Московский государственный технический университет
им. Н. Э. Баумана

В. В. Селиверстов, В. П. Обухов

СТАТИСТИКА ЭКСПЕРИМЕНТА В ТЕХНИКЕ И НАУКЕ. ОБРАБОТКА ВЫБОРКИ

Издательство МГТУ
1993

Московский государственный технический университет
им. Н.Э.Баумана

В.В.Селиверстов, В.П.Обухов

СТАТИСТИКА ЭКСПЕРИМЕНТА В ТЕХНИКЕ И НАУКЕ.
ОБРАБОТКА ВЫБОРКИ

Утверждено редсоветом МГТУ
в качестве учебного пособия

Издательство МГТУ
1993

ББК 22.172
С29

Рецензенты: Г.Д.Карташов, З.Н.Салтыкова,
Д.И.Тверетинов, О.И.Тескин

С29 Селиверстов В.В., Обухов В.П. Статистика эксперимента в технике и науке. Обработка выборки: Учеб. пособие. – М.: Изд-во МГТУ, 1993. – 103 с., ил.

ISBN 5-7038-1082-5

В пособии излагаются основы математической статистики в примерах и задачах, рассматриваются основные статистические методы и способы их применения в инженерных расчетах.

Большое количество приведенных задач из различных областей техники (30 вариантов) позволяет эффективно использовать настоящее пособие при организации самостоятельной работы студентов и аспирантов.

Предназначено для инженеров-исследователей, научных работников, а также студентов и аспирантов технических вузов.

Ил. 18. Табл. 5. Библиогр. 12 назв.

ББК 22.172

ISBN 5-7038-1082-5

Елисеев Александр Михайлович

ВВЕДЕНИЕ

В повседневной работе исследователь применяет некоторые приемы математической статистики, возможно, даже не зная об этом, чтобы затем эффективно использовать результаты статистического вывода в качестве прочной основы для последующих выводов и решений.

Цель настоящего пособия – познакомить инженера-исследователя и студентов технических специальностей с основными статистическими методами и научить правильно их применять в инженерных расчетах.

Для лучшего усвоения основных положений математической статистики последние подробно иллюстрируются поэтапным решением (по мере изложения теории) конкретной инженерной задачи: стойкость к разрушению образцов парашютной ткани в процессе их многократного растяжения (примеры 1-16). Такой подход к изучению основных задач математической статистики использован не столько для того, чтобы получился сборник готовых решений (что представляет несомненный интерес для инженеров-практиков), сколько для того, чтобы выделить из общей совокупности и четко разъяснить тот или иной аспект теории статистического вывода.

В пособие входит 30 задач (и заданий к ним), которые взяты из различных областей науки и техники. Такой набор задач несомненно окажется полезным для преподавателей при организации самостоятельной работы студентов и аспирантов и может быть эффективно использован при выполнении курсовых и дипломных проектов.

Предполагается, что читатель имеет подготовку в области классического математического анализа, знаком с элементами теории вероятностей и хочет приобрести практические навыки в теории статистического вывода; специальная подготовка в области математической статистики не требуется. Набор задач из различных областей техники особенно полезен студентам, самостоятельно изучающим методы математической статистики, а также преподавателям вузов.

Глава I. ПЕРВИЧНАЯ МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

§ I. Основные задачи математической статистики. Определения и понятия

Законы теории вероятностей не являются беспредметными абстракциями, лишенными конкретного (например, физического или технического) содержания; они представляют собой математическое выражение фактически существующих реальных закономерностей, характерных для массовых однородных случайных явлений природы.

За случайностями всегда скрывается необходимость, закономерность, которая определяет ход развития в природе и обществе.

Установление, выявление таких закономерностей основано на изучении статистических (опытных) данных и составляет предмет специальной науки – математической статистики, математическим аппаратом которой является теория вероятностей.

В инженерных расчетах установление статистических закономерностей, которым подчинены массовые однородные случайные явления, основано на изучении статистического (опытного) материала. При этом инженером-исследователем наиболее часто ставится одна из следующих задач математической статистики.

З а д а ч а I. Необходимо указать способы сбора статистического материала и последующей его группировки (если опытных данных очень много). Затем выполняется первичная математическая обработка статистического материала – графическое и табличное представление опытных данных.

З а д а ч а 2. Разработать методы анализа статистического материала в зависимости от поставленной инженером-исследователем цели. При этом по опытным данным желательно найти такую статистическую модель – статистическую (эмпирическую) функцию распределения $F^*(x)$ –, которая бы правильно описывала изучаемую статистическую закономерность. Такая функция распределения может зависеть еще и от нескольких параметров $\vartheta \equiv (\alpha_0, \alpha_1, \dots)$, которые в дальнейшем необходимо оценить по результатам наблюдений.

З а д а ч а 3. На практике часто вид статистической модели заранее известен (например, из тех или иных теоретических представлений об изучаемом явлении) и требуется лишь выполнить

статистическую оценку неизвестных параметров.

Этой цели служат специально разработанные в математической статистике методы и приемы, например, метод моментов, метод максимального правдоподобия и др. Эта задача является частным случаем задачи 2.

З а д а ч а 4. На основе статистического материала делается предположение (выдвигается гипотеза) о виде статистической модели. При этом возникает вопрос: насколько верна выдвинутая гипотеза? Ответ на поставленный вопрос дают специально разработанные математические приемы, например, критерий Пирсона, λ – критерий Колмогорова и др. Таков далеко не полный перечень задач математической статистики, которые ставит перед собой инженер-исследователь.

В основе научных знаний лежит наблюдение. Однако единичное наблюдение не отражает общей природы интересующего исследователя явления. Для обнаружения общей закономерности, которой подчиняется явление, необходимо многократно его наблюдать в одинаковых условиях. Теория вероятностей рассматривает те эксперименты, результат которых может меняться от опыта к опыту, причем имеет место статистическая устойчивость эксперимента [I]; теоретическое изучение таких случайных экспериментов и составляет предмет теории вероятностей.

Одним из основных понятий теории вероятностей и математической статистики является случайная величина ("изучаемый признак").

Случайная величина – это переменная величина, которая в результате опыта может принять то или иное значение, причем заранее неизвестно, какое именно. Иными словами можно сказать, что случайная величина – это функция, заданная на множестве исходов данного опыта. Случайные величины условились обозначать последними заглавными буквами латинского алфавита: например, $X, Y, Z, \dots$, а их значения – соответствующими малыми буквами, например, $x, y, z, \dots$.

Если случайная величина принимает только конечное или счетное множество значений на числовой оси, то она называется дискретной случайной величиной. Другими словами каждое значение дискретной случайной величины имеет некоторую ε-окрестность, в которую не попадают другие значения той же случайной величины.

Наоборот, значения непрерывных случайных величин сплошь заполняют некоторый интервал на числовой оси или всю числовую ось.

Примеры случайных величин:

I. Число циклов, которые выдерживают образцы некоторого материала до разрушения.

2. Число обрывов пряжи на 1000 веретен.

3. Разрывная нагрузка стальной проволоки.

4. Температура окружающей среды в красильном цехе в течение суток.

В примерах I, 2 рассмотрены дискретные, а в примерах 3, 4 непрерывные случайные величины.

Пусть требуется изучить совокупность однородных объектов относительно некоторого признака (случайной величины), характеризующего эти объекты. С этой целью иногда проводят сплошное обследование, т.е. обследуют каждый из объектов совокупности относительно признака, которым интересуются. Однако такое обследование применяется сравнительно редко. Если обследование объекта связано с его уничтожением или требует больших материальных затрат, то проводить сплошное обследование практически не имеет смысла. В подобных случаях случайно отбирают из всей совокупности ограниченное число объектов и подвергают их изучению.

Значения, которые приняла случайная величина (изучаемый признак) в процессе обследования отобранного объекта, принято называть наблюдаемыми значениями или реализациями случайной величины (в технической литературе часто используют термин "варианта").

Выборочной совокупностью (или просто выборкой) называют совокупность случайно отобранных объектов.

Генеральной совокупностью называют совокупность объектов, из которых производится выборка.

Выборка называется репрезентативной (предоставительной), если каждый объект генеральной совокупности имеет одинаковый шанс быть обследованным, т.е. имеет одинаковый шанс попасть в выборочную совокупность.

Согласно закону больших чисел [I, 4] выборка будет репрезентативной, если объекты для обследования отбираются случайно (наугад).

Объемом совокупности (выборочной или генеральной) называют число объектов этой совокупности. Например, если из 5000 бобин проволоки отобрано для обследования 100 бобин, то объем генеральной совокупности $N = 5000$, а объем выборки $n = 100$.

На практике, как правило, $n \ll N$, и в целях упрощения вычислений или облегчения теоретических выводов допускают, что генеральная совокупность состоит из бесконечного множества объектов.

§ 2. Графическое и табличное представление статистических данных

Пусть из генеральной совокупности извлечена выборка: пусть x_1 наблюдалось m_1 раз; x_2 наблюдалось m_2 раз и т.д. Причем

$$\sum_i m_i = n - \tag{I}$$

объем выборки.

Наблюдаемые значения случайной величины X ранжируют, нумеруют и записывают в порядке возрастания их номеров. Такая последовательность называется вариационным рядом. Элементы такого ряда называют вариантами (реализациями случайной величины).

Разность $\Delta x = x_{наиб} - x_{наим}$ называют размахом вариационного ряда.

Числа m_i, показывающие сколько раз встречается каждый вариант дискретного признака, называют частотами. При непрерывном изменении признака - это количество вариант, попавших в данный интервал значений признака.

Относительная частота (в технической литературе иногда используют термин "частость") - есть отношение частоты m_i к объему выборки n:

$$w_i = \frac{m_i}{n}. \tag{2}$$

Статистическим распределением выборки называют перечень вариант и соответствующих им частот. В случае непрерывной случайной величины такое распределение можно задать также в виде последовательности интервалов и соответствующих им частот (в качестве частоты, соответствующей интервалу, принимают приходящуюся на единицу длины интервала сумму частот, попавших в этот интервал).

Заметим, что в теории вероятностей под распределением понимают всякое соответствие между возможными значениями случайной величины и их вероятностями, а в математической статистике – соответствие между наблюдаемыми значениями (вариантами) и их частотами или относительными частотами.

В приложениях математической статистики к задачам инженерного содержания часто используют такие понятия, как накопленные частоты или относительные частоты.

Накопленной частотой (накопленной относительной частотой) для варианты x_i называется сумма частот (относительных частот) вариант, не превышающих x_p :

$$S(x_p) = m_1 + m_2 + \ldots + m_p = \sum_{x \le x_p} m_i \,, \qquad (3)$$

$$\tilde{S}(x_p) = w_1 + w_2 + \ldots + w_p = \sum_{x \le x_p} w_i \,. \qquad (4)$$

При больших объемах выборки n не обязательно сохранение точных первоначальных данных; здесь удобным оказывается группировка статистических данных. В процессе группировки первоначальные варианты случайной величины теряются, их роль выполняют представители каждого интервала, например срединное значение интервала. Если число интервалов (классов) выбрать правильно, то сохраняется удовлетворительное общее представление о генеральной совокупности.

Число классов g , на которое следует разбить статистический материал, не должно быть слишком большим, иначе статистический ряд становится невыразительным и частоты в нем обнаруживают случайные колебания, незакономерные для изучаемого явления. С другой стороны, оно не должно быть слишком малым; в этом случае статистический ряд слишком грубо описывает статистические закономерности, присущие изучаемому явлению.

В общем случае число классов, на которое следует разбить данные, т.е. оптимальная длина интервала, зависит как от объема выборки n , так и от размаха Δx статистических данных.

В самом общем случае выбор оптимального интервала подсказывается скорее интуицией и опытом исследователя, чем какими-

либо правилами. Практика статистических выводов в задачах инженерного содержания указывает на целесообразность следующих рекомендаций при выборе оптимальной длины интервала:

I. Обычно берется 8–20 классов группировки [2] в зависимости от объема выборки n :

Объем выборки n	40–60	60–I00	I00–200	200–400	Более 400
Число классов g	5–7	7–I0	I0–I4	I4–I8	I8–20

Иногда оптимальный интервал вычисляется по эмпирической формуле [2]

$$h = h(n, \Delta x) = \frac{\Delta x}{1 + \alpha \lg n} \,, \quad \alpha \in [3{,}21;\, 3{,}24], \qquad (5)$$

где $\Delta x = x_{наиб} - x_{наим}$ – размах; n – объем выборки.

2. Предпочтительно иметь интервалы одинаковой длины. Если же интервалы имеют разную длину, то площади (см. далее определение гистограммы частот) должны быть пропорциональны соответствующим частотам попадания в интервал.

3. Интервалы не должны пересекаться, не должно возникать никаких сомнений относительно того, в какой интервал попадает любое конкретное значение случайной величины.

4. При группировке статистического материала следует находить характеристики каждого класса: границы классов и их срединные значения.

Результаты такой первичной обработки удобно представлять в виде таблиц статистических данных.

Пример I. В табл. I даны числа циклов, которые выдерживают образцы парашютной ткани в процессе их многократного растяжения при заданной циклической деформации I2 %*. Выполнить группировку данных.

Решение. В данной задаче $x_{наим} = 116$; $x_{наиб} = 187$; $n = 216$; $\Delta x = x_{наиб} - x_{наим} = 71$. Согласно вышеприведенным рекомендациям число классов примем равным I8, тогда оптимальный интервал равен $h = \frac{\Delta x}{18} = 71/18 \approx 4{,}0$.

* Далее для краткости будем называть "число циклов растя-

Таблица I

Опытные данные к примеру I

I34	I5I	I62	I55	I38	I44	I58	I57	I35	I34	I6I	I34	I43	I44	I42	I66
I57	I65	I57	I45	I50	I54	I54	I25	I47	I33	I63	I37	I2I	I4I	I52	I37
I65	I6I	I50	I74	I58	I62	I50	I75	I44	I53	$\boxed{I87}$	I45	I64	I52	I50	I74
I63	I24	I42	I52	I83	I4I	I46	I72	I48	I33	I5I	I64	I53	I76	I60	I48
I36	I60	I54	I48	I6I	I49	I47	I6I	I58	I55	I75	I37	I5I	I43	I62	I79
II8	I59	I47	I68	I32	I6I	I54	I29	I32	I47	I58	I57	I56	I46	I46	I57
I36	I6I	I66	I52	I67	I60	I37	I56	I48	I7I	I66	I45	I5I	I73	I39	I55
I44	I46	I53	I62	I84	I45	I6I	I55	I42	I50	I59	I66	I73	I7I	I64	I63
I48	I49	I55	I3I	I39	I42	I56	I49	I63	I49	I64	I42	I36	I43	I48	I59
I46	I53	I54	I44	I72	I50	I43	I50	I69	I57	I64	I52	I65	$\boxed{II6}$	I77	I68
I68	I38	I43	I60	I37	I40	I64	I42	I32	I70	I38	I65	I58	I7I	I43	I72
I36	I6I	I69	I37	I56	I53	I34	I62	I50	I58	I67	I25	I70	I[illegible]3	I70	I52
I54	I45	I54	I34	I56	I53	I46	I38	I73	I79	I69	I67	I74	I57	I56	I35
I47	I27	I70	I65	I45	I80	I47	I54								

В табл. 2 статистические данные табл. I сгруппированы по классам. Во избежание попадания вариант на границы смежных интервалов левая граница первого интервала принята равной II5,5 (вместо $x_{наим}$ = II6).

