфerti (Wilson E. B. and Hilferty M. M.) (1931), The distribution of chi-square, *Proc. Nat. Acad. Sci.* **17**, 694.
- Уишарт (Wishart J.) (1929), The correlation between productmoments of any order in samples from a normal population, *Proc. Roy. Soc. Edin.* **49**, 1.
- Уишарт (Wishart J.) (1949), Cumulants of multivariate multinomial distributions, *Biom.* **36**, 47.
- Уишарт (Wishart J.) (1952), Moment coefficients of the k -statistics in samples from a finite population, *Biom.* **39**, 1.
- Феллер (Feller W.) (1950), *An Introduction to Probability Theory and its Applications*. John Wiley and Sons, New York. Chapman and Hall, London. (Есть русский перевод: В. Феллер, Введение в теорию вероятностей и ее приложения, М., 1952.)
- Финни (Finney D. J.) (1938), The distribution of the ratio of estimates of the two variances in a sample from a normal bivariate population, *Biom.* **30**, 190.
- Фишер (Fisher R. A.) (1915), Frequency-distribution of the values of the correlation coefficient in samples from an indefinitely large population, *Biom.* **10**, 507.

- Фишер (Fisher R. A.) (1920), A mathematical examination of the methods of determining the accuracy of an observation by the mean error and by the mean square error, *Month. Not. R. Astr. Soc.* **80**, 758.
- Фишер (Fischer R. A.) (1921a), On the mathematical foundations of theoretical statistics, *Phil. Trans. Roy. Soc. A*, **222**, 309.
- Фишер (Fisher R. A.) (1921b), On the probable error of a coefficient of correlation deduced from a small sample, *Metron* **1**, No 4, 1.
- Фишер (Fischer R. A.) (1928), Moments and product-moments of sampling distribution, *Proc. Lond. Math. Soc.* (2), **30**, 199.
- Фишер (Fisher R. A.) (1935), The mathematical distributions used in the common tests of significance, *Econometrica*, **3**, 353.
- Фишер и Иэйтс (Fisher R. A. and Yates F.) (1953), *Statistical Tables for use in Biological, Agricultural and Medical Research*, 4th edition, Oliver and Boyd, Edinburgh.
- Фишер, Корбе и Уилльямс (Fisher R. A., Corbet A. S. and Williams C. B.) (1943), The relation between the number of species and the number of individuals, *J. Animal Ecology* **12**, 42.
- Фишер и Типпетт (Fisher R. A. and Tippet L. H. C.) (1928), Limiting forms of the frequency-distribution of the largest or smallest member of a sample, *Proc. Camb. Phil. Soc.* **24**, 180.
- Фишер и Уишарт (Fisher R. A. and Wishart J.) (1931), The derivation of the pattern formulae of two-way partitions from those of simpler patterns, *Proc. Lond. Math. Soc.* **33**, 195.
- Фостер и Стюарт (Foster F. G. and Stuart A.) (1954), Distribution-free tests in time-series based on the breaking of records, *J. Roy. Statist. Soc. B*, **16**, 1.
- Фреше (Fréchet M.) (1927), Sur la loi de probabilité de l'écart maximum, *Annales de la Soc. Polonoise de Math.* **6**, 92.
- Фреше (Fréchet M.) (1937), *Recherches théoriques modernes*, Gauthier — Villars, Paris.
- Фриш (Frisch R.) (1926), Sur les semi-invariants et moments employés dans l'étude des distributions statistiques, Oslo, *Skrifter af det Norske Videnskaps Academie*, II, Hist.-Filos. Klasse, No. 3.
- Хартли и Дэвид (Hartley H. O. and David H. A.) (1954), Universal bounds for mean range and extreme observation, *Ann. Math. Statist.* **25**, 85.
- Хатке (Hatke M. A.) (1949), A certain cumulative probability function, *Ann. Math. Statist.* **20**, 461.
- Хельмерт (Helmert F. R.) (1876), Die Genauigkeit der Formel von Peters zur Berechnung des Wahrscheinlichen Beobachtungsfehlers direkter Beobachtungen gleicher Genauigkeit, *Astronomische Nachrichten* **88**, No. 2096.
- Хоёо (Hojo T.) (1931, 1933), Distribution of the median, quartiles and inter-quartile distance in samples from a normal population, *Biom.* **23**, 315; A further note on the relation between the median and the quartiles in small samples from a normal population, *Biom.* **25**, 79.
- Холдейн (Haldane J. B. S.) (1937), The approximate normalisation of a class of frequency-distribution, *Biom.* **29**, 392.
- Холдейн (Holdane J. B. S.) (1939), Cumulants and moments of the binomial distribution, *Biom.* **31**, 392.
- Холдейн (Haldane, J. B. S.) (1942), Mode and median of a nearly normal distribution with given cumulants, *Biom.* **32**, 294.
- Холдейн (Haldane J. B. S.) (1948), Note on the median of a multivariate distribution, *Biom.* **35**, 414.
- Хотеллинг (Hotelling H.) (1953), New light on the correlation coefficient and its transforms, *J. Roy. Statist. Soc. B*, **15**, 193.

- Хотеллинг и Соломонс (Hotelling H. and Solomons L. M.) (1932), The limits of a measure of skewness, *Ann. Math. Statist.* 3, 141.
- Хэстингс, Мостеллер, Тьюки и Уинсор (Hastings C., Mosteller F., Tukey J. W. and Winsor C. P.) (1947), Low moments for small samples, *Ann. Math. Statist.* 18, 413.
- Чу (Chu J. T.) (1955), The «inefficiency» of the sample median for many familiar symmetric distributions, *Biom.* 42, 520.
- Шарлье (Charlier C. V. L.) (1931), Applications [de la théorie des probabilités] à l'astronomie, Gauthier — Villars, Paris.
- Шентон (Shenton L. R.) (1954), Inequalities for the normal integral, including a new continued fraction, *Biom.* 41, 177.
- Шепард (Sheppard W. F.) (1898), On the application of the theory of error to cases of normal distributions and normal correlations, *Phil. Trans. A*, 192, 101 and *Proc. Roy. Soc.* 62, 170.
- Шепард (Sheppard W. F.) (1939, посмертное издание), The Probability Integral, *British Ass. Math. Tables*, vol. 7, Cambridge University Press.
- Эджворт (Edgeworth F. Y.) (1904), The Law of Error, *Trans. Camb. Phil. Soc.* 20, 36 and 113 (с дополнением в оттиках).
- Эйткин (Aitken A. C.) (1950), Statistical independence of quadratic forms in normal variates, *Biom.* 37, 93.
- Элдертон (Elderton Sir W. P.) (1938a), *Frequency Curves and Correlation*, 3rd ed., Cambridge University Press.
- Элдертон (Elderton Sir W. P.) (1938b), Correzioni dei momenti quando la Curva è simmetrica, *Giorn. Ist. Ital. Att.* 16, 145.
- Элфвинг (Elfvig G.) (1947), Asymptotical distribution of range in samples from a normal population, *Biom.* 34, 111.
- Энис и Ллойд (Anis A. A. and Lloyd E. H.) (1953), Range of partial sums of normal variates, *Biom.* 40, 35.
- Юл (Yule G. U.) (1910), On the distribution of deaths with age when the causes of death act cumulatively, *J. Roy. Statist. Soc.* 73, 26.
- Юл (Yule G. U.) (1927), On reading a scale, *J. Roy. Statist. Soc.* 90, 570.
- Юл (Yule G. U.) (1938), On some properties of normal distributions, univariate and bivariate, based on sums of squares of frequencies, *Biom.* 30, 1.
- Юэн (Yuan P. T.) (1933), On the logarithmic frequency distribution, *Ann. Math. Statist.* 4, 30.
- Ясукава (Yasukawa K.) (1926), On the probable error of the mode of frequency-distributions, *Biom.* 18, 263.

УКАЗАТЕЛЬ

- Абдель Ати (Abdel Aty S. H.), таблицы многомерных k -статистик 418
- Аборт, распределение женщин по времени между зачатием и абортом (таблица 1.22) 48
- Абсолютные моменты (absolute moments) 93, неравенство Ляпунова для а. м. 93 и (упражнение 3.15) 133
- Австралийские свадьбы, распределение а. с. (таблица 1.8) 24, моменты распределения а. с. (пример 3.1) 85, асимметрии и эксцесс распределения а. с. (пример 3.18) 126
- Антимода (antimode) 64
- Апостериорная (posterior) вероятность 276
- Аппроксимация выборочных распределений (approximation to sampling distributions) 380
- Аппроксимация с применением метода перевала (approximation by the method of the steepest descent) 373
- Априорная (prior) вероятность 276
- Арифметическое среднее (arithmetic mean), см. среднее арифметическое
- Асимметричные распределения (asymmetrical distributions) 23—25, см. также Асимметрия
- Асимметрия (skewness) 25; меры а. 124—125 и (упражнение 3.22) 135; стандартные ошибки b_1 325—336; выборочные моменты b_1 411—412 и (упражнение 12.9) 421
- Афолт (Uphoff W. M.), данные (таблица 1.5) 22
- Байеса теорема (Bayes' theorem) 275; постулат 276—282; максимальное правдоподобие 280—282 и (упражнения 8.1—2) 284—285; при выборочном изучении качественных признаков (упражнение 9.3) 315
- Бартлетт (Bartlett M. S.), характеристика нормального распределения 503
- Бернулли числа и полиномы (Bernoulli numbers and polynomials) 118—119 и (упражнение 4.20) 170
- Бёрр (Burr I. W.), подгонка ф. р. 242—243 и (упражнение 6.19) 248—249
- Бета-распределение (Beta distribution) первого рода, арифметическое среднее (пример 2.2) 58; геометрическое и гармоническое средние (пример 2.4) 59—60; дисперсия (пример 2.8) 72; первого и второго рода 210—212; полиномы Якоби при разложении в ряды 228; преобразование Джонсона (упражнения 6.14—15) 247—248; см. также Распределения типа I и VI
- Билетов выбор или лотерея (ticket sampling) 296
- Биномиальное распределение (binomial distribution), производящая функция 45; моменты (пример 2.3) 58 и (пример 2.9) 72; (пример 3.2) 86—88; производящая функция моментов (характеристическая функция) (пример 3.5) 92; факториальные моменты (пример 3.8) 99 и (упражнение 3.7) 131; факториальные семинварианты (пример 3.14) 111; асимметрия и эксцесс (пример 3.19) 126; асимптотическая нормальность (пример 4.6) 153—154; общие сведения 171—178; х. ф., моменты и семинварианты 173—175 и (упражнение 5.1) 201—202; функция распределения 175—178; таблицы 177—178; ошибки при нормальной аппроксимации 178; смешанные биномиальные совокупности 180—182 и (упражнение 5.6) 203; двумерное 199—201; неполные моменты (упражнения 5.2—3) 202; среднее отклонение (упражнение (5.4) 202; в выборочном изучении качественных признаков 305—307; (упражнение 9.14) 317; с отрицательным индексом, см. Отрицательное биномиальное распределение; распределение среднего (пример 11.10) 360
- Бобы, распределение (таблица 1.15) 40; гистограмма (рис. 1.6) 41; подгонка распределения типа Пирсона (пример 6.1) 214—216; подгонка с помощью рядов Грам-Шарлье (пример 6.2) 224—225; подгонка с помощью логнормального распределения (пример 6.5) 237—238
- Бозе (Bose S. S.), распределение дисперсионного отношения коррелированных величин (упражнение 16.16) 546—547
- Болдуин (Baldwin E. M.), таблица процентных точек t -распределения 521
- Большие выборки (large samples), аппроксимации в случае б. в., см. Стандартные ошибки
- Больших чисел закон (Law of large Numbers) 267—268 и (упражнение 7.20) 274; усиленный 268
- Борткевич (Bortkiewicz L.), данные о самоубийствах (таблица 1.6) 22
- Браун (Brown F. A. C.), см. Эйтчисон
- Бридж, вероятности при игре в б. (пример 7.2) 254—256
- Бросание костей (dice throw), данные Уэлдона (таблица 1.14) 38; (таблица 1.16) 45; (таблица 5.1) 173; (пример 9.10) 308
- Букер (Booker H. S.), данные (таблица 1.21) 47
- Бьенэме — Чебышева неравенство (Bienayme—Tchebycheff inequality) 128—129, 149, 267; в случае выборочного изучения