Используя штриховальный лист (столбец 3, табл. 2), данные табл. I следует разнести по классам. Подсчет численности каждого класса удобно вести по пять, применяя для их обозначения один из символов 卌, ⊠, ⁙. При оформлении штриховального листа следует соблюдать масштаб, а символы в каждом классе следует начинать заносить с одной и той же вертикальной прямой. "Огибающая" правильно оформленного штриховального листа, особенно полезна в дальнейшем при выдвижении гипотезы о виде статистического распределения (см. пример 6).

Наиболее распространенными способами графического представления статистических данных являются **полигон частот**, полигон накопленных частот, гистограмма.

<u>Полигоном частот</u> называют ломаную линию, отрезки которой соединяют точки $(x_1; m_1)$, $(x_2; m_2)$, ..., (x_k, m_k), ..., где x_k – варианты выборочной совокупности, m_k – соответствующие им частоты.

Таблица 2

Первичная статистическая обработка числа циклов (см. пример I)

Границы интервала	Срединное значение интервала x_i	Штриховальный лист	Частота m_i	Относительная частота w_i	Накопленная частота $S(x_p)$	Накопленная относительная частота $\tilde{S}(x_p)$
II5,5–II9,5	II7,5	||	2	0,0092	2	0,009
II9,5–I23,5	I2I,5	|	I	0,0046	3	0,0I4
I23,5–I27,5	I25,5	卌	5	0,023I	8	0,037
I27,5–I3I,5	I29,5	||	2	0,0092	I0	0,046
I3I,5–I35,5	I33,5	卌卌|	II	0,05I0	2I	0,097
I35,5–I39,5	I37,5	卌卌卌|	I6	0,0740	37	0,I7I
I39,5–I43,5	I4I,5	卌卌卌|	I6	0,0740	53	0,245
I43,5–I47,5	I45,5	卌卌卌卌||	22	0,I0I8	75	0,347
I47,5–I5I,5	I49,5	卌卌卌卌|	2I	0,0972	96	0,444
I5I,5–I55,5	I53,5	卌卌卌卌卌||	27	0,I250	I23	0,569
I55,5–I59,5	I57,5	卌卌卌卌|||	23	0,I065	I46	0,676
I59,5–I63,5	I6I,5	卌卌卌卌|	2I	0,0972	I67	0,773
I63,5–I67,5	I65,5	卌卌卌|||	I8	0,0833	I85	0,856
I67,5–I7I,5	I69,5	卌卌|	II	0,05I0	I96	0,907
I7I,5–I75,5	I73,5	卌卌|	II	0,05I0	207	0,958
I75,5–I79,5	I77,5	卌	5	0,023I	2I2	0,98I
I79,5–I83,5	I8I,5	||	2	0,0092	2I4	0,99I
I83,5–I87,5	I85,5	||	2	0,0092	2I6	I,000

Таблица 2

Первичная статистическая обработка числа циклов (см. пример I)

Границы интервала	Срединное значение интервала x_i	Штриховальный лист	Частота m_i	Относительная частота w_i	Накопленная частота $S(x_p)$	Накопленная относительная частота $\widetilde{S}(x_p)$
II5,5–II9,5	II7,5	‖	2	0,0092	2	0,009
II9,5–I23,5	I2I,5	|	I	0,0046	3	0,0I4
I23,5–I27,5	I25,5	卌	5	0,023I	8	0,037
I27,5–I3I,5	I29,5	‖	2	0,0092	I0	0,046
I3I,5–I35,5	I33,5	卌 卌 |	II	0,05I0	2I	0,097
I35,5–I39,5	I37,5	卌 卌 卌 |	I6	0,0740	37	0,I7I
I39,5–I43,5	I4I,5	卌 卌 卌 |	I6	0,0740	53	0,245
I43,5–I47,5	I45,5	卌 卌 卌 卌 ‖	22	0,I0I8	75	0,347
I47,5–I5I,5	I49,5	卌 卌 卌 卌 |	2I	0,0972	96	0,444
I5I,5–I55,5	I53,5	卌 卌 卌 卌 卌 ‖	27	0,I250	I23	0,569
I55,5–I59,5	I57,5	卌 卌 卌 卌 ‖|	23	0,I065	I46	0,676
I59,5–I63,5	I6I,5	卌 卌 卌 卌 |	2I	0,0972	I67	0,773
I63,5–I67,5	I65,5	卌 卌 卌 ‖|	I8	0,0833	I85	0,856
I67,5–I7I,5	I69,5	卌 卌 |	II	0,05I0	I96	0,907
I7I,5–I75,5	I73,5	卌 卌 |	II	0,05I0	207	0,958
I75,5–I79,5	I77,5	卌	5	0,023I	2I2	0,98I
I79,5–I83,5	I8I,5	‖	2	0,0092	2I4	0,99I
I83,5–I87,5	I85,5	‖	2	0,0092	2I6	I,000

Для рассмотренного выше примера полигон частот представляет собой "многоугольник" с вершинами в точках, соответствующих срединным значениям интервалов и их частотам. Как показано на рис. I, вершины имеют координаты (II3,5; 0), (II7,5; 2), (I2I,5; I), (I25,5; 5), ...

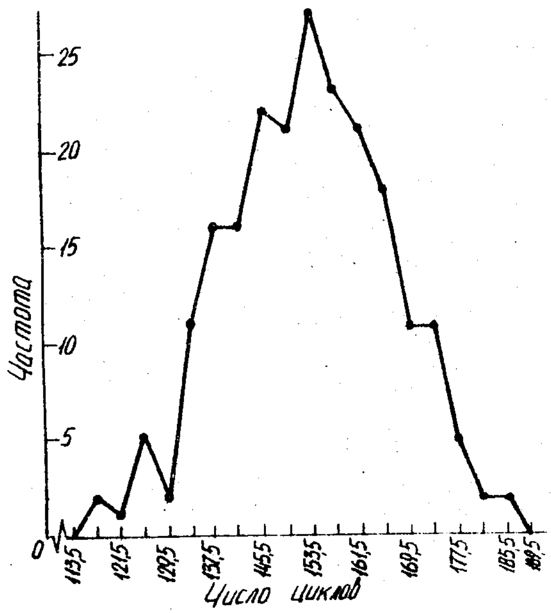


Рис. I. Полигон частот числа циклов растяжения

Обратим внимание на следующие два обстоятельства при построении полигона частот:

I. Излом оси абсцисс в самом ее начале. Это означает, что если оси пересекаются в точке (0; 0), то ось абсцисс обрывается до области наших данных.

2. Полигон частот должен быть замкнутой фигурой. Поэтому дополняем первый столбец табл. 2 двумя такими же интервалами, прилежащими к крайним интервалам.

Затем находим срединные значения этих интервалов, которым соответствуют равные нулю частоты (на рис. I соответственно точки (II3,5; 0) и (I89,5; 0).

Аналогично определяется и полигон относительных частот – ломаная линия, отрезки которой соединяют точки (x_1; w_1), (x_2, w_2), ... , (x_k, w_k), где x_k – варианты выборки; w_k – соответствующие им относительные частоты.

В технике часто используются (например, при контроле за ходом технологического процесса) кумулятивные кривые – полигоны накопленных частот. Вершины такого полигона имеют координаты, соответствующие верхней границе интервала и накопленной частоте этого интервала (рис. 2).

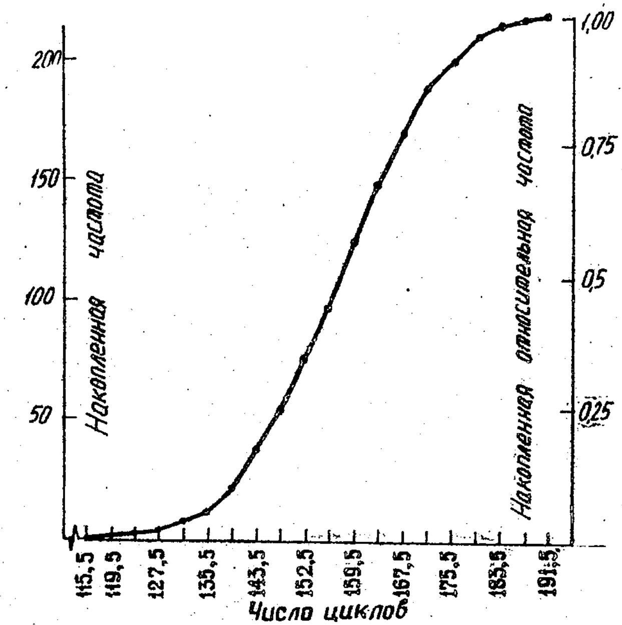


Рис. 2. Полигон накопленных частот числа циклов растяжения

Если случайная величина непрерывна или статистический материал сгруппирован, целесообразнее строить гистограмму, для чего интервал, в котором заключены все наблюдаемые значения случайной величины, разбивают на несколько частичных интервалов (классов) длиной h и находят для каждого класса сумму частот вариант m_i , попавших в i-й интервал (численность класса).

Если статистический материал сгруппирован, то за h удобно брать значение интервала группировки.

Гистограммой частот называют фигуру, состоящую из прямоугольников, основаниями которых служат интервалы длиной h, а высоты равны отношениям m_i/h (плотность частоты).

Площадь i-го интервала есть $m_i \cdot h \cdot \frac{m_i}{h}$, следовательно, площадь гистограммы частот равна сумме всех частот, т.е. объему выборки n (рис. 3).

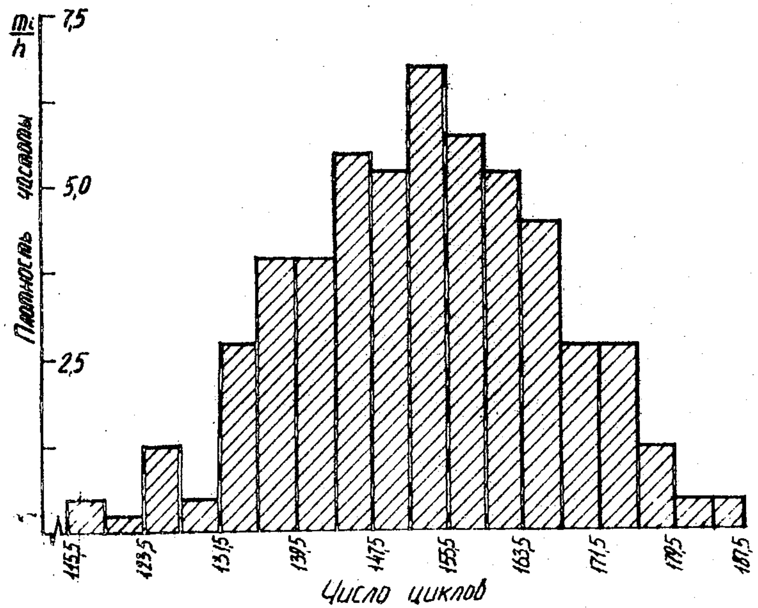


Рис. 3. Гистограмма распределения числа циклов растяжения образцов парашютной ткани

Аналогично гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат интервалы длиной h, а высоты равны отношениям w_i/h (плотность относительной частоты).

Очевидно, что площадь гистограммы относительных частот равна сумме всех относительных частот, т.е. равна единице.

§ 3. Статистическая функция распределения

Другим способом обработки выборки является построение статистической (эмпирической) функции распределения.

Введем обозначения: m_x – число наблюдений, при которых наблюдалось значение признака меньше x; n – общее число наблюдений (объем выборки). Тогда относительная частота события $X < x$ равна m_x/n. Если x будет изменяться, то, вообще говоря, будет изменяться и относительная частота, т.е. отношение m_x/n есть функция от x.

Итак, по определению

$$F^*(x) = \frac{m_x}{n} \equiv \frac{1}{n}\sum_{x_i < x} m_i . \qquad (6)$$

В отличие от статистической функции распределения (6) выборки интегральную функцию $F(x)$ распределения генеральной совокупности называют теоретической функцией распределения.

Различие между $F^*(x)$ и $F(x)$ состоит в том, что теоретическая функция распределения $F(x)$ определяет вероятность события; статистическая функция распределения $F^*(x)$ определяет относительную частоту того же события.

Статистическая функция распределения $F^*(x)$ обладает теми же свойствами [1], что и $F(x)$; она позволяет вычислить относительную частоту попадания случайной величины в заданный интервал:

$$w(x_1 \leqslant X < x) = F^*(x_2) - F^*(x_1) . \qquad (7)$$

Глава II. ВЫБОРОЧНЫЕ ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ

§ 1. Выборочные характеристики положения (выборочные средняя, медиана, мода)

Для вариационного ряда выборочная средняя есть средняя арифметическая всех вариант:

$$M^*X = \bar{x}_b = \frac{1}{n}\sum_{i=1}^{n} x_i . \qquad (8)$$

Если же статистические данные сгруппированы, то

$$M^*X = \bar{x}_в = \frac{1}{n}\sum_{i=1}^{g} x_i m_i = \sum_{i=1}^{g} x_i w_i, \tag{9}$$

где m_i – частота; w_i – относительная частота; n – объем выборки; g – число классов группировки.

Для каждой числовой характеристики случайной величины теории вероятностей существует аналогия в математической статистике. Например, для математического ожидания в теории вероятностей аналогией является выборочная средняя. Поэтому здесь и ниже такие аналогии будем обозначать теми же буквами, что и соответствующие числовые характеристики в теории вероятностей, снабжая их значком $*$.

Выборочная средняя есть фактически абсцисса центра тяжести гистограммы.

Выборочная средняя обладает теми же свойствами, что и математическое ожидание в теории вероятностей:

$$M^*[cX] = cM^*X, \quad M^*[X+c] = M^*X + c, \quad \text{где } c = \text{const}.$$

Из первых двух свойств следует, что

$$\bar{x}_в = M^*X = \frac{1}{n}\sum_{i=1}^{g} x_i m_i = \frac{k}{n}\sum_{i=1}^{g}\frac{x_i - c}{k} + c, \tag{10}$$

где k и c – произвольные постоянные.

При больших объемах выборки n или при больших значениях вариант определение выборочной средней по формуле (9) часто приводит к громоздким вычислениям. Поэтому переходят к так называемым условным вариантам

$$x_i = k u_i + c, \qquad u_i = \frac{x_i - c}{k}, \tag{11}$$

причем, согласно (10), имеем

$$\bar{x}_в = M^*X = kM^*U + c, \tag{10'}$$

где $M^*U = \frac{1}{g}\sum_{i=1}^{g} u_i$.

Постоянные c и k выбирают так, чтобы максимально упростить вычисления. В качестве c обычно принимается та варианта, которая занимает срединное положение в вариационном ряду или которая имеет наибольшую частоту. За k следует брать наибольший общий делитель разностей $x_i - c$, $i = \overline{1, n}$. В случае сгруппированного статистического материала таким наибольшим общим делителем будет значение интервала группировки h.

Например, в рассмотренном примере 1 целесообразно принять $c = 153{,}5$; $k = 4{,}0$. Значение c часто называют ложным (условным) нулем.

Медиана m_e представляет собой такое значение варианты $x = m_e$, которое делит вариационный ряд на две равные по числу вариант части. Медиана делит площадь гистограммы пополам.

В теории вероятностей [1] часто используют понятие квантиля уровня p – это такое значение x_p случайной величины X, левее которого случайная величина X находится с вероятностью p, т.е.

$$P(X < x_p) = \int_{-\infty}^{x_p} f(x)\,dx = p. \tag{12}$$

Иными словами, x_p – корень уравнения

$$F(x) = p. \tag{12'}$$

В этих терминах m_e есть выборочный квантиль уровня 1/2.

Чтобы определить медиану m_e по несгруппированным статистическим данным, находят срединное значение этих данных, если объем выборки n – нечетное число. Если же объем выборки n – четное число, например $n = 2r$, то за медиану m_e можно принять среднее арифметическое

$$m_e = \frac{1}{2}[x_r + x_{r+1}], \quad r = 1, 2, \ldots.$$

В случае сгруппированных статистических данных медиану m_e находим путем линейной интерполяции, используя полигон накопленных частот.

Если обозначить через $S(x_i) = \sum_{x < x_i} m_i$ накопленную частоту при $X = x_i$, то формула линейной интерполяции для двух точек $(x_1; S(x_1))$, $(x_2; S(x_2))$ и текущей точки $(m_e; S(m_e))$ имеет вид

$$m_e = x_1 + (x_2 - x_1)\,\frac{S(m_e) - S(x_1)}{S(x_2) - S(x_1)}\,. \qquad (13)$$

Пример 2. Найти медиану числа циклов растяжения, которые выдерживают образцы ткани (см. пример I).

Решение. В данном случае при $n = 216$ требуется найти такое значение варианты, при котором накопленная частота равна

$$S(m_e) = \frac{n}{2} = \frac{216}{2} = 108\,.$$

Из табл. 2 видно, что при $x_1 = 149{,}5$ накопленная частота равна $S(x_1) = 96$; при $x_2 = 153{,}5$ $S(x_2) = 123$. Полагая, что на отрезке $x \in [149{,}5;\ 153{,}5]$ увеличение накопленных частот происходит по линейному закону, т.е. согласно (13), получаем $m_e \approx 153{,}3$.

Модой m_0 дискретной случайной величины называется наиболее часто встречающаяся варианта.

Если статистические данные сгруппированы и построена гистограмма, то за моду m_0 можно приближенно принять середину интервала с наибольшей относительной частотой, т.е. абсциссу середины наиболее высокой ступеньки гистограммы.

В рассмотренном выше примере I за моду m_0 случайной величины, т.е. числа циклов растяжения можно принять $m_0 = 153{,}5$ (см. рис. 3). Вычисление выборочной средней приведем ниже (см. пример 3), $M^*X = \bar{x}_в = 153{,}0$, а здесь обратим внимание лишь на то, что $\bar{x}_в \approx m_e \approx m_0$. В общем же случае несимметричных распределений для этих выборочных характеристик положения справедливо [4]: $m_0 < m_e < \bar{x}_в$ – для правосторонних распределений, $\bar{x}_в < m_e < m_0$ – для левосторонних распределений.

§ 2. Выборочные характеристики рассеяния

Выборочной дисперсией $\sigma_в^2 = D^*X$ называют среднее арифметическое квадратов отклонения вариант от их выборочной средней $\bar{x}_в$.

Для несгруппированных статистических данных выборочную дисперсию определяют по формуле

$$D^*X = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_в)^2\,. \qquad (14)$$

Если же данные сгруппированы, то

$$D^*X = \frac{1}{n}\sum_{i=1}^{g} m_i (x_i - \bar{x}_в)^2\,. \qquad (15)$$

Корень квадратный из выборочной дисперсии называют выборочным средним квадратическим отклонением: $\sigma_в = \sqrt{D^*X}$.

Выборочная дисперсия обладает теми же свойствами, что и дисперсия в теории вероятностей:

$$D^*C = 0; \qquad D^*[CX] = C^2 D^*X\,, \quad C = \text{const}\,; \qquad (16)$$

$$D^*X = M^*X^2 - (M^*X)^2\,; \qquad (17)$$

$$D^*X = M^*[X+C]^2 - (M^*[X+C])^2\,; \qquad (18)$$

$$D^*[X+C] = D^*X\,, \quad C = \text{const}\,. \qquad (19)$$

Трудоемкость вычислений выборочной дисперсии (при больших n или больших значениях вариант) значительно уменьшается, если от вариант x_i перейти к условным вариантам (11):

$$u_i = \frac{x_i - C}{k}\,; \quad x_i = k u_i + C\,;$$
$$M^*X = kM^*U + C; \quad D^*X = k^2 D^*U\,. \qquad (20)$$

Последующие вычисления выборочных характеристик удобно проводить, например, методом произведений (или методом сумм) [4],[5]; методика таких вычислений использована в примере 3.

Пример 3. Вычислить методом произведений выборочную среднюю, выборочную дисперсию числа циклов растяжения (см. пример I).

Решение. Составим расчетную табл. 3. Для этого:

I. Запишем срединные значения интервалов x_i в первый столбец, а соответствующие им частоты – во второй; сумму частот (216) поместим в нижнюю клетку второго столбца.

Таблица 3

Расчет методом произведений выборочных характеристик

C = I53,5; k = 4,0

Срединное значение интервала x_i	Частота m_i	Условные варианты u_i	Условные моменты $m_i u_i$	$m_i u_i^2$	Контроль вычислений $m_i (u_i+1)^2$
II7,5	2	-9	-I8	I62	I28
I2I,5	I	-8	-8	64	49
I25,5	5	-7	-35	245	I80
I29,5	2	-6	-I2	72	50
I33,5	II	-5	-55	275	I76
I37,5	I6	-4	-64	256	I44
I4I,5	I6	-3	-48	I44	64
I45,5	22	-2	-44	88	22
I49,5	2I	-I	-2I	2I	0
I53,5	27	0	0	0	27
I57,5	23	I	23	23	92
I6I,5	2I	2	42	84	I89
I65,5	I8	3	54	I62	288
I69,5	II	4	44	I76	275
I73,5	II	5	55	275	396
I77,5	5	6	30	I80	245
I8I,5	2	7	I4	98	I28
I85,5	2	8	I6	I28	I62
Итого	2I6	–	-27	2453	26I5

2. В качестве ложного нуля C выберем срединное значение I53,5, которому соответствует наибольшее значение частоты 27. Затем в формулах (20) примем k равным 4,0 (интервал группировки).

3. Согласно (20) вычисляем условные варианты

$$u_1 = \frac{117{,}5-153{,}5}{4{,}0} = -9; \quad u_2 = \frac{121{,}5-153{,}5}{4{,}0} = -8 \text{ и т.д.}$$

Результаты помещаем в третий столбец. Затем вычисляем произведения частот m_i и условных вариант u_i и записываем в четвертый столбец; их сумму (-27) помещаем в нижнюю клетку столбца. Следует обратить внимание на то, что при использовании данной методики условные варианты являются целыми числами. Условная варианта равна нулю для ложного нуля C = I53,5; выше ложного нуля располагаются в порядке убывания отрицательные числа, ниже – натуральные числа в порядке их возрастания.

4. Произведения $m_i u_i^2$ запишем в пятый столбец; сумму чисел столбца (2453) помещаем в нижнюю клетку.

5. Для контроля вычислений воспользуемся следующим очевидным тождеством:

$$\sum_{i=1}^{g} m_i (u_i+1)^2 = \sum_{i=1}^{g} m_i u_i^2 + 2\sum_{i=1}^{g} m_i u_i + n. \qquad (2I)$$

С целью контроля произведения частот, увеличенных на единицу, и соответствующих условных вариант в квадрате, т.е. $m_i (u_i+1)^2$, запишем в шестой столбец; сумму чисел столбца (26I5) помещаем в нижнюю клетку.

Таким образом, согласно тождеству (2I), имеем

$$2615 = 2453-2\cdot27+216 = 2615.$$

Совпадение контрольных сумм свидетельствует о правильности вычислений.

6. Вычислим M^*U и M^*U^2, последовательно имеем

$$M^*U = \frac{1}{n}\sum_{i=1}^{g} m_i u_i = \frac{-27}{216} = -0{,}125;$$

$$M^*U^2 = \frac{1}{n}\sum_{i=1}^{g} m_i u_i^2 = \frac{2453}{216} = 11{,}3565.$$

7. По формулам (20) для выборочной средней и выборочной дисперсии имеем:

$$M^*X = \bar{x}_в = 4{,}0\cdot(-0{,}125) + 153{,}5 = 153{,}0;$$

$$D^*X = \sigma_в^2 = (4{,}0)^2\cdot(11{,}3565 - (-0{,}125)^2) = 181{,}4537.$$

Таким образом, $\bar{x}_в = 153,0$; $\sigma_в^2 = 181,4537$; $\sigma_в = 13,4705$.
Для коэффициента вариации $V_в$, определяемого отношением

$$V_в = \frac{\sigma_в}{x_в} \cdot 100\,\%, \tag{22}$$

имеем $V_в = 8,8\,\%$.

Заметим, что для исходных статистических данных (несгруппированных) среднее арифметическое равно 153,2; среднее квадратическое отклонение составляет 13,25.

§ 3. Выборочные начальные и центральные моменты. Метод произведений

Выборочная средняя и выборочная дисперсия являются частными случаями более общего понятия – статистических моментов выборки.

Начальным выборочным моментом ν_r^* порядка r называется среднее значение r-х степеней значений x_i :

$$\nu_r^* \equiv M^*X^r = \frac{1}{n}\sum_{i=1}^{g} m_i x_i^r, \tag{23}$$

где $r = 0, 1, 2, \ldots$. Частные случаи: 1) $\nu_0^* = 1$, 2) $\nu_1^* \equiv M^*X = \frac{1}{n}\sum_{1}^{g} m_i x_i = \bar{x}_в$ – выборочная средняя.

Центральным выборочным моментом μ_r^* порядка r называется среднее значение r-х степеней разностей $x_i - M^*X$:

$$\mu_r = M^*(X - MX)^r = \frac{1}{n}\sum_{i=1}^{g}(x_i - M^*X)^r m_i, \tag{24}$$

где $r = 1, 2, \ldots$. Частные случаи: 1) $\mu_1 = M^*(X - M^*X) = 0$, 2) $\mu_2 = M^*(X - MX)^2 = D^*X$ – выборочная дисперсия.

В самом общем случае центральные моменты можно выразить через начальные моменты:

$$\begin{aligned}
\mu_1^* &= 0;\\
\mu_2^* &= \nu_2^* - (\nu_1^*)^2;\\
\mu_3^* &= \nu_3^* - 3\nu_1^*\nu_2^* + 2(\nu_1^*)^3;\\
\mu_4^* &= \nu_4^* - 4\nu_1^*\nu_3^* + 6(\nu_1^*)^2\nu_2^* - 3(\nu_1^*)^4;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned} \tag{25}$$

Докажем, например, третье равенство: по формуле (24) имеем

$$\mu_3^* = M^*(X - 3M^*X \cdot X^2 + 3(M^*X)^2 \cdot X - (M^*X)^3).$$

Отсюда, используя свойства знака математического ожидания M и учитывая, что $M^*X = \mathrm{const}$, получим третье равенство в формулах (25).

Вычисление центральных моментов удобно также проводить методом произведений, вводя в рассмотрение, как обычно, условные варианты (20).

Подставив (20) в (25) и выполнив преобразования, получим:

$$\begin{aligned}
\mu_{1x}^* &= 0;\\
\mu_{2x}^* &= k^2[\nu_{2u}^* - (\nu_{1u}^*)^2];\\
\mu_{3x}^* &= k^3[\nu_{3u}^* - 3\nu_{1u}^*\nu_{2u}^* + 2(\nu_{1u}^*)^3];\\
\mu_{4x}^* &= k^4[\nu_{4u}^* - 4\nu_{1u}^*\nu_{3u}^* + 6(\nu_{1u}^*)^2\nu_{2u}^* - 3(\nu_{1u}^*)^4].
\end{aligned} \tag{26}$$

Здесь моменты, относящиеся к первоначальным x_i и условным u_i вариантам, помечены соответственно индексами x и u. Зависимости (26) удобны тем, что с их помощью удается избежать громоздких вычислений при переходе от условных вариант u_i к первоначальным x_i.

Пример 4. Вычислить центральные моменты первых четырех порядков числа циклов, которые выдерживают образцы парашютной ткани в процессе их многоциклового растяжения (см. примеры 1–3).

Решение. При вычислении методом произведений первых четырех начальных моментов ν_r^* составляем расчетную табл. 4, при оформлении которой следуем методике, подробно описанной в примере 3 (см. табл. 3).

Для контроля вычислений воспользуемся тождеством

$$\sum_{i=1}^{g} m_i(u_i+1)^4 = \sum_{i=1}^{g} m_i u_i^4 + 4\sum_{i=1}^{g} m_i u_i^3 + 6\sum_{i=1}^{g} m_i u_i^2 + 4\sum_{i=1}^{g} m_i n_i + n. \quad (27)$$

В данном примере (27) принимает вид:

$$83867 = 75437-4\cdot I599+6\cdot 2453-4\cdot 27+2I6 = 83867,$$

что свидетельствует о правильности проведенных вычислений (см. табл. 4).