- качественных признаков (упражнение 9.4) 315; в связи со стандартными ошибками 325
- Бэбингтон Смит (Babington Smith B.), данные о телефонных номерах (таблица 1.4) 22; данные о работе с латинком случайных чисел (пример 9.6) 291—292; таблицы случайных чисел 298, 304
- Бэйкер (Baker G. A.), распределение суммы случайных переменных, плотность распределения каждой из которых задается в виде ряда типа А (упражнение 11.10) 375
- Варианта (variate) 16, преобразования вариант 35—37, 43—44
- Вариация (variation), коэффициент в. 74—75; стандартная ошибка в. (пример 10.5) 324—325, 326, 336
- Ватсон (Watson G. S.), полиномы Якоби в разложении бета-распределения, 229
- Вероятность (probability) 250—285; исчисление в. 250—274; основные правила 252—253; важность точного описания множества элементарных исходов (примеры 7.1—2) 253—256 и (упражнения 7.6, 7.11) 271—272; в непрерывном случае 257—258; геометрическая в. (пример 7.4) 258 и (упражнения 7.9—11) 272; в. и статистические распределения 259—260; априорная и апостериорная в. 276; см. также Байеса теорема, Правдоподобие
- Бес, распределение мужчин по в. (таблица 1.10) 26
- Вист, распределение карт при игре в в. (таблица 5.5) 191
- Возвращение (replacement), выбор с и без в. 287; (пример 9.12) 312—313; (упражнения 9.8—9) 316
- Болд (Woid H.), поправки Шенпарда 113, 117—118, 124; критерий Карлемана для проблемы моментов в случае многомерных распределений 160; таблица случайных нормальных отклонений 299
- Вторая предельная теорема (Second Limit theorem), 165—167
- Бу (Woo T. L.), данные о черепах (таблица 1.18) 46
- Выбор (sampling) с и без возвращения 287; из гипотетической совокупности 288; билетов или лотерея 296; из непрерывных совокупностей 304—305; см. также Случайный выбор
- Выборочное распределение (sampling distribution) 265—266; роль в выборочной теории 311; точные в. р. 341—379; аналитический способ нахождения 341—351; геометрический способ нахождения 351—359; получение с помощью х. ф. 359—365; отыскание в. с. сумм с помощью метода индукции 365—367; отношений 367—373; многомерное 373; аппроксимация в. р. 380—445; см. также на названия распределений
- Выборочные моменты (sampling moments), см. *k*-статистики
- Выводенное распределение (unit distribution), х. ф. (пример 4.3) 143
- Гамбел (Gumbel E. J.), двумерные ф. р. (упражнение 1.22) 53; теория экстремальных (крайних) значений 453 (сноска), 456 (сноска), 460—461; редуцированный размах и редуцированная средняя размаха 469—471; таблицы распределений крайних значений 470—471; функция распределения размаха (упражнение 14.19) 476
- Гамбургер (Hamburger H.), проблема моментов 157 (сноска)
- Гамма-распределение (Gamma distribution), семинварианты и обращение х. ф. (пример 4.4) 143—144; однозначное определение по моментам (упражнение 4.6) 168; как распределение квадрата нормальной случайной величины (упражнение 4.18) 170; разложение Грам-Шарлье типа А (пример 6.3) 227—228; связь с полиномами Лагерра при разложении в ряд 228; приближение к нормальному (пример 6.4) 233—234; логарифмическое преобразование (Джойсона) (упражнение 6.13) 247; характеристизация (упражнение 15.22) 509; см. также распределение типа III
- Гарвуд (Garwood F.), аппроксимация χ^2 распределения (таблица 16.1) 516
- Гармоническое среднее (harmonic mean), см. Среднее гармоническое
- Гаусс (Gauss C.) и нормальное распределение, 192 (сноска)
- Гаусса — Винклера неравенство (Gauss — Winkler inequality) (упражнение 3.18) 134
- Гейен (Geyen A. K.), *z*-преобразованные коэффициента корреляции 540 (сноска)
- Гельдера неравенство (Hölder's inequality) 62 и (упражнение 3.15) 133
- Генеральная совокупность (population) как основное понятие статистической теории 15; существующая г. с. 38—39; гипотетическая г. с. 39; типы г. с., 287—288
- Геометрические вероятности (geometric probabilities) (пример 7.4) 258 и (упражнения 7.9—11) 272
- Геометрическое среднее (geometric mean), см. Среднее геометрическое
- Гипергеометрическое распределение (hypergeometric distribution) 188—191; х. ф. и моменты 189—190; предельные формы 197—198; факториальные моменты (упражнение 5.26) 207; при выборе для изучения качественных признаков (пример 9.12) 312—313
- Гипотезы 280—283
- Гипотетическая совокупность (hypothetical population) 39; выбор из нее 288
- Гри (Geary R. C.), распределение отношения 369 (упражнение 11.11) 375—376; независимость среднего и дисперсии выборки из нормальной генеральной совокупности (упражнение 11.19) 377—378 и (пример 12.7) 401—402
- Гистограммы (histogram) 19; двумерные 39—41
- Главное значение (principal value) в связи с математическим ожиданием 89 (сноска) и (упражнение 7.19) 274
- Годвин (Godwin H. J.), неравенства Чебышевского типа 130; распределение среднего отклонения 334; моменты порядковых статистик 450
- Гомоскедастичность (homoscedastic) 530
- Госсет (Gosset W. S.), см. Студент