Затем необходимо полученные результаты пересчитать по формулам (26), (20) в терминах первоначальных вариант x_i.

На практике наибольший интерес представляют первые четыре центральных момента, которые легко найти по формулам (26) без указанного предварительного пересчета.

Для рассматриваемого примера имеем:

$$\mu^*_{1x} = 0;$$

$$\mu^*_{2x} = D^*X = (4{,}0)^2\left[\frac{2453}{216} - \left(\frac{-27}{216}\right)^2\right] = 181{,}4537;$$

$$\mu^*_{3x} = (4{,}0)^3\left[-\frac{1599}{216} - 3\frac{-27}{216}\cdot\frac{2453}{216} + 2\left(\frac{-27}{216}\right)^3\right] = -201{,}4722; \quad (28)$$

$$\mu^*_{4x} = (4{,}0)^4\left[\frac{75437}{216} - 4\,\frac{-27}{216}\cdot\frac{-1599}{216} + 6\left(\frac{-27}{216}\right)^2\frac{2453}{216} - 3\left(\frac{-27}{216}\right)^4\right] = 88731{,}63$$

Как правило, на практике приходится иметь дело со случайными величинами X, имеющими размерность; их числовые характеристики также являются размерными величинами.

Этот факт создает определенные неудобства при последующем анализе числовых характеристик статистической совокупности, в особенности при дальнейшем сравнении числовых характеристик случайных величин разной размерности. По этой причине вместо μ^*_{3x} и μ^*_{4x} часто рассматривают относительные характеристики: коэффициенты асимметрии A^*_x и эксцесса $Э^*_x$, определяемые по формулам

$$A^*_x = \frac{\mu^*_{3x}}{(D^*X)^{3/2}} = \frac{\mu^*_{3x}}{\sigma_в^3}; \quad (29)$$

$$Э^*_x = \frac{\mu^*_{4x}}{(D^*X)^2} - 3 = \frac{\mu^*_{4x}}{\sigma_в^4} - 3. \quad (30)$$

Таблица 4

Расчет методом произведений сводных выборочных характеристик

$C = I53{,}5; \quad k = 4{,}0$

Срединное значение интервала x_i	Частота m_i	Условные варианты u_i	Условные моменты $m_i u_i$	$m_i u_i^2$	$m_i u_i^3$	$m_i u_i^4$	Контроль вычислений $m_i(u_i+1)^4$
II7,5	2	-9	-I8	I62	-I458	I3I22	8I92
I2I,5	I	-8	-8	64	-5I2	4096	2I0I
I25,5	5	-7	-35	245	-I7I5	I2005	6480
I29,5	2	-6	-I2	72	-432	2592	I250
I33,5	II	-5	-55	275	-I375	6875	28I6
I37,5	I6	-4	-64	256	-I024	4096	I296
I4I,5	I6	-3	-48	I44	-432	I296	256
I45,5	22	-2	-44	88	-I76	352	22
I49,5	2I	-I	-2I	2I	-2I	2I	0
I53,5	27	0	0	0	0	0	27
I57,5	23	I	23	23	23	23	368
I6I,5	2I	2	42	84	I68	336	I70I
I65,5	I8	3	54	I62	486	I458	4608
I69,5	II	4	44	I76	704	28I6	6875
I73,5	II	5	55	275	I375	6875	I4256
I77,5	5	6	30	I80	I080	6480	I2005
I8I,5	2	7	I4	98	686	4802	8I92
I85,5	2	8	I6	I28	I024	8I92	I3I22
Итого	2I6	-	-27	2453	-I599	75437	83867

Таблица 4

Расчет методом произведений сводных выборочных характеристик

$C = 153{,}5; \quad k = 4{,}0$

Срединное значение интервала x_i	Частота m_i	Условные варианты u_i	Условные моменты				Контроль вычислений
			$m_i u_i$	$m_i u_i^2$	$m_i u_i^3$	$m_i u_i^4$	$m_i (u_i+1)^4$
II7,5	2	–9	–I8	I62	–I458	I3I22	8I92
I2I,5	I	–8	–8	64	–5I2	4096	2I0I
I25,5	5	–7	–35	245	–I7I5	I2005	6480
I29,5	2	–6	–I2	72	–432	2592	I250
I33,5	II	–5	–55	275	–I375	6875	28I6
I37,5	I6	–4	–64	256	–I024	4096	I296
I4I,5	I6	–3	–48	I44	–432	I296	256
I45,5	22	–2	–44	88	–I76	352	22
I49,5	2I	–I	–2I	2I	–2I	2I	0
I53,5	27	0	0	0	0	0	27
I57,5	23	I	23	23	23	23	368
I6I,5	2I	2	42	84	I68	336	I70I
I65,5	I8	3	54	I62	486	I458	4608
I69,5	II	4	44	I76	704	28I6	6875
I73,5	II	5	55	275	I375	6875	I4256
I77,5	5	6	30	I80	I080	6480	I2005
I8I,5	2	7	I4	98	686	4802	8I92
I85,5	2	8	I6	I28	I024	8I92	I3I22
Итого	2I6	–	–27	2453	–I599	75437	83867

Пример 5. Используя результаты примера 4, вычислить коэффициенты асимметрии и эксцесса.

Решение. По формулам (29), (30) имеем:

$$A_x = \frac{-201,4722}{(13,4705)^3} = -0,0824; \qquad Э_x = \frac{88731,63}{(181,4537)^2} - 3 = -0,3051.$$

Глава III. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЙ

§ 1. Подбор распределения по статистическим данным

Пусть в результате эксперимента получена выборка объема n. Фактически мы не знаем точного закона распределения (теоретического распределения) изучаемой генеральной совокупности, из которой извлечена выборка. В этом случае разумно поставить вопрос: "Оправдано ли наше предположение, основанное на анализе статистических данных (выборки), о том, что имеет место тот конкретный вид распределения исследуемой случайной переменной, которому удовлетворяет наша выборка?".

Исследуя выборку, можно увидеть, во-первых, допускает ли она распределение некоторого вида. И, во-вторых, оправдано ли наше предположение о виде распределения в дальнейших исследованиях.

Пример 6. На основе анализа статистических данных, проведенного в гл. I, II, подобрать вид распределения для числа циклов, которые выдерживают образцы парашютной ткани в процессе их многоциклового растяжения (см. пример 1).

Решение. Проведенный статистический анализ указывает на следующее:

1. Выборочные средняя, мода и медиана примерно равны, что является характерным для симметричных распределений.

2. Коэффициенты асимметрии и эксцесса относительно малы (см. пример 5).

3. На гистограмму частот "на глаз" наложим плавную кривую (рис. 4), которая, как видно, близка к нормальному распределению. К такому же выводу можно прийти, анализируя вид штрихового листа (см. табл. 2).

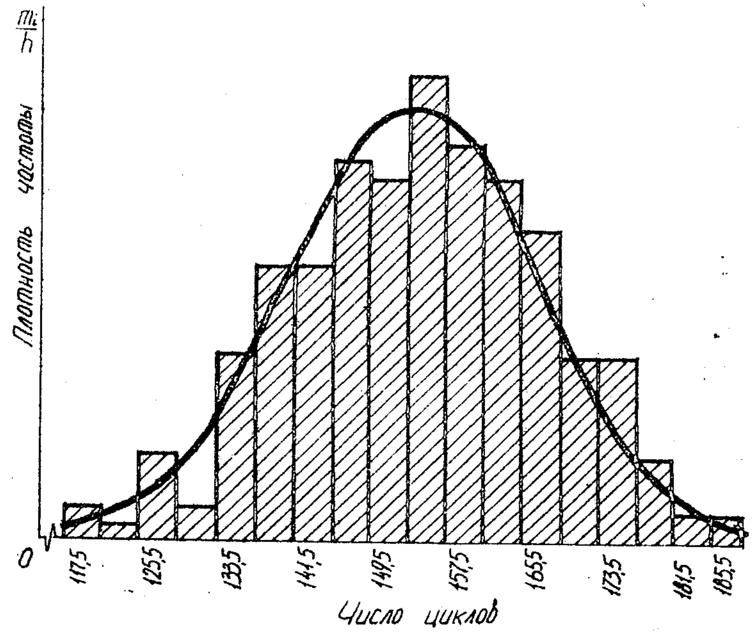


Рис. 4. Подбор статистической модели (распределения) по экспериментальным данным

Выше перечисленные особенности позволяют выдвинуть гипотезу: "Изучаемая случайная величина распределена по нормальному закону $N(a,\sigma)$" (см. гл. IV, § 4) с плотностью

$$f(x;\vartheta) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-a)^2}{2\sigma^2}}, \tag{31}$$

где $\vartheta = (a,\sigma)$, $a \in R$, $\sigma > 0$ – подлежащие определению параметры распределения.

Параметры $\vartheta = (a,\sigma), a \in R, \sigma > 0$ распределения (31) имеют простой вероятностный смысл: a является математическим ожиданием MX, а σ^2 – дисперсией $DX = MX^2 - (MX)^2$ нормально распределенной случайной величины X; их значения необходимо определить (оценить) по опытным данным (выборке).

В дальнейшем будем называть все подлежащие определению величины (вероятности событий, распределения случайных величин, их параметры и т.д.) вероятностными характеристиками, а найденные по результатам эксперимента, т.е. по выборке, их значения - оценками.

Последующий материал будет посвящен различным методам нахождения оценок и исследованию точности их приближения к оцениваемым характеристикам, а также рассмотрению методов проверки выдвинутых гипотез.

§ 2. Виды вероятностной сходимости

На основе вышеизложенных интуитивных соображений мы неизбежно приходим к выводу о том, что при увеличении объема выборки n частоты событий $\frac{m_i}{n}$ и выборочные средние $\bar{x}_в$ значений случайной величины должны бы приближаться к соответствующим вероятностям и математическим ожиданиям. Однако вследствие случайности результатов опыта о сходимости в обычном смысле (принятом в математическом анализе) здесь говорить не приходится.

Последовательность случайных величин $\{X_n\}$ называется <u>сходящейся в среднем квадратическом</u> (сокращенно: с.к. сходящейся) к случайной величине X, если

$$M(X_n - X)^2 \longrightarrow 0 \quad \text{при} \quad n \to \infty. \tag{32}$$

Последовательность случайных величин $\{X_n\}$ называется <u>сходящейся по вероятности</u> (п.в. сходящейся) к случайной величине X, если $\forall \varepsilon > 0$

$$P(|X_n - X| \geq \varepsilon) \to 0 \quad \text{при} \quad n \to \infty. \tag{33}$$

Последовательность случайных величин $\{X_n\}$ называется <u>сходящейся почти наверное</u> (п.н. сходящейся) к случайной величине X, если

$$P(X_n \to X) = 1 \quad \text{при} \quad n \to \infty. \tag{34}$$

Заметим, что вышеприведенные определения относятся и к сходимости последовательности случайных величин $\{X_n\}$ к неслучайной величине (например, к нулю), так как последнюю можно рассматривать как случайную величину с единственным возможным значением, имеющим вероятность, равную единице.

Из сопоставления этих определений интуитивно можно предположить, что всякая с.к. сходящаяся к X (или п.н. сходящаяся к X) последовательность случайных величин $\{X_n\}$ является п.в. сходящейся к тому же пределу X.

При доказательстве этого утверждения воспользуемся неравенством Чебышева [4] для случайной величины ξ:

$$P(|\xi| \geq \varepsilon) \leq \frac{1}{\varepsilon^p} M|\xi|^p, \quad \forall \varepsilon > 0, \ p > 0. \tag{35}$$

На практике число p подбирается так, чтобы существовали конечные начальные моменты порядка p, т.е. $M|\xi|^p < \infty$.

Полагая в (35) $p = 2$ и применяя неравенство (35) к случайной величине $\xi = X_n - X$, имеем

$$P(|X_n - X| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} M(X_n - X)^2.$$

Если $\{X_n\}$ - с.к. сходящаяся к X последовательность, т.е. если $M(X_n - X)^2 \to 0$ при $n \to \infty$, то $P(|X_n - X| \geq \varepsilon) \to 0$ при $n \to \infty$, т.е. $\{X_n\}$ - п.в. сходящаяся к X последовательность.

Докажем вторую часть этого утверждения. Из определения (33) следует, что

$\forall \varepsilon > 0$, $P(|X_m - X| \geq \varepsilon$ хотя бы при одном $m \geq n) \to 0$ при $n \to \infty$.

Кроме того,

$$P(|X_n - X| \geq \varepsilon) \leq P(|X_m - X| \geq \varepsilon \text{ хотя бы при одном } m \geq n).$$

Таким образом, из п.н. сходимости $\{X_n\}$ к X следует п.в. сходимость $\{X_n\}$ к тому же пределу X.

Заметим, что в самом общем случае обратное утверждение является неверным: из п.в. сходимости $\{X_n\}$ не всегда следует с.к. сходимость (или п.н. сходимость) той же последовательности $\{X_n\}$. Для доказательства этого приведем пример.

<u>Пример 7</u>. Пусть случайная величина X_n распределена по закону Коши:

$$f_n(x) = \frac{1}{\pi} \frac{n}{1 + n^2 x^2}, \quad n = 1, 2, \ldots \tag{35'}$$

Вычислим вероятность события $|X| \geq \varepsilon;\ \forall \varepsilon > 0$ при $n \to \infty$:

$$P(|X| \geq \varepsilon) = 1 - P(|X| < \varepsilon) = 1 - \int_{-\varepsilon}^{\varepsilon} f_n(x)\,dx = 1 - \frac{2}{\pi}\operatorname{arctg}(n\varepsilon) \to 0$$

Значит последовательность $\{X_n\}$ является п.в. сходящейся к нулю. С другой стороны,

$$MX_n^2 = \frac{n}{\pi}\int_{-\infty}^{+\infty} \frac{x^2}{1+n^2x^2}\,dx = \infty, \quad \forall n \in N,$$

т.е. случайная величина, распределенная по закону Коши, не имеет конечных начальных моментов, вследствие чего последовательность $\{X_n\}$ не является с.к. сходящейся.

Резюмируя вышеизложенное, отметим [7], что как множество с.к. сходящихся последовательностей, так и множество п.н. сходящихся последовательностей являются подмножествами последовательностей п.в. сходящихся, причем эти два подмножества имеют непустое пересечение (рис. 5).

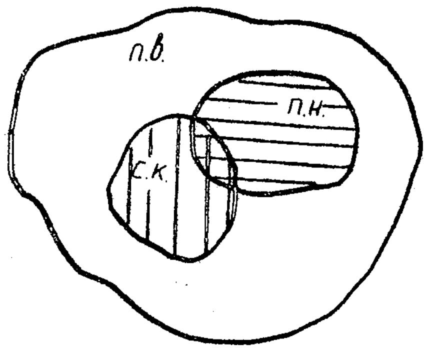


Рис. 5. Виды вероятностной сходимости

В заключение обратим внимание на следующие два обстоятельства:

I. Так как для случайной величины S справедливо

$$DS = MS^2 - (MS)^2 \Longleftrightarrow MS^2 = DS + (MS)^2,$$

то для $S = X_n - X$ имеем

$$M(X_n - X)^2 = D(X_n - X)^2 + (MX_n - MX)^2, \tag{36}$$

т.е. для с.к. сходимости к X последовательности $\{X_n\}$ необходима и достаточна сходимость последовательности математических ожиданий $\{MX_n\}$ к MX и сходимость последовательности дисперсий разностей $X_n - X$ к нулю.

Из формулы (36) также видно, что для с.к. сходимости $\{X_n\}$ к неслучайной величине a необходима и достаточна сходимость $\{MX_n\}$ к a и сходимость последовательности дисперсий $\{DX_n\}$ к нулю.

2. Сходимость по вероятности, сходимость в среднем квадратическом и сходимость почти наверное представляют собой особые, вероятностные виды сходимости; в частном случае последовательности неслучайных величин все эти три вида сходимости совпадают с обычной сходимостью классического математического анализа.

§ 3. Оценивание. Состоятельные, несмещенные и эффективные оценки

Любая функция результатов эксперимента $x_1, x_2, \ldots, x_n$, которая не зависит от неизвестных параметров, называется <u>статистикой</u> $S = S(x_1, x_2, \ldots, x_n)$.

Так как теория вероятностей имеет дело только со случайным экспериментом, то любая статистика представляет собой случайную величину.

<u>Оценкой</u> $\hat{\vartheta}$ вероятностной характеристики ϑ называется статистика $\hat{\vartheta} = S(x_1, x_2, \ldots, x_n)$, реализация которой, полученная в результате случайного эксперимента, принимается за неизвестное истинное значение параметра ϑ.

Различают оценки точечные и интервальные (интервальные оценки будут рассмотрены ниже).

<u>Точечной</u> называют статистическую оценку параметра, характеризуемого одним определенным значением (числом), которое находят из подходящих для этой цели статистик $S(x_1, x_2, \ldots, x_n)$ и

которое затем стремятся максимально приблизить к значению соответствующего параметра.

Ясно, что не всякая статистика может служить удовлетворительной оценкой данной вероятностной характеристики ϑ ; для того, чтобы оценки давали "хорошие" приближения оцениваемых вероятностных характеристик, они должны удовлетворять определенным требованиям.

Оценка $\hat{\vartheta}$ вероятностной характеристики ϑ называется <u>состоятельной</u>, если она является п.в. сходящейся к ϑ ;

$$P(|\hat{\vartheta}-\vartheta|\geq\varepsilon)\to 0, \quad \forall\varepsilon>0, \quad n\to\infty.$$

Из закона больших чисел /см. (35), (36)/ вытекает: чтобы оценка $\hat{\vartheta}$ была состоятельной достаточно выполнения условий

$$M\hat{\vartheta}\to\vartheta, \quad D\hat{\vartheta}\to 0 \quad \text{при} \quad n\to\infty.$$

Оценка $\hat{\vartheta}$ вероятностной характеристики ϑ называется <u>несмещенной</u>, если

$$M\hat{\vartheta}=\vartheta, \tag{37}$$

и смещенной, если

$$M\hat{\vartheta}\neq\vartheta$$

при любом объеме выборки n ; разность

$$M\hat{\vartheta}-\vartheta \tag{38}$$

называют смещением оценки.

Рассмотрим статистики (8), (I4), соответствующие оценки которых получены в примере 3, и вычислим их математическое ожидание:

$$M\bar{x}_в=M\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)=\frac{1}{n}\sum_{i=1}^{n}MX_i=\frac{1}{n}\,n\bar{x}_г=\bar{x}_г,$$

$$Mx_в=\bar{x}_г=MX, \tag{39}$$

где $\bar{x}_г$ - среднее значение генеральной совокупности, $\bar{x}_г=MX$;

$$M\sigma_в^2=M\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x}_в)^2\right)=\frac{1}{n}\sum_{i=1}^{n}MX_i^2-M(\bar{x}_в)^2=\sigma_г^2-\frac{1}{n}\sigma_г^2,$$

$$M\sigma_в^2=\frac{n-1}{n}\sigma_г^2\neq\sigma_г^2. \tag{40}$$

Следовательно, среднее значение $\bar{x}_в$ выборки является несмещенной оценкой для математического ожидания $MX=\bar{x}_г$ генеральной совокупности. Напротив, статистика (I4) является смещенной оценкой для дисперсии $DX=\sigma_г^2$ генеральной совокупности.

По этой причине на практике предпочтительнее пользоваться (особенно в случае малых объемов выборки n) статистикой

$$S^2=\frac{n}{n-1}\sigma_в^2=\frac{1}{n-1}\sum_{i=1}^{g}m_i(x_i-\bar{x}_в)^2, \tag{41}$$

называемой <u>"исправленной" дисперсией</u>.

Разумеется, при больших объемах выборки, когда $n-1\approx n$, различия между этими двумя оценками несущественны.

Обычно исправленная выборочная дисперсия вычисляется по формулам

$$S^2=\frac{1}{n-1}\sum_{i=1}^{g}m_i x_i^2-\frac{1}{n(n-1)}\left(\sum_{i=1}^{g}m_i x_i\right)^2, \tag{42}$$

или в условных вариантах (20)

$$S_u^2=\frac{1}{n-1}\sum_{i=1}^{g}m_i u_i^2-\frac{1}{n(n-1)}\left(\sum_{i=1}^{g}m_i u_i\right)^2, \tag{43}$$

$$S^2=k^2S_u^2, \quad x_i=ku_i+C.$$

Для статистики (4I) при $n\to\infty$ справедливы соотношения:

$$DS^2=\left(\frac{n}{n-1}\right)^2 D\sigma_в^2\to 0, \tag{44}$$

$$MS^2=\frac{n}{n-1}M\sigma_в^2=\sigma^2=DX, \tag{45}$$

означающие, что S^2 является состоятельной несмещенной оценкой для дисперсии $\sigma^2=DX$ генеральной совокупности.

Пример 8. Вычислить несмещенную оценку генеральной дисперсии для числа циклов растяжения.

Решение. Воспользовавшись результатами табл. 3, по формулам (43) найдем

$$S_u^2 = \frac{216 \cdot 2453 - (-27)^2}{216 \cdot 215} = 11,3956;$$

$$S^2 = (4,0)^2 \cdot 11,3956 = 182,2976;$$

$$S = 13,5018.$$

Иногда в инженерных расчетах для получения более точной оценки генеральной дисперсии вводят поправку Шеппарда на группировку [6, 7]. Поправка состоит в уменьшении дисперсии S_u^2 (для условных вариант) на величину $h^2/12$, где h – интервал группировки.

Следовательно, в случае оценки, рассмотренной в предыдущем примере, дисперсия $S_u^2 = 11,3976$ уменьшится до $S_u^2 - h^2/12 = 10,8936$. Тогда с учетом поправки Шеппарда имеем

$$S = \sqrt{(4,0)^2 \cdot 10,8936} = 13,2.$$

Качество оценки принято характеризовать математическим ожиданием квадрата смещения оценки /см. формулу (38)/

$$\delta = M(\hat{\vartheta} - \vartheta)^2 \geq 0, \qquad (46)$$

называемого средним квадратом ошибки. Очевидно, в случае несмещенной оценки (37) средний квадрат ошибки равен дисперсии оценки $\hat{\vartheta}$.

Множество значений $\delta \geq 0$, соответствующее всем оценкам $\hat{\vartheta}$ вероятностной характеристики ϑ при заданном n, как всякое множество неотрицательных чисел, имеет точную нижнюю грань

$$\delta_0 = \inf M(\hat{\vartheta} - \vartheta)^2. \qquad (47)$$

Естественно, нужно стремиться к тому, чтобы искомая оценка $\hat{\vartheta}$ имела δ, равное δ_0 или близкое к нему значение. Такие оценки называются эффективными. Это требование к оценкам имеет особенно простой вероятностный смысл в случае несмещенных оценок.

Пусть даны две несмещенные оценки $\hat{\vartheta}_1 = S_1(x_1, x_2, \ldots, x_n)$

$\hat{\vartheta}_2 = S_2(x_1, x_2, \ldots, x_n)$ одной и той же вероятностной характеристики ϑ:

$$M\hat{\vartheta}_1 = M\hat{\vartheta}_2 = \vartheta. \qquad (48)$$

Если степень "концентрации" значений статистик S_1 и S_2 вокруг их математического ожидания (48) различна, то для использования в качестве оценки для ϑ более выгодно применять ту статистику, для которой степень указанной "концентрации" выше, т.е. применять ту статистику, у которой меньшая дисперсия. Таким образом, если для статистики $S = S_0$ выборочная дисперсия DS_0 достигает минимального возможного значения по сравнению с дисперсиями любых других статистик S, удовлетворяющих условию (37), то статистика S_0 является эффективной оценкой вероятностной характеристики ϑ.

В теории вероятностей доказывается [9], что любая эффективная оценка является несмещенной.

§ 4. Основные методы нахождения точечных оценок. Достаточные статистики

Если распределение генеральной совокупности описывается функцией $F(x, \vartheta)$ или $f(x, \vartheta)$ заданного вида с подлежащими определению параметрами $\vartheta = \{\alpha_0, \alpha_1, \ldots, \alpha_m\}$, то каждая числовая характеристика распределения MX, DX, ν_k и т.д. является функцией параметров ϑ. В частности, $\nu_k = \nu_k(\vartheta)$, $k = 1, 2, \ldots$, если эти моменты существуют. Методом моментов определяют оценки $\hat{\vartheta}$ соответствующих параметров ϑ посредством уравнений

$$\nu_k(\vartheta) = \nu_k^*, \quad k = 1, 2, \ldots, \qquad (49)$$

полученных путем приравнивания первых k выборочных моментов ν_k^* соответствующим моментам ν_k генеральной совокупности. Получаемые таким образом точечные оценки $\hat{\vartheta}$ будут функциями от выборочных моментов.

Пример 9. Методом моментов найти оценки параметров $\vartheta = (\alpha, \sigma)$, $\alpha \in R$, $\sigma > 0$ нормально распределенной случайной величины (см. гл. IУ, § 4):

$$f(x;\alpha,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\alpha)^2}{2\sigma^2}} \qquad (50)$$

Решение. Вычисляя MX и DX, находим

$$MX=\alpha, \quad DX=\sigma^2, \qquad (51)$$

и система уравнений (49) принимает вид

$$\alpha=\bar{x}_в, \quad \sigma^2=\sigma^2_в. \qquad (52)$$

Пример 10. Методом моментов найти оценки параметров $\vartheta=(\lambda,\eta)$, $\lambda>0$, $\eta>0$ случайной величины X, имеющей гамма-распределение (см. гл. IУ, § 4):

$$f(x;\vartheta)=\begin{cases}\frac{\lambda^\eta}{\Gamma(\eta)}x^{\eta-1}e^{-\lambda x}, & x\geq 0;\\ 0, & x<0,\end{cases} \qquad (53)$$

где

$$\Gamma(\eta)=\int_0^\infty t^{\eta-1}e^{-t}dt - \qquad (54)$$

полная гамма-функция, которая табулирована (табл. IШ).

Решение. Вычисляя по формулам (26) два первых начальных момента $\nu_1(\lambda,\eta)$ и $\nu_2(\lambda,\eta)$, имеем

$$MX=\eta/\lambda, \quad DX=\eta/\lambda^2. \qquad (55)$$

Система уравнений (49) принимает вид

$$\begin{cases}\eta/\lambda=\bar{x}_в,\\ \eta/\lambda^2=\sigma^2_в\end{cases}$$

и приводит к оценкам

$$\hat{\lambda}=\bar{x}_в/\sigma^2_в, \quad \hat{\eta}=\bar{x}^2_в/\sigma^2_в. \qquad (56)$$

Точечные оценки для ϑ можно получить путем минимизации среднего квадрата ошибки /см. формулу (46)/. При этом фактически производится поиск несмещенных оценок с наименьшей дисперсией (метод наименьших квадратов) [9].

В случае нормально распределенной случайной величины (50) этот метод приводит к оценкам (52).

Наиболее важным, с теоретической точки зрения, общим методом нахождения оценок является метод максимального правдоподобия Фишера. Этот метод состоит в том, что за оценку неизвестных параметров $\vartheta=(\alpha_0,\alpha_1,\dots,\alpha_m)$ принимаются такие их значения, при которых плотность распределений результатов опытов $X_1, X_2,\dots,X_n$ достигает наибольшего значения при заданных реализациях $x_1, x_2,\dots,x_n$:

$$f(x_1,x_2,\dots,x_n/\vartheta)=\max. \qquad (57)$$

Функция (57) называется функцией правдоподобия, а значения $\hat{\vartheta}=(\hat{\alpha}_0,\hat{\alpha}_1,\dots,\hat{\alpha}_m)$, при которых она достигает наибольшего значения, – оценкой максимального правдоподобия параметров ϑ (МП-оценкой). На практике такие оценки находят, максимизируя $\ln f(x_1,x_2,\dots,x_n)$ и пользуясь тем, что $y=\ln x$ – строго возрастающая функция.

Так как результаты опытов, как правило, независимы, то

$$\ln f(x_1,x_2,\dots,x_n/\vartheta)=\sum_{i=1}^n \ln f(x_i/\vartheta), \qquad (58)$$

где $f(x/\vartheta)$ – плотность наблюдаемой случайной величины X.

Если $f(x/\vartheta)$ $\forall x\in R$ имеет непрерывную производную по ϑ (точнее непрерывный градиент по $\vartheta\equiv\{\alpha_0,\alpha_1,\dots,\alpha_m\}$), то для оценок максимального правдоподобия получаем систему уравнений

$$\sum_{i=1}^n \frac{\partial}{\partial\vartheta}\ln f(x_i/\vartheta)=0. \qquad (59)$$

Определение. Статистика $S=S(x_1,x_2,\dots,x_n)$ для характеристики ϑ называется достаточной, если при любом другом выборе другой статистики S', для которой нет функциональной зависимости $\varphi(S,S')=0$, условное распределение статистики S' при заданном значении S статистики $S(x_1,x_2,\dots,x_n)$ не зависит от ϑ. Таким образом, по определению имеем

$$f(S,S'/\vartheta)=f_1(S/\vartheta)\cdot f_2(S'/S),$$

или

$$f_2(S'/S)=\frac{f(S,S'/\vartheta)}{f_1(S/\vartheta)}\neq\varphi(\vartheta), \qquad (60)$$

где $\mathcal{S}'$ – любая другая статистика $\mathcal{S}(x_1, x_2, \dots, x_n)$, $f(\mathcal{S}, \mathcal{S}'/\vartheta)$ – совместная плотность случайных величин $\mathcal{S}$, $\mathcal{S}'$, зависящая от ϑ, $f_1(\mathcal{S}/\vartheta)$ – плотность случайной величины $\mathcal{S}$.