- Грам-Шарлье ряды (Gram-Charlier series) типа A 219—228; соответствующие ф. р. 224; подгонка к данным о бобах (пример 6.2) 224—225; условия сходимости 225—227 и (упражнение 6.9) 246; двумерные 244 и (упражнение 6.16) 248; (упражнения 6.10—11) 246—247; распределение сумм (упражнение 11.10) 375; см. также Эджворта ряды типов В и С 228 и (упражнения 6.8) 246
- Графическое представление частотных распределений (graphical representation of frequency-distributions), см. Полигон частот, Гистограмма
- Гринберг (Greenberg В.), таблица дисперсий порядковых статистик в нормальном случае 452
- Гривуд (Greenwood М.), данные о нечастотных случаях на производстве (таблица 5.3) 183
- Гулдберг (Guldberg S.), семинварианты многомерных распределений (упражнение 5.19) 206
- Гэлбран (Galbrun Н.), сходимость рядов Грам-Шарлье 226
- Дайсон (Dyson F. J.), эксцесс (упражнение 3.20) 134
- Даннэлс (Daniels Н. Е.), распределение отклонений 369, (упражнение 1.24) 379, метод перевала 373—374
- Датчик случайных чисел (randomizing machine) (пример 9.3) 291—292
- Двойное экспоненциальное (Лапласа) распределение (double exponential (Laplace) distribution), обращение х. ф. (упражнение 4.3) 168; дисперсия средней широты (пример 14.5) 469
- Двумерное (bivariate) биномиальное распределение
- Двумерное (binomial distribution) 199—201; нормальное распределение (normal distribution) (пример 1.1) (пример 3.17) 121; моменты (пример 3.19) 126; однозначность определения по моментам (упражнение 4.17) 170; совместная х. ф. квадратов величин (упражнение 4.19) 170; как предел двумерного биномиального распределения 200—201; двумерные ряды Грам-Шарлье и дифференциальное уравнение (упражнения 6.16—17) 248; х. ф. сумм случайных величин (упражнение 7.13) 273; распределение отношения случайных величин (упражнение 11.22) 378—379; моменты ковариации выборки (пример 13.3) 442—443; дисперсии и ковариации k -статистик (упражнения 13.1—2 и 13.7—8) 443, 444; (пример 15.1) 481; теорема Шеппарда с медианой дихотомии 485; (упражнения 15.7—8) 506—507; условная ковариация при линейных связях (упражнение 15.18) 508; недостаточность знания одномерных нормальных распределений (упражнение 15—20) 508; выбор из д. н. р. 529—544; распределение коэффициента корреляции выборки 529—543; распределение коэффициентов регрессии выборки 542—544; пуассоновское распределение (упражнение 5.11)
- Двумерные (bivariate) распределения (distributions) 39—44; (упражнение 1.22) 53; подгонка 244
- k -статистики (k -statistics) 424—430
- Двумерные моменты и семинварианты (moments and cumulants) 120—124; стандартные ошибки их 327—328
- Дербини (Durbin J.), полиномы Якоби в разложении бета-распределения 229
- Децилы (deciles) 65; интердецильные широты 67; стандартные ошибки д. 337; см. также Порядковые статистики
- Джеймс (James G. S.), преобразование случайной переменной (упражнение 13.6) 444; разложение суммы квадратов 497
- Дженкинсон (Jenkinson A. F.), асимптотическое распределение экстремальных значений 454
- Джеффрейс (Sir Harold Jeffreys) и постулат Бафеса 277 (сноска)
- Джини (Gini C.), коэффициент средней разности 74, 333—335; коэффициент рассеяния 75
- Джонсон (Johnson N. L.), среднее отклонение биномиального распределения (упражнение 5.4) 202; система преобразований вариант 234—235, 239, 242; подгонка двумерных распределений 244; моменты дисперсионного отношения 414; моменты порядковых статистик 452 и (упражнения 14.1 и 14.3) 472—473
- Джонсона преобразования (Johnson transformations) 234—243; использование при подгонке распределения к данным о бобах (примеры 6.5, 6.7) 237—238, 241—242; к данным об облачности (пример 6.6) 239—240; применение к гамма-распределению (упражнение 6.13) 247
- Джэксон (Jackson D.) о медиане 63
- Дискретные распределения (discontinuous distributions) 21—23, 28—29, 32; обращение соответствующих им х. ф. 140—141
- Дисперсионного отношения распределение (variance-ratio distribution) (пример 11.20) 369—371; общие сведения 521—529; отношение между F - и z -распределениями 522—523; вид F -распределения 523; х. ф. z -распределения 523; приближение z -распределения к нормальному 524; ф. р. F - и z -распределений и таблицы 524; распределения величин $1/F$ и $1/z$ 524—525; аппроксимации z -распределения 525—528; отношение к нормальному, χ^2 - и t -распределению 528—529; моменты F -распределения (упражнение 16.1) 544; распределение в случае коррелированных нормальных случайных величин (упражнение 16.16) 546—547
- Дисперсию уравнивающее преобразование (variance-stabilizing transformations) (упражнение 10.17) 339 и (упражнения 16.18—19) 547
- Дисперсия (variance) 68; как половина средней квадратичной разности 74; стандартная ошибка 320, 336; точные среднее и дисперсия выборочного распределения д. (упражнение 10.13) 339, 381 (пример 12.1) 383 и (пример 12.3), 397; точное распределение в случае выборок из нормальной генеральной совокупности (пример 11.7) 354; выборочная дисперсия д. выборки из конечной генеральной совокупности (упражнение

- 12.11) 422; см. также стандартное отклонение
- Доход, распределение лиц по доходам 18, 20, 31
- Дэвид (David H. A.), верхние границы для средних значений порядковых статистик 462, 463; границы для среднего значения размаха 467; распределение отношения случайного отклонения от среднего выборки к стандартному отклонению выборки (упражнение 16.20) 547
- Дэвид Ф. (David F. N.), таблицы симметрических функций 102 (сноска), 383, распределение разности случайных переменных с распределениями III типа (упражнение 11.15) 377; моменты симметрических функций 413; моменты дисперсионного отношения 414; моменты порядковых статистик 452 и (упражнения 14.1, 14.3) 472, 473; таблицы коэффициентов корреляции 537, 541
- Дэс (Das S.), отношение Миллса (упражнение 5.13) 204
- Закон больших чисел (Laws of Large Numbers) 267, 268 и (упражнение 7.20) 274
- Замена переменных (change of variables) 44, 342—344
- Зелен (Zelen M.), двумерный нормальный интеграл 483
- Ирвин (Irwin J. O.), распределение среднего в случае выборок из совокупности с распределением II типа (упражнение 11.14) 377; разность двух пуассоновских переменных (упражнение 11.16) 377; выбор из конечных совокупностей 416
- Интервал (interval), см. Класс-интервал
- Интердецильная широта (interdecile gap) 68
- Интерквартильная широта (interquartile range) 67; стандартная ошибка семинтерквартильной широты (пример 10.8) 332; стандартная ошибка н. ш. в нормальном случае 450
- Иэйтс (Yates F.), данные о высоте растений (пример 9.1) 289; таблицы случайных чисел 299; таблица процентных точек χ^2 -распределения 513; таблица процентных точек t -распределения 520; таблицы процентных точек F - и z -распределений 524
- Йоргенсен (Jorgensen N. R.), таблицы полиномов Чебышева — Эрмита 218
- Йохансен (Johansen W.), данные о бобах, цитированные Преториусом (таблица 1.15) 40
- Кавата (Kawata T.), характеристика нормального распределения 502
- Каллбэк (Kullback S.), распределения геометрических средних (упражнения 11.3, 11.9) 374, 375
- Каплан (Kaplan E. Z.), тензорные формулы для k -статистик 440 и (упражнение 13.9) 444
- Каплански (Kaplansky J.), эксцесс (упражнение 3.21) 135
- Карлсман (Carleman T.), критерий единственности для проблемы моментов 159
- Карлтон (Carlton G. A.), выбор из прямоугольного (равномерного) распределения (упражнения 14.17—18) 475—476
- Картофель, систематическая ошибка (пристрастие) в предсказании урожая (пример 9.4) 292
- Качественные признаки (attributes), выборочное изучение качественных признаков 305—315; (упражнения 9.1—9.12 и 9.14) 315—317; (упражнение 10.4) 338
- Квадратичные формы от нормальных случайных переменных (quadratic forms in normal variates) 489—500 и (упражнения 15.14—15) 508; их независимость 493—496; независимость от линейных форм 496 и (упражнение 15.13) 507; разложение сумм квадратов 497—500
- Квантили (quantiles) 65; графический способ определения 67; асимптотическая нормальность и стандартные ошибки 328—332, 337; ковариации, к. 331; см. также Экстремальные значения, Порядковые статистики
- Квартили (quartiles) 65; интерквартильная широта как мера рассеяния 67; стандартные ошибки 337
- Кендалл (Kendall D. G.) и Рао (Rao K. S.), последовательности ф. р. 149; вторая предельная теорема 161—167; стремление старших моментов к бесконечности, когда последовательность ф. р. сходится к нормальной (пример 4.12) 169
- Кендалл М. (Kendall M. G.), цифры из телефонной книги (таблица 1.4) 22; таблицы симметрических функций 102—103 (сноска) и 383; поправки Шепарда 117; двумерные ряды Грам-Шарлье (упражнения 6.16—17) 248; о теориях вероятностей 251; данные о случайности (таблица 9.3) 291; таблица случайных чисел 298—299; случайные перестановки 302 (сноска); выбор из конечных совокупностей 416; многомерные k -статистики 418; п. ф. м. k -статистик (упражнения 12.15—16) 422—423; формальный метод получения многомерных k -статистик из одномерных 428; дисперсия ближайшего наблюдаемого значения к истинному среднему в нормальном случае (упражнения 14.10—11) 474; многомерный нормальный интеграл 483, 486; преобразования коэффициента корреляции (упражнение 16.19) 547
- Класс-интервал (class-interval) 17; замечания о к.-и. 19—21
- Класс-частота (class-frequency) 18
- Ковариация (covariance) 120; моменты к. (пример 13.3) 442—443; матрица к., см. Рассеяния матрица
- Кокран (Cochran W. G.), случайные перестановки 302 (сноска); независимость квадратичных форм 494; теорема о разложении суммы квадратов 497—500 и (упражнение 15.17) 508; аппроксимация z -распределения 526
- Кокран теорема (Cochran's theorem) 497—500
- Кокс Г. (Cox G. R.), случайные перестановки 302 (сноска)