Иными словами, статистика $\mathcal{S}$ достаточна, если значения любых других статистик $\mathcal{S}'$ не дают никакой дополнительной информации о характеристиках ϑ сверх той, которую дает статистика $\mathcal{S}$.

В математической статистике доказывается [6, 8], что:

1. Если существует достаточная статистика для ϑ, то все решения системы (59) являются функциями этой достаточной статистики. По этой причине все решения системы (59), не зависящие от результатов опытов $x_1, x_2, \dots, x_n$, следует отбросить (как не дающие никакой статистической информации).

2. Если существует эффективная оценка $\hat{\vartheta}$ для ϑ, то она является единственным решением системы (59).

3. Оценки максимального правдоподобия при достаточно общих условиях состоятельны, асимптотически (т.е. при больших объемах n) нормальны и асимптотически эффективны.

Доказательство этих утверждений выходит, к сожалению, за рамки настоящего пособия.

Пример 11. Найти МП-оценки параметров $\vartheta \equiv (a, \sigma)$, $a \in R$, $\sigma > 0$, нормально распределенной случайной величины X (см. пример 6).

Решение. В выборке примера 1 содержится $n = 216$ результатов эксперимента $X_1, X_2, \dots, X_n$, которые можно рассматривать как независимые, одинаково распределенные случайные величины с функцией максимального правдоподобия

$$L = f(x_1, \dots, x_n/\vartheta) = \prod_{i=1}^{n} f(x_i/\vartheta) = (2\pi)^{-n/2} \sigma^{-n} \exp\left\{-\sum_{i=1}^{n} \frac{(x_i - a)^2}{2\sigma^2}\right\}, \tag{61}$$

$$\vartheta = (a, \sigma), \quad a \in R, \quad \sigma > 0, \quad n = 216.$$

Приравнивая нулю частные производные по a и σ, имеем систему (59)

$$\sum_{i=1}^{n} (x_i - a) = 0,$$

$$-n\sigma^2 + \sum_{i=1}^{n} (x_i - a)^2 = 0,$$

решение которой приводит к МП-оценкам

$$a = \bar{x}_в, \quad \sigma^2 = D^* X = \frac{1}{n} \sum_{i=1}^{n} (x_i - x_в)^2. \tag{61'}$$

Заметим, что оценки (61') состоятельны, нормально распределены и асимптотически эффективны ($n = 216$). Эти оценки в случае нормального закона совпадают с оценками, ранее полученными методом моментов (см. пример 9) и МНК-методом [9].

Пример 12. Найти МП-оценки параметров: а) гамма-распределения /см. формулу (53)/; б) распределения Вейбулла

$$f(x;\vartheta) = \begin{cases} \dfrac{b}{c}\left(\dfrac{x}{c}\right)^{b-1} e^{-(x/c)^b}, & x \geq 0; \\ 0, & x < 0, \end{cases} \tag{62}$$

где $\vartheta = (b, c)$, $b > 0$, $c > 0$.

Решение: а. Функция максимального правдоподобия (57) имеет вид

$$f(x_1, \dots, x_n/\vartheta) = \lambda^{n\eta} \Gamma^{-n}(\eta) \prod_{i=1}^{n} x_i^{\eta-1} \cdot \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}, \tag{ж}$$

где $\vartheta = (\lambda, \eta)$, $\lambda > 0$, $\eta > 0$.

Логарифмируя (ж) и приравнивая нулю частные производные по λ и η, для оценок параметров λ и η получаем систему нелинейных уравнений /см. систему (59)/

$$\begin{cases} \eta/\lambda = \bar{x}_в, \\ \Gamma'(\eta)/\Gamma(\eta) = \ln\lambda - \overline{\ln x}. \end{cases} \tag{63}$$

Здесь и ниже в примере черта означает усреднение по выборке объема n, например $\overline{\ln x} = \frac{1}{n}\sum_{i=1}^{n} \ln x_i$; штрих – производная полной гамма-функции (54) по параметру η.

б. В случае распределения Вейбулла (62) –

$$\ln f(x_1, \ldots, x_n / \vartheta) =$$

$$= n \ln b + (b-1)\sum_{i=1}^{n} \ln x_i - n b \ln c - \sum_{i=1}^{n} (x_i/c)^b . \qquad (**)$$

Аналогично, для параметров $a > 0$, $b > 0$ после несложных преобразований получим систему

$$\begin{cases} c^b = \overline{x^b}, \\ b = \dfrac{\overline{x^b}}{\overline{x^b \ln x} - \overline{x^b}\cdot\overline{\ln x}}, \end{cases} \qquad (62')$$

для которой не выполняется условие (60), т.е. рассматриваемая статистика не является достаточной.

Замечания. I. Если в выражении (53) $\eta = I$, $\Gamma(1) = I$, то получаем частный случай экспоненциального распределения (см. гл. IУ, § 4):

$$f(x;\vartheta) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0; \\ 0, & x < 0, \end{cases} \qquad (64)$$

где $\vartheta = \lambda$.

Система (63) приводит к оценке максимального правдоподобия параметра λ :

$$\lambda = 1/x_в . \qquad (64')$$

2. В случае гамма-распределения (53) метод приводит к необходимости решения системы нелинейных уравнений, что приводит к трудоемким вычислениям с использованием ЭВМ. Этого можно избежать, применяя метод максимального правдоподобия по методике, рассмотренной в работе [II].

3. В случае распределения Вейбулла (62) метод моментов (49) приводит к нелинейным уравнениям

$$\begin{cases} c\,\Gamma(1+1/b) = x_в , \\ c^2\left\{\Gamma(1+2/b) - [\Gamma(1+1/b)]^2\right\} = D^* X \end{cases} \qquad (62'')$$

для оценок параметров b и c.

§ 5. Оценивание с помощью доверительных интервалов

Пусть генеральная совокупность имеет функцию распределения вероятностей $F(x, \vartheta)$, зависящую от одного параметра $\vartheta = \alpha$. Допустим также, что $S_1 = S_1(x_1, x_2, \ldots, x_n)$ и $S_2 = S_2(x_1, x_2, \ldots, x_n)$ такие две статистики, для которых вероятность одновременного осуществления двух неравенств

$$S_1(x_1, x_2, \ldots; x_n) < \alpha_1 , \qquad S_2(x_1, x_2, \ldots, x_n) > \alpha_1$$

равна $1 - \alpha$. Иными словами, пусть соблюдается условие

$$P\{S_1 < \alpha_1 < S_2\} = 1 - \alpha \qquad (65)$$

независимо от того, чему в действительности равно значение параметра α_1. Интервал (S_1, S_2) является случайным, его длина и положение на числовой оси меняются от выборки к выборке.

Случайный интервал, полностью определенный результатами опытов и независящий от неизвестных параметров, который с наперед заданной вероятностью $(1-\alpha)$ накрывает неизвестную скалярную характеристику $\vartheta = \alpha_1$, называется доверительным интервалом; его границы S_1, S_2 – доверительными границами. Значение вероятности $(1-\alpha)$, называемой доверительной (коэффициентом доверия или надежностью), из теории не следует и задается исследователем; на практике за $(1-\alpha)$ берут значения 0,95 или 0,99, которые соответственно согласуются с правилом 2-х и 3-х сигм. Число α называют уровнем значимости.

Если генеральная совокупность кроме параметра α_1 зависит и от других параметров $\alpha_2, \alpha_3, \ldots$, то доверительный интервал (точнее доверительная область) определяется требованием соблюдения условия (65) при любых значениях всех параметров $\vartheta = (\alpha_1, \alpha_2, \ldots)$.

Доверительные интервалы (S_1, S_2), являющиеся оценками вероятностных характеристик ϑ, имеют простой вероятностный смысл. Если после извлечения выборки объема n и нахождения (по заданному $(1-\alpha)$) отвечающего ей доверительного интервала (S_1, S_2) считать, что $\vartheta \in (S_1, S_2)$, то доля тех случаев, когда это заключение окажется верным, будет по закону больших чисел близка к $(1-\alpha)$. Доля же случаев, когда вероятностная характеристика $\vartheta \in (\mathcal{S}_1, \mathcal{S}_2)$ окажется ложной, будет примерно равна α.

Очевидно, что доверительный интервал для $\vartheta \equiv \alpha_1$ или доверительная область для $\vartheta \equiv (\alpha_1, \alpha_2, \dots)$ определяется доверительной вероятностью $(1-\alpha)$ не однозначно; существует бесконечное множество доверительных интервалов, соответствующих одному и тому же значению $(1-\alpha)$. Обычно при заданном $(1-\alpha)$ стремятся определить такой доверительный интервал, который имеет минимальный размер. Этой цели служат следующие методы.

Первый метод, удобный в случае положительного параметра $\vartheta \equiv \alpha_1$, основан на нахождении распределения отношения оценки $\hat{\vartheta}$ параметра ϑ к самому параметру ϑ. Если оценка такова, что это распределение не зависит от неизвестных вероятностных характеристик, то, зная это распределение, можно найти вероятность попадания отношения $\hat{\vartheta}/\vartheta$ в любой интервал и, наоборот, по заданному $(1-\alpha)$ найти интервал, вероятность попадания отношения $\hat{\vartheta}/\vartheta$ в котором равна $(1-\alpha)$.

Вычислим вероятность неравенства (относительной ошибки)

$$|\hat{\vartheta} - \vartheta|/\hat{\vartheta} < \Delta_\alpha, \quad \forall \Delta_\alpha > 0, \quad \vartheta \geq 0$$

и приравняем ее заданному $(1-\alpha)$:

$$P\{|\hat{\vartheta} - \vartheta|/\hat{\vartheta} < \Delta_\alpha\} = P\{(1-\Delta_\alpha)\hat{\vartheta} < \vartheta < (1+\Delta_\alpha)\hat{\vartheta}\} = 1-\alpha. \quad (66)$$

Обычно стремятся получить доверительный интервал, симметричный относительно оценки $\hat{\vartheta}$ /что не всегда возможно, так как нижняя граница в (66) может оказаться отрицательной/. Поэтому в (66) для положительного параметра ϑ следует определить доверительные границы из условия

$$\max\{0; (1-\Delta_\alpha)\hat{\vartheta}\} < \vartheta < (1+\Delta_\alpha)\hat{\vartheta} \Longleftrightarrow \max\{0; 1-\alpha\} < \vartheta/\hat{\vartheta} < 1+\Delta_\alpha,$$

или

$$\frac{1}{1+\Delta_\alpha} < \frac{\hat{\vartheta}}{\vartheta} < \frac{1}{\max\{0; 1-\Delta_\alpha\}}. \quad (67)$$

При $\Delta_\alpha \in (0; 1]$ интервал (67) симметричен относительно $\hat{\vartheta}$; при $\Delta_\alpha > 1$ симметрия нарушается. Таким образом, этот метод приводит к уравнению

$$P\left\{\frac{1}{1+\Delta_\alpha} < \frac{\hat{\vartheta}}{\vartheta} < \frac{1}{\max\{0; 1-\Delta_\alpha\}}\right\} = 1-\alpha \quad (66')$$

для нахождения доверительного интервала для ϑ при заданной вероятности $(1-\alpha)$. Подчеркнем, что закон распределения $\hat{\vartheta}/\vartheta$ в (66') считается известным и не зависящим от неизвестных вероятностных характеристик.

Второй метод. Для каждого значения неизвестного параметра $\vartheta \equiv \alpha_1$ выбирают такую область $D_\alpha(\vartheta)$, содержащую ϑ, в которую оценка $\hat{\vartheta}$ попадает с заданной доверительной вероятностью $(1-\alpha)$:

$$P\{\hat{\vartheta} \in D_\alpha(\vartheta)\} = 1-\alpha. \quad (68)$$

Но множество значений ϑ в общем случае зависит от $\hat{\vartheta}$ и $(1-\alpha)$, т.е.

$$\{\vartheta : \hat{\vartheta} \in D_\alpha(\vartheta)\} \equiv \Delta_\alpha(\hat{\vartheta}).$$

Иными словами, если для данной выборки $x_1, x_2, \dots, x_n$ $\hat{\vartheta} \in D_\alpha(\vartheta)$, то $\vartheta \in \Delta_\alpha(\hat{\vartheta})$ и наоборот. Таким образом, имеем уравнение

$$P(\hat{\vartheta} \in D_\alpha(\vartheta)) = P\{\vartheta \in \Delta_\alpha(\hat{\vartheta})\} = 1-\alpha \quad (69)$$

для определения доверительного интервала $\Delta_\alpha(\hat{\vartheta})$.

Здесь $(1-\alpha)$ – заданная доверительная вероятность, оценка $\hat{\vartheta}$ есть фиксированная выборкой $x_1, x_2, \dots, x_n$ величина; закон распределения в (69) считается известным и не зависящим от неизвестных вероятностных характеристик.

Третий метод [3]. Предположим, что существует статистика $S(\vartheta, \hat{\vartheta})$ такая, что:

1) закон распределения статистики S известен и не зависит от оцениваемого параметра ϑ;

2) функция $S(\vartheta, \hat{\vartheta})$ непрерывна и строго монотонна по ϑ.

Пусть далее $(1-\alpha)$ – заданная доверительная вероятность, а $s_{\alpha/2}$ и $s_{1-\alpha/2}$ – квантили распределения статистики S уровней $\alpha/2$ и $1-\alpha/2$ соответственно. Тогда с вероятностью $(1-\alpha)$ выполняется неравенство

$$s_{\alpha/2} < S(\hat{\vartheta}, \vartheta) < s_{1-\alpha/2}. \qquad (70)$$

Решая неравенство (70) относительно ϑ, найдем границы ϑ_1 и ϑ_2 доверительного интервала ϑ.