- Кокс Д. (Cox D. R.), распределение размаха 466 и (упражнение 14.15) 475; редуцированный размах 470
- Колонии, распределение числа индивидуумов в к. (упражнения 5.7—8) 203 и (упражнения 5.21—2) 206
- Конверты и письма, задача об их соответствии (пример 7.3) 256—257 и (упражнение 7.4) 271
- Конечные совокупности (finite populations) 287; k -статистики при выборе из к. с. 414—418; двумерный случай 429—430
- Кооператив Стади (Co-operative Study), о распределении коэффициента корреляции 536 и (упражнение 16.15) 546
- Корбе (Corbet A. S.), логарифмическое распределение (таблица 5.4) 188
- Корниш (Cornish E. A.), нормализация распределений 233; аппроксимация z -распределения 525
- Корреляции коэффициент (correlation coefficient), стандартная ошибка (пример 10.6) 328, 337; преобразование, выравнивающее дисперсию (упражнение 10.17) 339; выборочное распределение при выборке из двумерной нормальной генеральной совокупности 529—539; таблицы к. к. 537; z -преобразование Фишера 539—541, (упражнение 16.21) 548; рекуррентное соотношение для ф. п. (упражнение 16.14) 546; дифференциальное уравнение для ф. п. (упражнение 16.15) 546; среднее и дисперсия к. к. для выборки из генеральной совокупности с распределением, имеющим непрерывные независимые компоненты (упражнение 16.17) 547
- Коши — Буяковского неравенство (Cauchy — Schwarz inequality) 62
- Коши распределение (Cauchy distribution), бесконечность моментов (пример 3.3) 89; х. ф. (пример 3.13) 109; обращение х. ф. (пример 4.2) 142; совпадение с распределением среднего выборки (пример 11.1) 344—346 и (пример 11.17) 366; как распределение отношения квадратов нормальных случайных величин (пример 11.21) 372; как распределение отношения не нормально распределенных случайных величин (упражнение 11.23) 379
- Коши типа распределения в теории экстремальных значений (Cauchy-type distributions in extreme-value theory) 457
- Крамэр (Cramér H.), условия для того, чтобы функция была х. ф. 146—147; критерий Карлемана для проблемы моментов в многомерном случае 160; сходимость рядов Грам-Шарлье 225—227, 234; распределение отношения 369 (goodness of-fit distribution) пример 15.3) 490—492 и (упражнение 15.21) 508—509
- Крэйг А. (Craig A. T.), квадратичные формы 491, 499. (упражнение 15.16) 508
- Крэйг С. (Crsig C. S.), поправки к моментам дискретных распределений 117 и (упражнение 3.13) 132
- Критерий согласия χ^2 и χ^2 -распределение
- Крэйг Р. (Craig R.), данные (таблица 1.6) 22
- Кук (Cook M. B.), моменты и семинварианты многомерных распределений 123 (сноска); выборочные семинварианты многомерных k -статистик 428, 442 и (упражнение 13.4) 444
- Кэнуй (Quepouille M. H.), отрицательное биномиальное распределение как комбинация пуассоновского и логарифмического распределений (упражнение 5.21) 206; случайные отклонения 299
- Лагерра полиномы и разложения гамма-распределения (Laguerre polynomials and Gamma distribution expansions) 228
- Лаплас (Marquis de Laplace P. S.) и нормальное распределение 192 (сноска); непрерывная дробь для нормального распределения 194—195 и (упражнение 5.12) 204; лапласовское правило следования (пример 8.3) 278—279 и (упражнение 8.4) 285
- Лапласа распределение (Laplace distribution), см. Двойное экспоненциальное распределение
- Лаха (Laha R. G.), характеристика гамма-распределения (упражнение 15.22) 509
- Леви (Levy P.) о х. ф. 136
- Ли (Lee A.), данные (таблица 1.19) 46
- Либерман (Lieberman G. J.), таблицы гинергеометрического распределения 191
- Линдеберг (Lindeberg J. W.), центральная предельная теорема 269
- Линейные функции (linear functions), стандартные ошибки л. ф. от случайных переменных 323—324; распределение л. ф. от независимых нормальных случайных переменных (пример 11.2) 346—347, (пример 11.5) 352—353; независимость л. ф. от квадратичных форм 496
- Ллойд (Lloyd E. H.), матрица рассеяния многомерного нормального распределения (упражнение 15.11) 509
- Логарифмическое распределение (logarithmic distribution) 186—188; х. ф. и моменты 187—188; отрицательное биномиальное распределение как комбинация логарифмического и пуассоновского распределений (упражнение 5.21) 206
- Логистическое распределение (logistic distribution), х. ф. (упражнение 4.21) 170; ф. р. экстремального значения при выборке из генеральной совокупности с л. р. (упражнение 14.2) 472—473
- Логнормальное распределение (lognormal distribution) 235—238
- Лоли (Lawley D. N.), выборочные семинварианты k -статистик в нормальном случае 411
- Ломиники (Lomnicki Z. A.), стандартная ошибка средней разности Джини 334—335
- Лукач (Lukacs E.) х. ф. 136 (сноска). 147, 167; независимость среднего и дисперсии выборки как характеристическое свойство нормального распределения (упражнение 11.19) 377—378, (пример 12.7) 401—402 и 500—503; характеристика 505

- Лэнкастер (Lancaster H. O.), независимость квадратичных форм от нормальных случайных переменных 494, 499; разложение сумм квадратов 497; характеристизация нормального распределения 504—506; условие для того, чтобы одна квадратичная форма была частью другой (упражнение 15.15) 508
- Ляпунов А., неравенство для моментов 93 и (упражнение 3.15) 133
- Мак Кей (McKay A. T.), распределение порядковых статистик в нормальном случае (упражнение 14.14) 475
- Максимальное правдоподобие (maximum likelihood), см. Правдоподобие
- Маргинальное (частное) распределение (marginal distribution) 42
- Марков А., неравенство 129, (упражнение 7.16) 273
- Мартин (Martin E. S.), поправки к группировке 117
- Марцинкевич (Marcinkiewicz J.), условия для того, чтобы экспонента от полинома была х. ф. 147 и (упражнение 4.11) 169
- Математические таблицы, распределение цифр из м. т. 298
- Матерн (Matern B.), независимость неотрицательных квадратичных форм 495
- Медиана (median) 162; вычисление (пример 2.5) 81; неоднозначность определения 62—63; двумерного распределения 63; связь со средним и модой 64—65 и (упражнение 6.20) 249; как значение, минимизирующее среднее отклонение (упражнение 2.1) 80; стандартная ошибка (пример 10.7) 330, 336, (пример 11.4) 351; распределение в нормальном случае 447—450; распределение в случае распределения Коши (упражнение 14.1) 472; границы для дисперсии (упражнение 14.7) 474
- Медианная дихотомия (median dichotomy) 485
- Менделя закон (Mendelian law) в выборе при изучении качественных признаков 306—307
- Меррингтон (Merrington M.), таблица процентов точек распределения дисперсионного отношения 524
- Мерсер (Mercer W.), данные (таблица 1.20) 47
- Мизес (R. von Mises), асимптотические распределения экстремальных значений 456
- Миллс отношение (Mills's ratio) 193; (упражнения 5.13—16) 204—205
- Многомерное нормальное распределение (multivariate normal (multinormal) distribution), правила для k -статистик (упражнение 13.12) 445; общие сведения 477—509; вырожденные (сингулярные) 478 и (пример 15.3) 490—492; х. ф. и моменты 478—481; линейные функции нормальных переменных 481—482; нормальность маргинальных распределений 482; интеграл 483—489 и (упражнения 16.6, 15.12) 506, 507; квадратические формы 489—500 и (упражнение 15.13) 507; характеристизация 500—506; нормальность условных распределений (упражнение 15.1) 506; условия на дисперсионную матрицу (упражнение 15.9—11) 507; условная ковариация при наличии линейных связей (упражнение 15.19) 508; недостаточность нормальности маргинальных распределений для совместной нормальности (упражнение 15.20) 508
- Многомерные (multivariate) распределения 39—45; моменты и семинварианты 120—124, х. ф. 154—156; выборочные распределения 373; k -статистики 424—430, 438—443
- Мода (mode) 63—64; связь со средним и медианой 64—65 и (упражнение 6.20) 249
- Моменты (moments) 68—69, 83—84; поправки к группировке 72—73, 111—119; вычисление относительно одной точки через значения относительно другой точки 84—85; вычисление 85—90; выражение через факториальные моменты 95; связь с семинвариантами 101—106; многомерные м. 120—121; как характеристики распределения 126—128; неоднозначность определения ф. р. по моментам, см. Проблема моментов; при подгонке распределений Пирсона 214—216; стандартные ошибки 318—328 и (упражнение 10.5) 338; ковариации (упражнения 10.7, 10.16) 338, 339; распределения 362—365, см. Абсолютные моменты, Факториальные моменты, Неравенства для м.
- Моменты вероятностей (frequency-moments) нормального распределения (упражнение 5.10) 203—204; в подгонке распределений 243—244
- Мономиальные симметрические функции (monomial symmetric functions) 243—244
- Моригути (Moriguti S.), границы для моментов вероятностей 244; границы для средних значений порядковых статистик 463
- Морэн (Moran P. A. P.), многомерный нормальный интеграл 487, 489
- Муавр (DeMoivre A.), первооткрыватель нормального распределения 192 (сноска)
- Мультимодальность (multimodal) 64
- Мультиномиальное распределение (multinomial distribution) 198—199; семинварианты многомерного м. р. 201; семинварианты (упражнение 5.19) 206; дисперсия функций от частот (упражнение 10.2) 337—338
- Мур (Moore G.), данные (таблица 1.19) 46
- Мэллоуз (Mallows C. L.), неравенства чебышевского типа 130
- Мэртн (Murry V. N.), распределение отношений наибольших значений в случае прямоугольных распределений (упражнение 11.20) 378
- Мюллер-Лебедева (Myller-Lebedeff V.), разложение в ряды по производным 128 (сноска)
- Наибольшее ожидаемое значение (expected largest value) 456 (сноска)
- Независимость (Independence) 42—43; в случае двумерного биномиального распределения 200—201

- Нейман (Neuman J.), неймановское распределение типа А, получаемое как комбинация двух пуассоновских распределений (упражнение 5.7) 203
- Неполные моменты (incomplete moments) 76; н. м. биномиального распределения (упражнения 5.2—3) 202—203
- Непрерывные распределения (continuous distributions) 23, 29—31; н. р. как результат обращения х. ф. 139—140; выбор из генеральной совокупности с н. р. 304—305
- Неравенства для моментов (inequalities for moments), Ляпунова 93 и (упражнение 3.15) 133; Гаусса — Вилкнера (упражнение 3.18) 134; (упражнение 3.19) 134
- Несмещенные оценки (unbiased estimates) 309, 314; (упражнения 9.13—14) 317
- Несчастные происшествия, распределение водителей лондонских автобусов в соответствии с числом несчастных случаев (таблица 1.21) 47; пуассоновское и отрицательное биномиальное распределение несчастных случаев (таблица 5.3) 183
- Новорожденные, распределение областей по числу н. (таблица 1.1) 17; полигон частот (рис. 1.1) 20
- Нормализация функций плотности (normalization of frequency functions) 229—234; и. ф. п. гамма-распределения (упражнение 6.4) 245
- Нормальное распределение (normal distribution), моменты и п. ф. м. (пример 3.4) 90; семинварианты (пример 3.11) 108—109; обращение х. ф. (пример 4.1) 141—142; как предел биномиального распределения (пример 4.6) 153—154; одностепенность определения по моментам (пример 4.7) 160; как предел распределения Стюдента (пример 4.8) 164—165; как предел пуассоновского распределения (пример 4.9) 166—167; х. ф. квадрата нормальной случайной величины (упражнение 4.18) 170; ф. р. 192—193 и (упражнения 5.12—18) 204—206; таблицы 195—196, среднее отклонение 196; подгонка к данным о ростах (таблица 5.6) 196; моменты вероятностей (упражнение 5.10) 203—204; со случайными параметрами (упражнения 5.24—25) 207; как распределение Пирсона (упражнение 6.1) 245; в центральной предельной теореме 268—270; х. ф. совместного распределения нормальной случайной величины и ее квадрата (упражнение 7.15) 273; случайные нормальные отклонения 299; в теории больших выборок 307—312; стандартная ошибка медианы (пример 10.7) 330—331; стандартная ошибка семинтерквартильной широты (пример 10.8) 332; методы оценки стандартного отклонения исходного распределения 332; моменты среднего отклонения 333—334; моменты средней разности 334—335; дисперсия оценки выборочной дисперсии (упражнение 10.9) 338; распределение суммы квадратов 342—343; (пример 11.6) 353—354, (пример 11.16) 364—365 и (пример 11.18) 366—367; нормальность распределения линейной функции от независимых нормальных случайных величин (пример 11.2) 346—347 и (пример 11.5) 352—353; распределение среднего выборки (пример 11.7) 354—355, (пример 11.12) 361; независимость среднего и дисперсии выборки как характеристическое свойство (пример 11.7) 354—355, (пример 12.7) 401—402, (примеры 15.4—6) 495—496, 499 и 500—503; распределение отношения нормальных случайных величин (пример 11.21) 372—373, (упражнение 11.11) 375—376 и (упражнение 11.22) 378—379; математические ожидания отношений моментов (упражнение 11.5) 374; некоррелированность k -статистик (пример 12.5) 400; выборочные семинварианты k -статистик 409—411; независимость среднего выборки и любой другой k -статистики как характеристическое свойство, некоррелированность среднего выборки и любой другой k -статистики как характеристическое свойство (упражнение 12.20) 423; комбинаторные правила для k -статистик (упражнения 13.10—11) 444—445; моменты медианы и других порядковых статистик 447—450, (пример 14.1) 451—452 и (упражнение 14.14) 475; таблицы моментов порядковых статистик 452; распределение экстремального значения (пример 14.3) 548; распределение размаха и его таблицы 452; дисперсия середины размаха (пример 14.4) 469; дисперсия ближайшего к истинному среднему наблюдаемого значения (упражнения 14.10—11) 474; квадратичные формы от нормальных случайных величин 489—500; характеристизации нормального распределения 500—506; неравенство для интеграла от нормальной плотности (упражнения 15.5) 506; распределения, связанные с нормальным 510—548; отношение между нормальным и связанными с ним распределениями 528—529; распределение отношения случайного отклонения от среднего выборки к стандартному отклонению выборки (упражнение 16.20) 547; см. также Двумерное нормальное распределение, Многомерное нормальное распределение
- Нормированное распределение (standard measure) 74—75; преобразование семинвариантов при переходе к и. р. 100—101; преобразование х. ф. при переходе к и. р. (пример 4.6) 153—165, (пример 4.9) 166—167
- Нэйр (Nair U. S.), моменты средней разности 335
- Облачность, данные об интенсивности о. (таблица 1.11) 26; подгонка с помощью преобразования Джонсона и распределений Пирсона I типа (пример 6.6) 239—240
- Обратные вероятности (inverse probability), см. Байеса теорема
- Обратный выбор для изучения качественных признаков (inverse sampling for attributes) (пример 9.13) 313—314
- Ожидание, среднее значение (expectation, expected value), 79—80
- Ортогональные полиномы Чебышева — Эрмита 217—219; Лагера и Якоби 228—229

- Ортогональные преобразования 346—347; Хельмерта о. п. (пример 11.3) 347—348; характеристическое свойство нормально-го распределения 504—505
- Отношение, стандартная ошибка о. 324; распределение 367—373 и (упражнения 11.11—13, 11.22) 375—377; 378—389; моменты 413—414
- Отрицательное биномиальное распределе-ние (negative binomial distribution) 185—186; как пуассоновское распределе-ние со случайным параметром, имею-щим распределение типа III 184; х. ф. и семинварианты 186; как комбинация пуассоновского и логарифмического распределений (упражнение 5.21) 206; в последовательном выборе для изуче-ния качественных признаков (пример 9.13) 313—314; (упражнение 9.12) 316
- Отрицательное экспоненциальное распре-деление (negative exponential distribu-tion), см. Экспоненциальное распреде-ление
- Оуэн (Owen D. B.), таблицы гипергеомет-рического распределения 191; многомер-ный нормальный интеграл 483
- Оценивание (estimation), при изучении ка-чественных признаков 305—314; в слу-чае больших выборок 312
- Ошибки (errors), см. Стандартные ошиб-ки
- Параметры (parameters) 55
- Парето кривая (Pareto curve) и средняя разность (упражнение 2.19) 82
- Первая предельная теорема (First limit theorem) 149—151; обращение 151—153
- Перевала метод, использование при по-строении аппроксимации 373—374, 471
- Перестановки случайные (random permu-tations) (пример 9.7) 302
- Пирс (Pearse G. E.), данные (таблица 1.11) 26
- Пирс (Pearse T. V.), данные (таблица 1.22) 48
- Пирсон Е. (Pearson E. S.), выборочное распределение $\sqrt{b_1}$ 411—413; выборочное распределение b_2 (упражнение 12.10) 421; распределение размаха 466 н 471; таблицы распределения размаха 466 и 471; таблицы χ^2 512; распределение от-ношения случайного отклонения от среднего к стандартному отклонению выборки (упражнение 16.20) 547
- Пирсон К. (Pearson K.), данные о три-папозомах (таблица 1.13) 28, данные о плодovitости конематок (таблица 1.19) 46; коэффициент вариации 75; поправ-ки к группировке 117; мера асимметрии 124, 210; неравенство чебышевского ти-па 129; данные о распределении карт в висте (таблица 5.5) 191; моменты би-номиального распределения (упражне-ние 5.5) 202; табулирование тетрахори-ческих функций 225; распределение раз-ности случайных переменных III типа (упражнение 11.15) 377; моменты поряд-ковых статистик в нормальном случае 450; разложение вариантов в терминах ее ф. р. 450—451; критерий согласия χ^2 и χ^2 -распределение 492
- Пирсоновские распределения (Pearson di-tributions) как предел гипергеометри-ческих 198; общие сведения 208—216; рекуррентные соотношения между мо-ментами 209; асимметрия 210; критерий для различения типа 212; подгонка 214—216; рекуррентные соотношения для х. ф. и семинвариантов (упражнение 6.7) 246; см. также Типы I—VII
- Письма и конверты, см. Конверты и пись-ма
- Питмэн (Pitman E. J. G.), производные х. ф. и моменты 163; характеристиче-ское свойство гамма-распределения (упражнение 15.22) 509; моменты коэф-фициента корреляции выборки в случае независимых наблюдений (упражнение 6.17) 547
- Плодовитость, распределение конематок в соответствии с п. (таблица 1.19) 46
- Плэкетт (Plackett R. L.), границы для среднего значения размаха 467; много-мерный нормальный интеграл 483
- Подвыборки среднее значение (subsample mean), стандартная ошибка в разности между средними выборки и подвыбор-ки (упражнение 10.20) 339
- Подходящий налог, распределение лиц в соответствии с п. н. 18, 19, 31
- Поля (Pólya G.), Поля—Эйли распреде-ление (упражнение 5.8) 203; неравенство для нормального интеграла (упражне-ние 5.17) 206; нормальная ф. р. (упражне-ние 5.18) 206
- Полигон частот (frequency polygon) 19; двумерный случай 39—41
- Полсон (Paulson E.), нормализация рас-пределения дисперсионного отношения 527—528
- Поправки к группировке (corrections for grouping) 57, 72—73, 111—120, 124 и (упражнения 3.10, 3.13) 132; см. также Шепарда поправки
- Порядковые статистики (order-statistics) 349, 446—476; распределение 319—351; совместное распределение (упражнение 11.7) 50, 446, 463, 446—476; моменты в нормальном случае 447—450; таблицы в нормальном случае 452; асимптотиче-ские распределения 452—461 и (упражне-ние 14.21) 476; верхние границы для ожидаемых значений 462—463; размах и средина размаха 464—471; таблицы рас-пределений крайних значений 479; рек-куррентные соотношения между момен-тами и распределениями 461 н (упражне-ние 14.5) 473; дисперсия (упражнение 14.6) 473
- Последовательный выбор при изучении качественных признаков (sequential sampling for attributes) (пример 9.13) 313—314
- Правдоподобие (likelihood) 276; принцип максимального правдоподобия 280—284; и теорема Байеса 280—283 н (упражне-ние 8.1—2) 284—285; принцип м. п. в случае выбора для изучения качествен-ных признаков (упражнения 9.5, 9.10) 315, 316
- Правило следования Лапласа (succession rule of Laplace) (пример 8.3) 278—279 и (упражнение 8.4) 285
- Предельные теоремы (limit theorems), см. Первая предельная теорема. Вторая предельная теорема. Центральная пре-дельная теорема. Больших чисел са-кон