§ 6. Общие сведения о распределении χ^2 и распределении Стьюдента

Распределения статистик, вычисляемых по выборке из нормально распределенной генеральной совокупности, представляют особый практический интерес и связаны с распределениями $\chi^2(k)$ ("хи-квадрат" распределение) и Стьюдента (t-распределение). Такие статистики будут рассмотрены в следующем параграфе; квантили этих распределений приведены в приложении (табл. ШУ и ПУ). Приведем общие сведения и некоторые свойства этих распределений.

Распределение $\chi^2(k)$ (распределение Пирсона). Пусть имеется k независимых нормально распределенных по закону $N(0, 1)$ случайных величин U_i, $i = 1, 2, \ldots, k$. В теории вероятностей доказывается [4, 6, 9], что случайная величина

$$\chi^2(k) = U_1^2 + \ldots + U_k^2 \qquad (71)$$

имеет распределение χ^2 с плотностью

$$f_{\chi^2}(x) \equiv f(x;k) = \begin{cases} \dfrac{1}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k-2}{2}} e^{-x/2}, & x \ge 0; \\ 0, & x < 0, \end{cases} \qquad (71')$$

где $\Gamma(t)$ – полная гамма-функция /см. формулу (54)/; параметр k принято называть числом степеней свободы распределения. Графическая иллюстрация, основные числовые характеристики распределения приводятся в § 4 гл. IУ.

Практическая значимость этого распределения в статистических вычислениях объясняется, в частности, нижеследующей теоремой [1, 8].

Пусть $x_1, x_2, \ldots, x_n$ – выборка из нормально распределенной генеральной совокупности $N(a, \sigma)$, а a, s – соответственно выборочное среднее /см. формулу (8)/ и выборочная "исправленная" дисперсия /см. формулу (41)/. Тогда статистики $\bar{x}_в$ и S^2 – независимые случайные величины, причем статистика $\dfrac{(n-1)S^2}{\sigma^2}$

имеет распределение $\chi^2(n-1)$.

Отсюда, в частности, следует, что если в выражении (71) слагаемые разбить на две произвольные подгруппы с числом слагаемых k_1 и k_2 соответственно ($k = k_1 + k_2$), то случайные величины $\chi^2(k_1)$ $\chi^2(k_2)$ – независимые, имеют распределение χ^2 (71') с k_1 и k_2 степенями свободы соответственно. Причем сумма этих случайных величин имеет распределение (71') с $k_1 + k_2$ степенями свободы:

$$\chi^2(k_1) + \chi^2(k_2) = \chi^2(k_1 + k_2).$$

Заметим, что распределение (71') при больших k ($k > 30$) с достаточной для практических расчетов точностью аппроксимируется нормальным распределением. Этот факт позволяет [3] выразить квантили распределения $\chi^2(k)$ через квантили u_p нормального распределения $N(0, 1)$:

$$\chi_p^2(k) \approx \frac{1}{2}\left(u_p + \sqrt{2k-1}\right)^2, \qquad (72)$$

где $k \ge 30$, $p \ge 0{,}5$, относительная погрешность формулы лежит в пределах 1 %;

$$\chi_p^2(k) \approx k\left(1 - \frac{2}{9k} + 4u_p\sqrt{\frac{2}{9k}}\right)^3 \qquad (72')$$

(применяется для вычисления квантилей малого уровня).

Упражнение 1. Вычислить квантили

$$\chi^2_{0,01}(12),\ \chi^2_{0,99}(40),\ \chi^2_{0,01}(40),\ \chi^2_{0,025}(215),\ \chi^2_{0,975}(215).$$

Решение. По табл. ПУ находим $\chi^2_{0,01}(12) = 3{,}57$; по формуле (72) имеем $\chi^2_{0,99}(40) \approx \frac{1}{2}(2{,}326 + \sqrt{2\cdot 40 - 1})^2 \approx 62{,}878$, где $u_{0,99} = 2{,}326$ (см. табл. Ш).

Затем по формуле (72') получаем

$$\chi^2_{0,01}(40) \approx 40\left(1 - \frac{2}{9\cdot 40} - 2{,}326\sqrt{\frac{2}{9\cdot 40}}\right)^3 = 22{,}14,$$

так как $u_{0,01} = -u_{0,99} = -2{,}326$ (см. табл. Ш).

Аналогично получаем значения для $\chi^2_{0,975}(215) \approx 257,0165$:

$$\chi^2_{0,025}(215) \approx 215\left(1-\frac{2}{9\cdot 215}-1,96\sqrt{\frac{2}{9\cdot 215}}\right)^3 = 176,902.$$

Заметим, что $u_{0,025} = -u_{0,975} = -1,96$ найдено по табл. ПII.

<u>Распределение Стьюдента с k степенями свободы</u> (псевдоним английского математика В.Госсета). Пусть случайная величина U имеет нормальное распределение $N(0,1)$. Тогда случайная величина $T(k)$, равная отношению двух независимых случайных величин U и $\sqrt{\chi^2(k)/k}$, имеет [1, 4, 8] распределение

$$f_T(x) \equiv f(x,k) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma(k/2)\sqrt{k\pi}}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \quad x\in R. \tag{73}$$

Квантили этого распределения приведены в табл. ПV; графическая иллюстрация (73), основные числовые характеристики даны в § 4 гл. IV. Так как функция (73) является четной, то для квантилей этого распределения справедливо

$$t_p(k) = -t_{1-p}(k).$$

Заметим, что при неограниченном возрастании объема выборки n (в практических расчетах $k\geqslant 30$) распределение Стьюдента (73) стремится к нормальному $N(0,1)$, так как справедливо

$$\lim_{k\to\infty}\frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma(k/2)} = \frac{1}{\sqrt{2\pi}}, \quad \lim_{k\to\infty}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}} = e^{-x^2/2}$$

Графическая иллюстрация этого факта приведена в § 4 гл. IV; при больших k ($k\geqslant 30$) для квантилей $t_p(k)$ распределения справедливо

$$t_p(k) \approx u_p,$$

или более точно [4]

$$t_p(k) \approx u_p\left[\left(1-\frac{1}{4k}\right)^2-\frac{u_p^2}{2k}\right]^{-1/2}. \tag{73'}$$

<u>Упражнение 2</u>. Вычислить квантили $t_{0,9}(8)$ и $t_{0,1}(8)$, $t_{0,9}(40)$.

<u>Решение</u>. По табл. ПV находим $t_{0,9}(8) = 1,397$; $t_{0,1}(8) = -t_{0,9}(8) = -1,397$; $t_{0,9}(40) = 1,303$.

Для сравнения последний квантиль вычислим приближенно по формуле (73'): $u_{0,9} = 1,282$ (табл. ПII), $t_{0,9}(40) = 1,297$; относительная ошибка – менее 0,5 %.

§ 7. <u>Интервальные оценки параметров нормального распределения</u>

На практике особый интерес представляют доверительные интервалы для параметров $\vartheta = (a,\sigma)$, $a\in R$, $\sigma>0$ нормально распределенной случайной величины /см. формулу (50)/. Здесь в зависимости от цели возможна постановка следующих задач.

<u>Оценка математического ожидания при известном σ</u>. Пусть случайная величина X распределена нормально /см. формулу (50)/, причем среднее квадратическое отклонение $\sigma = \sqrt{DX}$ известно. Оценим параметр a по выборочной средней $\bar{x}_в = S(x_1,\ldots,x_n)$, являющейся случайной величиной $\bar{X}_в$ ($\bar{x}_в$ меняется от выборки к выборке), а реализации случайной величины $x_1, x_2, x_3, \ldots, x_n$ будем рассматривать соответственно как одинаково распределенные независимые случайные величины $X_1, X_2, \ldots, X_n$ (также меняющиеся от выборки к выборке). Значит $MX_i = a$, $DX_i = \sigma^2$, $i = 1,2,\ldots,n$. Кроме этого, имеем

$$\bar{X}_в = \frac{1}{n}\sum_{i=1}^{n}X_i, \quad M\bar{X}_в = M\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \frac{1}{n}\cdot na = a;$$

$$D\bar{X}_в = D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}DX_i = \frac{\sigma^2}{n}.$$

Таким образом, получаем

$$M\bar{X}_в = a, \quad D\bar{X}_в = \sigma^2/n, \quad \sigma(\bar{X}_в) = \sigma/\sqrt{n}. \tag{74}$$

Согласно (74) рассмотрим статистику

$$S \equiv U = \frac{\bar{X}_в - a}{\sigma/\sqrt{n}}, \tag{75}$$

имеющую нормальное распределение $N(0,1)$ независимо от значения параметра $a \in R$. В этом случае U как функция параметра a непрерывна и строго монотонна и, следовательно, по известной формуле теории вероятностей [I]

$$P(u_{\alpha/2} < U < u_{1-\alpha/2}) = 2F(u_{\alpha/2}) - 1 = 1 - \alpha, \qquad (75')$$

где $u_{\alpha/2}$, $u_{1-\alpha/2}$ – квантили нормального распределения $N(0,1)$ уровня $\alpha/2$, $1-\alpha/2$ соответственно, которые по заданной доверительной вероятности $(1-\alpha)$ можно найти по табл. П II;

Решая неравенство

$$u_{\alpha/2} < \frac{\overline{X}_в - a}{\sigma/\sqrt{n}} < u_{1-\alpha/2}$$

относительно a, имеем интервальную оценку

$$\overline{x}_в - \Delta < a < \overline{x}_в + \Delta \qquad (76)$$

параметра a при известном значении σ; величину

$$\Delta = \frac{\sigma}{\sqrt{n}} u_{\alpha/2} \qquad (76')$$

называют <u>точностью оценки</u>.

Полученную интервальную оценку часто называют классической, для которой характерны следующие особенности.

С одной стороны, увеличение объема выборки n, согласно (76'), приводит к уменьшению Δ (сужение доверительного интервала), т.е. к увеличению точности оценки; с другой стороны, стремление увеличить доверительную вероятность $(1-\alpha)$ приводит к увеличению аргумента $t = \frac{\Delta}{\sigma/\sqrt{n}}$, так как функция $F(t)$ в (75') возрастающая, и, таким образом, приводит согласно (76') к увеличению Δ. Увеличение $(1-\alpha)$ (т.е. уменьшение α) влечет за собой уменьшение точности оценки (уширение доверительного интервала).

Заметим, что если требуется оценить математическое ожидание с наперед заданной точностью Δ и надежностью $(1-\alpha)$, то минимальный объем выборки n, который обеспечивает эту точность, равен, согласно (76'), $n = \left(\frac{\sigma}{n} u_{\alpha/2}\right)^2$.

<u>При оценке математического ожидания с неизвестным средним квадратическим отклонением σ</u> нельзя воспользоваться предыдущими результатами, так как в статистике (75) содержится неизвестная вероятностная характеристика σ. В этом случае по выборке объема n вместо $\overline{X}_в$ рассматривают новую случайную величину T (ее реализации будем обозначать через t):

$$T = \frac{\overline{X}_в - a}{S/\sqrt{n}}, \qquad (77)$$

которая имеет распределение Стьюдента (73) с $k = n-1$ степенями свободы (см. § 6 данной главы); S^2 – "исправленная" дисперсия /см. формулу (41)/. Затем, используя распределение (73), вычислим /см. формулу (66)/ вероятность события (77) и приравняем заданной $(1-\alpha)$;

$$P\left(\overline{x}_в - t_{1-\alpha/2}(k)\cdot\frac{S}{\sqrt{n}} < a < \overline{x}_в + t_{1-\alpha/2}(k)\frac{S}{\sqrt{n}}\right) = 2\int_0^{t_{1-\alpha/2}} f_T(x)\,dx = 1-\alpha.$$

Таким образом, используя понятие квантиля $t_p(k) = -t_{1-p}(k)$ распределения Стьюдента, для параметра a имеем интервальную оценку

$$\overline{x}_в - t_{1-\alpha/2}\frac{S}{\sqrt{n}} < a < \overline{x}_в + t_{1+\alpha/2}\frac{S}{\sqrt{n}}. \qquad (78)$$

<u>Пример I3</u>. С доверительной вероятностью $(1-\alpha) = 0,95$ найти интервальную оценку для математического ожидания $a = MX$ нормально распределенной случайной величины X генеральной совокупности – числа циклов растяжения (см. пример I).

<u>Решение</u>. В этой задаче инженерного содержания $n = 216$ (см. пример I), $\overline{x}_в = M^*X = 153,0$ (см. пример 3) и $S = 13,5018$ (см. пример 8); случайная величина X распределена по нормальному закону (см. пример 6). По доверительной вероятности $(1-\alpha) = 0,95$ при $k = n-1 = 215$ по табл. П V находим $t_{0,975}(215) \approx 1,97$ или более точно по формуле (73')

$$t_{0,975}(215) = 215\left[1 - \frac{1}{4\cdot 215} - \frac{(-1,96)^2}{2\cdot 215}\right]^{-1/2} = 1,97.$$

Таким образом, для параметра a при неизвестном σ имеем интервальную оценку

$$151{,}1915 < a < 154{,}8085 .$$

Оценка среднего квадратического отклонения нормального распределения $N(a,\sigma)$. Пусть $x_1, x_2, \dots, x_n$ – выборка из нормально-распределенной генеральной совокупности $N(a,\sigma)$. Тогда, согласно § 6 данной главы, статистика $\frac{n-1}{\sigma^2} S^2$ имеет распределение $\chi^2(n-1)$, причем статистики $\bar{x}_в = \frac{1}{n}\sum_i X_i$ и $S^2 = \frac{1}{n-1}\sum_i (X_i - \bar{x}_в)^2$ – независимые случайные величины. Необходимо найти интервальные оценки для σ по s "исправленному" среднему квадратическому отклонению /см. формулу (41)/. Функция $\frac{n-1}{\sigma^2} s^2$ непрерывна и строго монотонна по σ. Значит, согласно § 5 данной главы, с вероятностью $(1-\alpha)$ выполняется неравенство

$$V_1 < \frac{n-1}{\sigma^2} s^2 < V_2 , \quad \text{или} \quad \frac{n-1}{S^2} V_2 < \sigma^2 < \frac{n-1}{S^2} V_1 ,$$

где $V_2 > V_1$.

Следуя 3-му методу (§ 5 данной главы: $V_1 = \chi^2_{\alpha/2}(k)$, $V_2 = \chi^2_{1-\alpha/2}(k)$ – квантили распределения Стьюдента уровня $\alpha/2$ и $1-\alpha/2$ соответственно), получим интервальные оценки для дисперсии и среднего квадратического отклонения:

$$\frac{n-1}{\chi^2_{1-\alpha/2}} s^2 < \sigma^2 < \frac{n-1}{\chi^2_{\alpha/2}} s^2 , \tag{79}$$

$$\frac{\sqrt{n-1}}{\chi_{1-\alpha/2}} s < \sigma < \frac{\sqrt{n-1}}{\chi_{\alpha/2}} s . \tag{80}$$

Пример 14. С доверительной вероятностью $(1-\alpha) = 0{,}95$ найти интервальную оценку для среднего квадратического отклонения σ нормально распределенной случайной величины X – числа циклов растяжения (см. примеры 1, 8).