- Предложения, распределение по их длине (таблица 1.17) 46
- Предсказание урожая, пристрастие (систематическая ошибка) при п. у., (пример 9.4) 292—293
- Преобразования 35—37, 43—45, 342—344; п., уравнивающие дисперсию (упражнение 10.17) 339; ортогональные 346—348; оставляющие порядки величин семинвариантов без изменения (упражнение 13.6) 444; коэффициент корреляции, См. корреляции коэффициент
- Преториус (Pretorius S. T.) данные об австралийских свадьбах (таблица 1.8) 24; данные о бобах (таблица 1.15) 40 н (таблица 6.1); двумерные распределения 223
- Пристрастие (bias) (систематическая ошибка) 289—294; при выборе растений (пример 9.1) 289—290; при считывании со шкал (пример 9.2) 290—291; при считывании результатов случайного эксперимента (пример 9.3) 291—292; в предсказании урожая (пример 9.4) 292—293
- Проблема моментов (problem of moments) (упражнение 3.12) 132; общие сведения 156—161 и 166—167; (упражнения 4.13, 4.17) 169, 170; п. м. и сходность рядов типа А (упражнение 6.9) 246
- Производящая функция моментов (moment-generating functions) (п. ф. м.) 91—93
- Производящая функция частот (frequency-generating function), 45; нахождение среднего (пример 2.3) 58 и дисперсии (пример 2.9) 72
- Простой случайный выбор (simple random sampling) 286—287
- Прямоугольное распределение (uniform distribution), см. Равномерное распределение
- Псевдослучайные числа (pseudo-random numbers) 304
- Пшеницы растения, распределение по высоте (таблицы 9.1) 289
- Пуассоновское распределение (Poisson distribution), производящая функция частот (упражнение 1.13) 52; х. ф. и семинварианты (пример 3.10) 108 и 178—179; моменты (упражнение 3.1) 130; разность двух пуассоновских вариантов (упражнение 3.8) 131, (пример 4.5) 144—146 и (упражнение 11.16) 377; сходимость к нормальному распределению (пример 4.9) 166—167; однозначность определения по моментам (упражнение 4.4) 168; со случайным параметром (упражнение 4.16) 170, 183—188 и (упражнение 5.23) 207; общие сведения 178—185; ф. р. и таблицы 179—180; смесь распределений 180—185; распределение Неймана типа А со случайными параметрами как результат комбинации двух пуассоновских распределений (упражнение 5.7) 203; разбиение совокупности на группы и распределение Пойа — Эйлера (упражнение 5.8) 203, трехмерное (упражнение 5.11) 204; в сочетании с логарифмическим распределением (п. р.) приводит к отрицательно биномиальному распределению (упражнение 5.21) 206; х. ф. сумм случайного числа (с п. р.) случайных величин (упражнение 5.22) 206; факториальные моменты (упражнение 5.26) 207; сумма независимых пуассоновских величин есть пуассоновская величина (упражнение 7.14) 273; при изучении качественных признаков (упражнение 9.5) 315; распределение среднего выборки (пример 11.11) 360; условное в связи с критерием согласия χ^2 (упражнение 15.21) 508—509; распределение временных интервалов (упражнение 16.6) 545
- П. ф. м., производящая функция моментов (moment-generating function) 91—93
- П. ф. с., производящая функция семинвариантов (cumulant-generating function) 100
- П. ф. ф. м., производящая функция факториальных моментов (factorial moment-generating function) 98—99; обращение (упражнение 3.24) 135
- Пэрмэи (Pairman E.), поправки к группировке 117
- Равномерное распределение (rectangular distribution), дискретное (х. ф. и семинварианты (упражнение 3.23) 135, преобразования к р. р. 37, (упражнение 1.21) 53; дисперсия, среднее отклонение и средняя разность (упражнение 2.5) 81; х. ф. и семинварианты (упражнение 3.23) 135; моменты средней разности 335; дисперсия медианы (упражнение 10.15) 339 и (пример 14.2) 452; дисперсия среднего отклонения (упражнение 10.19) 339; точное распределение среднего выборки (пример 11.9) 357—359, (пример 11.14) 361—362; распределение среднего геометрического (пример 11.15) 363—364; распределение произведения равномерно распределенных случайных величин (упражнение 11.6) 374—375 и (упражнение 16.4) 544; распределение отношения наибольших значений (упражнение 11.20) 378; распределение произведения наибольших значений (упражнение 11.21) 378; распределения порядковых статистик — бета-распределение 447; совместное распределение размаха и средины размаха (упражнение 14.12) 474; распределение средины размаха (упражнение 14.13) 474—475; распределение размаха (упражнение 14.16) 475; распределение отношения средины размаха к размаху (упражнение 14.17) 475; распределение разности двух средних размаха (упражнение 14.18) 475—476
- Разложение суммы квадратов (decomposition of sums of squares) 497—500 и (упражнение 15.16—17) 508
- Размах (range) 67—68; точное распределение 464—466 и (упражнение 14.19) 476; границы для математического ожидания 466—467; асимптотическое распределение 467—469 и (упражнение 14.20) 476; редуцированный размах 469—471; существование математического ожидания (упражнение 14.15) 475
- Райдер (Rider P. R.), распределение произведений наибольших значений в выборках из равномерных распределений (упражнение 11.21) 378

- Рао (Rao K. S.), см. Кендалл Д.
- Расположение (location), меры р. 54—67
- Распределение, зависящее от случайных параметров (contagious distribution), см. Типа А распределение
- Рассеяние (dispersion), меры р. 67—79
- Рассеяния коэффициент (concentration coefficient) 75—77; кривая рассеяния 76
- Рассеяния матрица (dispersion matrix 480—481; (упражнения 15.1—3, 15.9—11) 506, 507
- Расширенные симметрические функций (augmented symmetric functions) 382
- Рафф (Raff M. S.), ошибка при аппроксимации биномиального распределения нормальным 178
- Регрессия (regression), линейность в случае двумерного нормального распределения 529—530; коэффициенты 541—542; распределение коэффициентов в случае выборки из генеральной совокупности с двумерным нормальным распределением 542—544; в линейной регрессии одной модели (упражнение 16.5) 544—545; стандартные ошибки коэффициентов р. (упражнения 16.13) 546
- Редуцированные размах и середина размаха (reduced range and midrange) 469—471
- Романовский В., о моментах биномиального распределения (упражнение 5.2) 202
- Рост, распределение мужчин по ростам (таблица 1.7) 23; полигон частот (рис. 1.3) 24; среднее (пример 2.1) 56—58; медиана (пример 2.5) 63; квантили (пример 2.6) 66; кривая распределения (рис. 2.1) 66; среднее отклонение и стандартное отклонение (пример 2.7) 70—71; средняя разность (пример 2.10) 78—79; факториальные и обычные моменты (пример 3.7) 97—98; семинварианты (пример 3.9) 107; подгонка с помощью нормального распределения (таблица 5.6) 197; стандартная ошибка среднего (пример 10.1) 321; стандартные ошибки старших моментов (упражнение 10.1) 337
- Рубин (Ruben H.), моменты наибольших порядковых статистик для выборки из нормальной генеральной совокупности 461; многомерный нормальный интеграл 488
- Сакамото (Sakamoto H.), характеристическое свойство нормального распределения 502
- Самоубийства, данные о с. (таблица 1.6) 22; с. и пуассоновское распределение 183
- Сархан (Sarhan A. E.), таблица дисперсий порядковых статистик в нормальном случае 452
- Сас (Szász O.), условия для х. ф. 147
- Свертка (convolution) случайных величин, см. Сумма с. в.
- Семинвариантов производящая функция (п. ф. с.) (cumulant-generating function) или логарифм характеристической функции 100
- Семинварианты (cumulants), определение 99—100; производящая функция с. 100; свойства инвариантности 100; связь с моментами 101—106; существование 106—107; вычисление 107—110; многомерные 121—124; поправки к группировке 119—120, 124 и (упражнение 3.13) 132; обращение в нуль начиная со второго (пример 4.3) 143, после второго (упражнение 4.11) 169; стандартные ошибки (пример 10.4) 324 и (упражнение 10.12) 338—339; структурные свойства 438—440; см. также *k*-статистики
- Семинтерквартильная широта (с. и. ш.) (semi-interquartile range), см. Интерквартильная широта
- Сёре (Szegő), см. Ортогональные многочлены 228
- Сигнум-функция (sgn *x*) 138
- Симметрические функции (symmetric functions), таблицы 102—103 (сноска), 383; общие сведения 380—390, (упражнение 12.19) 423
- Симметричные распределения (symmetrical distributions), 23—25; действительность х. ф. (упражнение 4.1) 168; нулевая корреляция между средним и четными центральными моментами в случае выборки большого объема (пример 10.3) 322; нулевая корреляция между каждым моментом нечетного порядка и моментом четного порядка в случае выборки большого объема (упражнение 10.6) 338; неотрицательная корреляция между четными моментами выборки (упражнение 10.8) 338; нулевая корреляция между *k*-статистиками нечетного и четного порядка (пример 12.9) 418 и (упражнение 12.13) 422; верхние границы для средних значений порядковых статистик 462—463; распределение редуцированного размаха и редуцированной середины размаха для распределений экспоненциального типа 469—471; граница для дисперсии медианы (упражнение 14.7) 474
- Сичел (Sichel H. S.), моменты вероятностей 244
- Скачок ф. р. (saltus in d. f.) 31, 140
- Случай Е. Е., таблицы χ^2 -распределения 517
- Случайность (randomness) 285—291
- Случайные величины (random variables) 259—265; суммы с. в. 260—261, 267—270; 365; разности с. в. 261—264; преобразование с. в. 342—344
- Случайные нормальные отклонения (random normal deviates) 299
- Случайные перестановки (random permutations), (пример 9.7) 302
- Случайные числа (random sampling numbers) 296—305; таблицы 298—299
- Совместимые события (compatible events), вероятность осуществления (упражнение 7.1—5) 271
- Совокупность (population), см. Генеральная совокупность
- Соломонс (Solomons L. M.), мера асимметрии (упражнение 3.22) 135
- Средина размаха (midrange) 65; распределение в нормальном случае 449—460; сходимость к истинному значению 467—468; дисперсия в нормальном случае (пример 14.4) 469; в случае двойного эк-

- спонциального распределения (пример 14.5) 469; редуцированная средина размаха 469—471, (упражнение 14.4) 473
- Среднее арифметическое (mean arithmetic), определение 55—56; вычисление 56—58; с. а. больше, чем геометрическое и гармоническое средние 59—61; связь с медной и модой 63—65 и (упражнение 6.20) 249; как первый момент 83; стандартная ошибка с. а. 320, 336; распределение среднего выборки, когда истинное распределение является распределением Коши (пример 11.1) 344—346; распределение в случае нормальных совокупностей (пример 11.7) 354—355; независимость от дисперсии выборки в нормальном случае (пример 11.12) 361; распределение в случае равномерного исходного распределения (примеры 11.9, 11.14) 357—359, 361—362; распределение в случае биномиального распределения (пример 11.10) 360; в случае пуассоновского распределения (пример 11.11) 360; в случае распределений III типа (пример 11.13) 361; моменты в случае выборки из конечных совокупностей (пример 12.8) 416—417; некоррелированность с каждой k -статистикой четного порядка (пример 12.9) 418; см. также Центральная предельная теорема
- Среднее гармоническое (mean harmonic) 59—60; с. г. меньше, чем арифметическое и геометрическое среднее 60—61
- Среднее геометрическое (mean geometric) 59—60; соотношение с арифметическим и гармоническим средними 60—61; распределение 362—363; в случае равномерного исходного распределения (пример 11.15) 363—364; в случае распределений III типа (упражнения 11.3, 11.9) 374, 375
- Среднее отклонение (mean deviation) 68; вычисление (пример 2.7) 70—71, (упражнение 2.21) 82; относительно среднего значения не превышает стандартное отклонение 71 (сноска); обращение в минимум относительно медианы (упражнение 2.1) 80; стандартные ошибки 333—334, 336
- Средние значения (mean values) 79—80
- Средняя разность (mean difference) 73—74 и (упражнения 2.4, 2.9, 2.19) 80, 81, 82; вычисление 77—79 и (упражнение 2.10) 81; стандартная ошибка 334—336
- Стандартное отклонение (standard deviation) 68—69; вычисление (пример 2.7) 70—71; связь со средним отклонением и средней разностью 71 (сноска) и (упражнение 2.4) 80; с. о. и поправка Шеннарда (упражнение 2.8) 81; стандартная ошибка с. о. выборки 336; см. также Дисперсия
- Стандартные ошибки (standard errors) 308—313 и 318—340; таблицы 336—337; сравнение с поправками Шеннарда 326—327
- Статистика (statistics), определение 15—16; как множество данных 16
- Статистические гипотезы 281—283
- Статистические таблицы, распределение цифр в с. т. 298
- Степень свободы (degrees of freedom) 373 (сноска), 511, 517, 522
- Стереогрaмма (stereogram) 41
- Стильтес (Stieltjes J.), проблема моментов (упражнение 3.12) 132; 157 (сноска), 159
- Стильтеса интеграл (Stieltjes integral) 34—35
- Стирлинга ряды (Stirling's series), определение 119
- Стуффер (Stouffer S. A.), распределение разности величин, имеющих распределение III типа (упражнение 11.15) 377
- Стюарт (Stuart A.), матрица рассеяния (упражнение 15.9) 507
- Стюдент («Student», W. S. Gossett), смешанное биномиальное распределение (упражнение 5.6) 203; 518; таблица t -распределения 520—521; распределение коэффициента корреляции 535
- Стюдента распределение (Student's distribution), х. ф. (пример 3.13) 109—110; асимптотическая нормальность (пример 4.8) 164—165; вывод (пример 11.8) 355—357 и (пример 11.19) 368; общие сведения 517—521; сходимости к нормальному распределению, х. ф., моменты 518; ф. р. 518—521 и (упражнение 16.10) 546; таблицы 520—521; связь с нормальным χ^2 и F -распределениями 528—529
- Сумма вариант (sum of variates), распределение с. в. 261—264, 267—270, 365—367
- Суммы квадратов (sums of squares), разложение с. к. 497—500 и (упражнение 15.9) 507
- Суммы степеней (power-sums) как симметрические функции 383
- Считывание шкал (scale-reading), систематические ошибки (пристрастие) в с. ш. (пример 9.2) 290—291
- Сэмфорд (Samford M. R.), отношение Миллса (упражнение 5.16) 205
- Сэндфорд (Sandiford P. J.), аппроксимация гипергеометрического распределения биномиальным 191
- Сэндон (Sandon F.), поправки к группировке 117
- Сюй (Hsu C. T.), выборочные семинварианты k -статистик в нормальном случае 411
- Табулирование частотных распределений (tabulation of frequency-distributions) 17—21
- Тейкроу (Teichroew D.), таблица средних значений порядковых статистик в нормальном случае 452
- Телефонная книга, распределение цифр в т. к. (таблица 1.4) 22, 297—298
- Теорема обращения х. ф. (Inversion theorem on c. f. s.) 137—146, (упражнение 4.22) 170
- Тетрахорические функции (tetrachoric functions) 225
- Тиле (Thiele T. N.), семинварианты 101, 280
- Типа A распределение Неймана (type A contagious distribution) (упражнение 5.7) 203

- Типа А, В, С ряды (types A, B, C series), см. Грама-Шарлье ряды
- Типа f (Пирсона) распределение (Type I (Pearson) distribution), поправки к группировке 113—115; общие сведения 210—213; подгонка к данным об облачности (пример 6.6) 239—240; см. также Бета-распределение первого рода
- Типа II (Пирсона) распределение (type II (Pearson) distribution) (упражнение 6.1) 245; распределение среднего (упражнение 11.14) 374
- Типа III (Пирсона) распределение (type III (Pearson) distribution), х. ф. и моменты (пример 3.6) 92—93; семинварианты (пример 3.12) 109; отрицательное биномиальное распределение как пуассоновское распределение со случайным параметром, имеющим распределение типа III 183—185 и (упражнение 5.23) 207; общие сведения 213—214; бесконечность χ^2 -критерия (упражнение 6.5) 245; распределение среднего (пример 11.13) 247—248; распределения средних геометрических (упражнения 11.3, 11.9) 374, 375; распределение разности двух величин, имеющих распределение типа III (упражнение 11.15) 377; см. также Гамма-распределение
- Типа IV (Пирсона) распределение (type IV (Pearson) distribution), моменты (упражнение 3.3) 130; общие сведения 213; подгонка к данным о бобах (пример 6.1) 214—216; таблицы 216
- Типа V (Пирсона) распределение (type V (Pearson) distribution), моменты и семинварианты (упражнение 3.2) 130; равенство единице χ^2 -критерия (упражнение 6.6) 245
- Типа VI (Пирсона) распределение (type VI (Pearson) distribution) 212 и (упражнение 6.4) 245; см. также Бета-распределение второго рода
- Типа VII (Пирсона) распределение (type VII (Pearson) distribution), моменты (пример 3.3) 130; х. ф. (пример 3.13) 132; (упражнение 6.1) 245, см. также Стьюдента распределение
- Типпетт (Tippett L. H. C.), таблицы случайных чисел 298; асимптотическое распределение экстремальных значений 254—257; точное распределение наибольшего значения выборки из нормальной генеральной совокупности 462; распределение размахов выборок из нормальной генеральной совокупности 464—471
- Томпсон (Thompson C.), таблица процентных точек для распределения χ^2 512; таблица процентных точек для t -распределения 521; таблица процентных точек для F -распределения 524
- Точер (Tocher J. F.), данные (таблица 1.24) 50
- Трехмерное пуассоновское распределение (threevariate Poisson distribution) (упражнение 5.11) 209
- Трипанозомы, данные о т. (таблица 1.13) 28
- Тьюки (Tukey J. W.), многомерные k -статистики 418
- Уайз (Wise M.), гипергеометрическая ф. р. 191
- Уикселл (Wickell S. D.), логнормальное распределение 237
- Уильямс (Williams C. B.), логарифмическое распределение (таблица 5.4) 188
- Уильямс (Williams J. L.), неравенство для нормального интеграла (упражнение 15.5) 506
- Уилсон (Wilson E. B.), аппроксимация χ^2 -распределения 513—517
- Уишарт (Wishart J.), семинварианты многомерных мультиномиальных распределений 201; таблицы многомерных k -статистик 418; шаблонная функция (упражнение 12.7) 421; дисперсии и ковариации двумерных k -статистик (упражнение 13.7—8) 444; комбинаторные правила для отыскания k -статистик (упражнение 13.10—12) 444—445
- Унимодальные распределения (unimodal distributions) 54
- Унитарные симметрические функции (unitary symmetric functions) 383
- Урезанные распределения (truncated distributions) 27, 244
- Урожай зерна, данные (таблица 1.20) 47
- Усиленный закон больших чисел (strong law of large numbers) 268
- Условные распределения (conditional distributions) 41—42
- Уэлдон (Weldon W. F. R.), данные о результатах бросания костей (таблица 1.14) 38; (таблица 1.16) 45; (таблица 5.1) 173; (пример 9.10) 308
- Факторные моменты (factorial moments) 93—99; в терминах обычных моментов 95; вычисление 96—98; поправки Шепарда 117—118
- Факторные семинварианты (factorial cumulants), 110—111
- Фармер (Farmer E.), данные (таблица 1.21) 47
- Федериги (Federighi E. T.), таблицы t -распределения 521
- Феллер (Feller W.) «Введение в теорию вероятностей и ее приложения» 270
- Фиеллер (Fieller E. C.), таблицы случайных коррелированных нормальных отклонений 299; распределение отношения коррелированных нормальных величин (упражнение 11.22) 378—379
- Финни (Finney D. J.), распределение отношения дисперсий для коррелированных нормальных величин (упражнение 16.16) 546—547
- Фишер (Sir Ronald A. Fisher), поправка Шепарда (упражнение 3.17) 133—134; логарифмическое распределение 187; нормированное распределение 233; таблицы случайных чисел 299; дисперсия среднего отклонения 333; дисперсия функций от частот мультиномиального распределения (упражнение 10.2) 337—338; дисперсионного отношения F - и χ^2 -распределения, см. Дисперсионное отношение; k -статистики 385; выборочное распределение $\sqrt{b_1}$ 413; шаблонная

- функция (упражнения 12.4—5, 12.7) 420—421; производящая функция для моментов k -статистик (упражнение 12.14) 422; комбинаторные правила для k -статистик в нормальном случае (упражнения 13.10—12) 444—445; асимптотические распределения крайних значений 453—457; таблицы процентных точек распределения χ^2 513, 551; аппроксимация χ^2 -распределения 513—517; таблицы t -распределения 520; таблицы процентных точек F - и z -распределений 524; аппроксимация z -распределения 526—527; распределение коэффициента корреляции 532; z -преобразование коэффициента корреляции 539—541; ф. р. t -распределения (упражнение 16.10) 546
- Фостер (Foster F. G.), матрица рассеяния (упражнение 15.9) 507
- Ф. п., см. Функция плотности (плотность распределения)
- Ф. р., см. Функция распределения
- Франел (Franel J.), распределение цифр в математических таблицах 298 (списка)
- Фреше (Fréchet M.), о соглашениях по поводу средних значений 89 (сноска) и (упражнение 7.19) 274; асимптотическое распределение крайних значений 457
- Фриш (Frisch R.), факториальные моменты 95; семинварианты и моменты биномиального распределения (упражнения 5.1, 5.3) 201, 202
- Функции от случайных величин (functions of random variables), их х. ф. 156; стандартные ошибки 322—325; ковариация 323
- Функция плотности (ф. п.) (frequency function (f. f.)) или плотность 28—33; определение посредством х. ф. 137—141; нормализация ф. п. 229—231
- Функция распределения (ф. р.) (distribution function (d. f.)), 28—33; графики ф. р. для определения квантилей 67; определение по х. ф. 137—141; предельные свойства 147—153; ф. р. и моменты 156—160; двойственность с х. ф. 167; ф. р. и вероятность 259—260
- Фурье преобразование (Fourier transforms и х. ф.) 136, 167
- Характеризации (characterizations) нормального распределения 500—506 и (упражнение 15.16) 508; гамма-распределения (упражнение 15.22) 509
- Характеристическая функция (characteristic function) (х. ф.) как п. ф. м. 91—92; сведения 136—170; связь с ф. р. 136—146 и 167; условия для того, чтобы функция была х. ф. 146—147 и (упражнения 4.5, 4.10—11) 168, 169; предельные свойства 147—153; многомерные х. ф. 154—156; х. ф. функции от вариантов 156; разложение 161—162; производные х. ф. и существование моментов 91—92 и 160—163; отношения между производными (упражнение 4.15) 169; х. ф. суммы 262 и (пример 7.6) 264—265; использование при отыскании точных выборочных распределений 359—365; в методе перевала 373—374
- Харди, Литтлвуд и Полиа (Hardy G. H., Littlewood and Polya) «Неравенства», ИЛ, Москва, 1948
- Хартли (Hartley H. O.), распределение среднего отклонения 334; верхние границы для средних значений порядковых статистик 462—463; таблицы распределений размаха в нормальном случае 466, 471; границы для средних значений размаха 467; таблицы χ^2 -распределения 512; распределение отношения случайного отклонения от среднего выборки к стандартному отклонению выборки (упражнение 16.20) 547
- Хатке (Hatke M. A.), подгонка ф. р. 242—243
- Хелмерт (Helmert W.), дисперсия среднего отклонения 334; ортогональное преобразование (пример 11.3) 347—348
- Хи-квадрат (chi-square), см. χ^2
- Хилфerti (Hilferty M. M.), аппроксимация χ^2 -распределения 513—517
- Хогг (Hogg R. V.), разложение квадратичных форм (упражнение 15.16) 508
- Хойо (Hoio T.), моменты порядковых статистик 449—450
- Холдейн (Haldane J. B. S.), семинварианты биномиального распределения (упражнение 3.1) 130; связь между средним, медианой и модой в разложении Эджворта (упражнение 6.20) 249; аппроксимация χ^2 -распределения (упражнение 16.8) 545
- Холл А. (Hall A. D.), данные относительно пшеничных полей (таблица 1.20) 47
- Холл П. (Hall P.), распределение среднего в выборках из равномерного распределения (пример 11.9) 357—359
- Хотеллинг (Hotelling H.), мера асимметрии (упражнение 3.22) 135; распределение коэффициента корреляции 536—537; преобразования коэффициента корреляции 541 и (упражнение 16.19) 547; соотношения для функций плотности коэффициента корреляции (упражнения 16.14—15) 546
- Х. ф., см. Характеристическая функция
- Хэстингс (Hastings C.), моменты порядковых статистик 450
- Цель, вероятность поражения (упражнение 15.7) 509
- Центр распределения (centre of a distribution) 65, см. также Среднее значение
- Центральная предельная теорема (central limit theorem) 268—270, 312, 325—326; неприменимость к распределению Коши (упражнение 11.1) 374; для равномерного распределения (пример 11.9) 359
- Це-це мух, данные о них (таблица 1.13) 28
- Цифры, распределение в телефонной книге (таблица 1.4) 22
- Чакраварти (Chakravarti M. C.), среднее отклонение и стандартное отклонение 71 (сноска)

Частот полигон (frequency polygon), см.

Полигон частот

Частота (frequency) (класс-частота), 17—18

Частотные распределения (frequency distributions) 15—53; дискретные 21—23; непрерывные 23—28; симметричные 23—25; асимметричные 25; J-образные 25; U-образные 26—27; генезис ч. р. 37—39; многомерные 39—45; условные 41—42; маргинальные (частные) 42

Чебышев П. Л., см. Бьенэме

Чебышева—Эрмита полиномы 217—219

Чебышевского типа неравенства, 128—130

Чеймберс (Chambers E. G.), данные (таблица 1.21) 47

Чу (Chu J. T.), граница для дисперсии медианы (упражнение 14.7) 474

Шаблонная функция (pattern function) в определении выборочных семинвариантов k -статистик 392—401 и (упражнения 12.4—5, 12.7) 420, 421

Шарлье (Charlier C. V. L.), полиномы Чебышева—Эрмита, 218, см. также Грам-Шарлье ряды

Шварца неравенство (Schwarz inequality) 62 (сноска)

Шентон (Shenton L. R.), непрерывные дробь в связи с нормальным распределением 195 и (упражнения 5.14—15) 204—205

Шеппард (Sheppard W. F.) таблицы нормального распределения 195—196; теорема о медианной дихотомии для двумерных нормальных распределений 485

Шеппарда поправки к группировке (Sheppard's corrections for grouping) 72—73; для стандартного отклонения (упражнения 2.8) 81; общие сведения 111—120, 124; (упражнение 3.17) 133; ш. п. к стандартному отклонению и их сравнение с выборочными флуктуациями 325—327

Эджворт (Edgeworth F. Y.), данные Уэлдона (таблица 1.14) 38; форма рядов Грам-Шарлье 222

Эджворта ряды (Edgeworth series) 220—222; связь между средними, медианой и модой (упражнение 6.20) 249; метод перевала в сравнении с разложением Эджворта 374

Эйткин (Aitken A. C.), независимость квадратичных форм в случае коррелированных нормальных случайных величин 499 и (упражнение 15.13) 507

Эйтчисон, Браун (Aitchison J., Brown J. A. C.), логнормальное распределение 237

Эйлера—Маклорена формула суммирования (Euler—MacLaurin sumformula) 112

Экспоненциального типа распределения в теории экстремальных значений (exponential-type distributions in extreme-value theory) 456, 469—471

Экспоненциальное распределение (exponential distribution), среднее, стандартное отклонение, средняя разность и интерквартильная широта (упражнения

2.6) 81; семинварианты (упражнения 3.6) 131; моменты средней разности 335; распределение порядковых статистик 447; дисперсия наименьшего значения (упражнения 14.8) 474; временные интервалы (упражнение 16.6) 545; см. также Гамма-распределение, Двойное экспоненциальное распределение

Экстремальные (крайние) значения (extreme values) 453; асимптотические распределения э. з. 453—461; таблицы 472

Экцесс (kurtosis, excess) 125—126; (упражнения 3.20—1) 134—135; стандартная ошибка b_2 336; моменты b_2 (упражнения 12.10) 421

Элдертон (Sir William P. Elderton), поправки к группировке для моментов 117 и (упражнения 3.10) 132; подгонка распределений Пирсона 214

Элфвинг (Elfvig G.), асимптотическое распределение размаха 468—471 и (упражнения 14.20) 476

Эннс (Anis A. A.), матрица рассеяния многомерного нормального распределения (упражнения 15.11) 507

Эрмита полиномы (Hermite polynomials), см. Чебышева—Эрмита полиномы

Юл (Yule G. V.), данные о длине предположений (таблица 1.17) 46; данные о несчастных случаях на производстве (таблица 5.3) 183; отрицательные биномиальные распределения 184—185; моменты вероятностей нормального распределения (упражнения 5.10) 203—204; пристрастие (систематические ошибки) при считывании со шкал (пример 9.2) 290—291; таблица χ^2 -распределения с одной степенью свободы 513

Юэи (Yuan P. T.), логнормальное распределение 237

Якоби полиномы в разложении по произвольным плотностям бета-распределения 228

Якобиан (детерминант), преобразования 44, 342—344

$b_1, b_2, \beta_1, \beta_2$, см. Асимметрия, Экцесс

γ_1, γ_2 , см. Асимметрия, Экцесс

F-распределение, см. Дисперсионного отношения распределение

J-образные распределения 25—27

κ — величина, служащая критерием при различении типа распределений Пирсона 212; бесконечность κ для III типа (упражнения 6.5) 245; равенство κ единице для V типа (упражнения 6.6) 245

***k*-статистики**, 385—445; определение 385; как семинварианты 385; в терминах симметрических произведений и сумм 388—390; правила для нахождения выборочных семинвариантов 390—401; правила для выборочных семинвариантов в случае нормального исходного распределения (пример 12.5) 400; формулы для выборочных семинвариантов 402—411; формулы для выборочных семинвариантов в нормальном случае 409—411; в выборках из конечной совокупности 414—419; многомерные *k*-статистики для оценивания нелинейных функций от семинвариантов 418—420 и (упражнения 12.18—19) 423; п. ф. м. *k*-статистик (упражнения 12.14—16) 422—423; двумерный случай 424—430; доказательство правил для нахождения выборочных семинвариантов 430—437; структурные свойства 438—440; тензорная запись формул 440—443

***m*-е значения**, см. Экстремальные значения, Порядковые статистики

***t*-распределение**, см. Стьюдента распределение

U-образные распределения 26—27

χ^2 -распределение 510—517; форма 511; х. ф. и моменты, сходимость к нормальному, ф. р. 511—512 и (упражнение 16.7) 545; таблицы 512—513; аппроксимации χ^2 -распределения 513—517 и (упражнение 16.8) 545; связь с нормальным, *t*- и *F*-распределениями 528—529; рекуррентные соотношения для центральных моментов (упражнение 16.3) 544; см. также Гамма-распределение и Типа iii распределение

***z*-распределение**, см. Дисперсионного отношения распределение

***z*-преобразование**, см. Корреляции коэффициент

М. Кендалл, А. Стьюарт

Теория распределений

М., 1966 г., 588 стр. с илл.

Редактор *Л. Н. Большев*

Техн. редактор *К. Ф. Брудно*

Корректор *С. Н. Емельянова*.

Сдано в набор 29/VIII 1966 г. Подписано к печати 8/XII 1966 г. Бумага 60×90^{1/16}. Физ. печ. л. 36,75. Условн. печ. л. 36,75. Уч.-изд. л. 37,47. Тираж 8 000 экз. Цена книги 2 р. 79 к. Заказ № 332.

Издательство «Наука»

Главная редакция физико-математической литературы.

Москва, В-71, Ленинский проспект, 15.

Ленинградская типография № 2

имени Евгении Соколовой

Главполнграфпрома

Комитета по печати

при Совете Министров СССР.

Измайловский проспект, 29.