Решение. Для этой инженерной задачи имеем: $(1-\alpha) = 0{,}95$; $n = 216$ (см. пример 1); $s = 13{,}5013$ (см. пример 8); $\chi^2_{0{,}975}(215) = 257{,}0165$ и $\chi^2_{0{,}025}(215) = 176{,}9020$ (см. упр. 1 данной главы). По формуле (80) находим доверительный интервал

$$12{,}3487 < \sigma < 14{,}8843 .$$

Глава IУ. СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

§ 1. Основные сведения о статистических критериях

По статистическому материалу объема n желательно иметь довольно четкое представление о распределении случайной величины X генеральной совокупности. Статистической гипотезой называется предположение о виде неизвестного распределения или о параметрах известного распределения. Обычно выдвинутую гипотезу называют основной (нулевой) гипотезой H_0, наряду с которой имеет смысл рассматривать еще и альтернативную (конкурирующую) гипотезу H_1.

Например, требуется проверить гипотезу $\widetilde{H}_0$ о равенстве параметра ϑ некоторому заданному значению ϑ_0 для нормально распределенной случайной величины X, или кратко $\widetilde{H}_0 : \vartheta = \vartheta_0$. А по результатам примеров 3, 6, 14 настоящего пособия относительно случайной величины X можно выдвинуть гипотезы:

$H_0^{(1)}$: "случайная величина X распределена по закону $N(a,\sigma)$, $a = 153{,}0$; $\sigma = 13{,}5$";

$H_0^{(2)}$: "случайная величина X распределена по закону $N(a,\sigma)$, $a = 153{,}0$; $12{,}35 < \sigma < 14{,}9$".

Если проверяется гипотеза $\widetilde{H}_0$, то в качестве альтернативной гипотезы можно рассматривать одну из гипотез:

$$\widetilde{H}_1^{(1)}: \vartheta < \vartheta_0 ; \quad \widetilde{H}_1^{(2)}: \vartheta > \vartheta_0 ; \quad \widetilde{H}_1^{(3)}: \vartheta \neq \vartheta_0 ; \quad \widetilde{H}_1^{(4)}: \vartheta = \vartheta_1 ,$$

где заданное значение ϑ_1 не равно ϑ_0. Очевидно вид как основной, так и альтернативной гипотезы определяется исследователем в зависимости от поставленной цели и конкретной формулировки исследуемой задачи.

Статистическая гипотеза называется простой, если она однозначно определяет распределение случайной величины X, в противном случае гипотеза H будет сложной. В вышерассмотренных гипотезах простыми будут $\tilde{H}_0$, $\tilde{H}_1^{(4)}$, $H_0^{(1)}$.

Статистическая гипотеза называется параметрической, если известно распределение случайной величины X и по выборке объема n необходимо проверить лишь предположение о значении параметров этого распределения.

Всякое правило, по которому принимается решение о принятии (или отклонении) основной гипотезы H_0, называется критерием K. Статистикой критерия U (или просто критерием U) называют случайную величину U, с помощью которой проверяют нулевую гипотезу H_0. Эмпирическим (наблюдаемым) значением $U_в$ называют то значение критерия U, которое вычислено по выборке (на основе опытных данных). Так, например, при проверке вышерассмотренной простой параметрической гипотезы $\tilde{H}_0: \vartheta = \vartheta_0$ в качестве подходящей статистики критерия лучше выбрать ту же статистику, что и для оценки параметра ϑ, т.е. $\hat{\vartheta}$ (см. гл. Ш, § 7). Именно этот принцип чаще всего используется при проверке параметрических гипотез. Примеры построения подходящих статистик для проверки непараметрических гипотез будут расcмотрены ниже.

В основу статистической проверки гипотез обычно закладывается принцип, согласно которому маловероятные случайные события считаются практически невозможными, а события с вероятностью, близкой к единице, - практически достоверными. Перед проведением эксперимента исследователь должен в зависимости от поставленной им цели и конкретной формулировки задачи количественно определить понятие практической невозможности (или практической достоверности) случайного события. С этой целью задаются некоторой малой вероятностью α, называемой уровнем значимости. Дальнейшие рассуждения удобно провести опираясь на теоретико-множественный подход. Пусть Y - множество значений статистики U, а $Y_k \subseteq Y$ - подмножество, для которого при условии истинности гипотезы H_0 вероятность попадания статистики U в Y_k равна α, т.е.

$$P(U \in Y_k / H_0) = \alpha . \qquad (81)$$

Используя формулу (8I), критерий (правило), основанный на количественной оценке понятия практической невозможности, можно сформулировать так: отклонить гипотезу H_0, если $U_в \in Y_k$; гипотеза H_0 не противоречит полученным результатам случайного

эксперимента, если $U_в \in Y \setminus Y_k$.

Критерий, построенный с помощью заранее заданного уровня значимости α, называют критерием значимости. Сразу же заметим, что критерием согласия называют критерий проверки гипотезы о предполагаемом законе (виде) неизвестного распределения генеральной совокупности. Принципы построения таких критериев будут рассмотрены ниже.

Подмножество Y_k (подмножество $Y \setminus Y_k$) всех значений статистики U критерия, при которых принимается решение отклонить гипотезу H_0 (принять гипотезу H_0), называют критической областью (областью принятия гипотезы H_0). Точки, отделяющие критическую область от области принятия гипотезы, называют критическими точками $U_{кр}$. Уровень значимости α как количественная мера понятия практической невозможности случайного события определяет "размах" критической области. Положение же критической области на множестве значений статистики U всецело зависит от выбранной исследователем альтернативной гипотезы H_1. Сказанное поясним на примере вышерассмотренных (и специально для этих целей помеченных знаком $\sim$) основной гипотезы $\tilde{H}_0: \vartheta = \vartheta_0$ и одной из альтернативных гипотез $\tilde{H}_1^{(1)}$, $\tilde{H}_1^{(2)}$, $\tilde{H}_1^{(3)}$. Пусть $f_U(u/H_0)$ - плотность распределения статистики U критерия значимости при условии, что истинна гипотеза H_0. Тогда, если проверяется гипотеза $H_0: \vartheta = \vartheta_0$ и H_1 задается исследователем в виде $\tilde{H}_1^{(1)}: \vartheta < \vartheta_0$, то согласно формуле (8I) критическая область определяется неравенством $U < u_\alpha$, где u_α - квантили распределения статистики U (в рассматриваемом примере квантили нормального распределения уровня α). Такая критическая область находится на левом "хвосте" распределения статистики U (рис. 6а) и называется левосторонней критической областью. В случае же альтернативных гипотез $\tilde{H}_1^{(1)}$ и $\tilde{H}_1^{(3)}$ аналогичные рассуждения соответственно приводят к понятию правосторонняя ($U > u_{1-\alpha}$) и двусторонняя ($u_{\alpha/2} < U < u_{1-\alpha/2}$) критические области (рис. 6б,в). Аналогичные рассуждения справедливы и в общем случае произвольной плотности распределения статистики U (по этой причине знак $\sim$ в обозначениях к рис. 6 опущен).

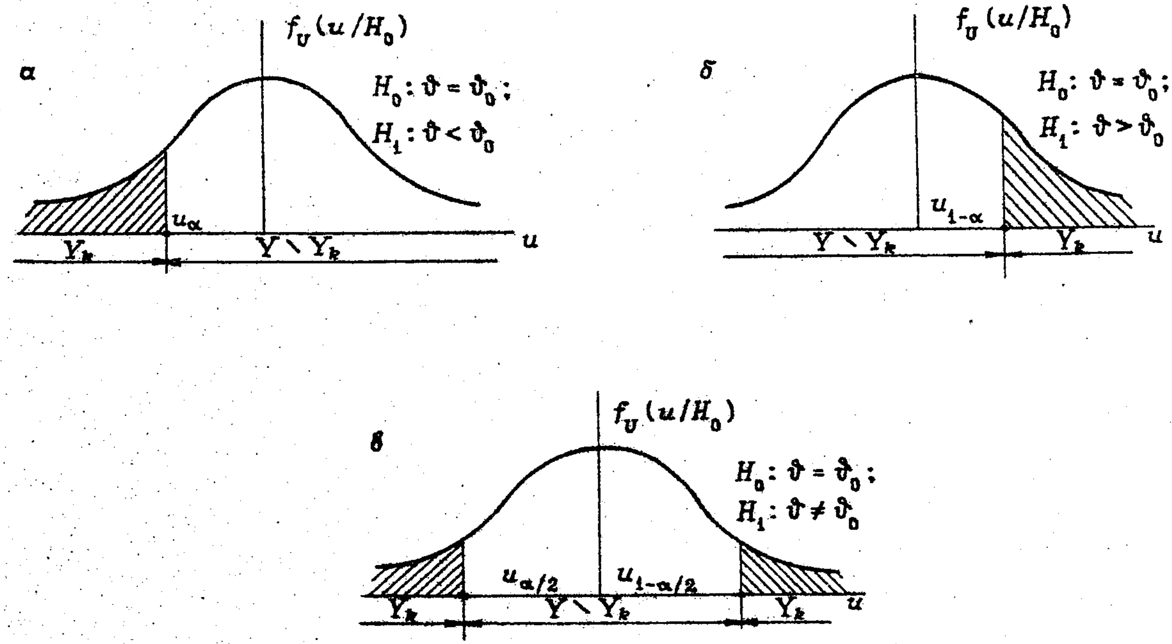


Рис. 6. Левосторонняя а), правосторонняя б) и двусторонняя в) критические области

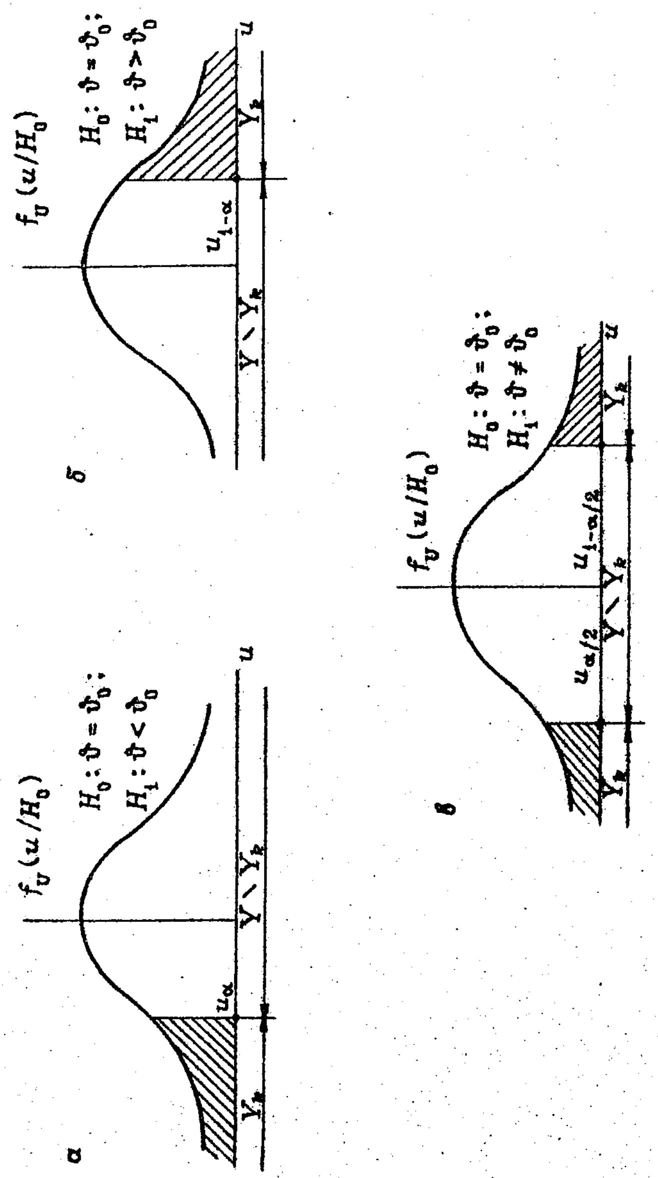


Рис. 6. Левосторонняя а), правосторонняя б) и двусторонняя в) критические области

К сожалению, в настоящем пособии не представляется возможным более детально рассмотреть вопросы, связанные с параметрическими гипотезами и их приложениями. Такой материал можно найти в [3, 4, 8]. Здесь же обратим внимание на следующие обстоятельства, имеющие практическое значение.

I. Проверка статистических гипотез с использованием критериев значимости может быть выполнена в рамках интервальных оценок (см. гл. Ш, § 5-7). В этом случае одностороннему критерию значимости (см. рис. 6а,б) ставится в соответствие односторонний доверительный интервал; двустороннему критерию значимости (см. рис. 6в) – двусторонний доверительный интервал. Основная гипотеза принимается, если значение ϑ_0 накрывается соответствующим доверительным интервалом; в противном случае гипотеза H_0 отвергается. Если проверяется гипотеза H_0: $\vartheta_1 = \vartheta_2$ (например, сравнение двух выборочных средних произвольно распределенных генеральных совокупностей по их большим независимым выборкам), то рассматривается доверительный интервал для разности $\vartheta_1 - \vartheta_2$. Гипотеза H_0 принимается, если соответствующий доверительный интервал накрывает нулевое значение.

2. Статистическое решение может быть и ошибочным: наблюдаемое значение критерия $U_в$ может оказаться в критической области не потому, что основная гипотеза неверна, а по другим причинам (малый объем выборки, просчеты в методике проведения эксперимента, неудачная группировка статистического материала и т.д.). При этом различают ошибки первого и второго рода (вида). Ошибка первого рода состоит в том, что основная гипотеза H_0 отклоняется, в то время как она истинна; согласно формуле (8I) и рис. 6 вероятность такой ошибки равна уровню значимости α. Ошибка второго рода состоит в том, что гипотеза H_0 принимается, но в действительности верна альтернативная гипотеза H_1. Вероятность ошибки второго рода $(1-\alpha)$ вычисляется (при простой альтернативной гипотезе H_1) по формуле (см. рис. 6)

$$P[U \in (Y \setminus Y_k)/H_1] = 1 - \alpha. \qquad (82)$$

Разумеется, что последствия, связанные с этими ошибками, могут оказаться также различными. Например, если на основе анализа статистического материала о выделении газа в шахте отвергнуто правильное решение "продолжать эксплуатацию шахты", то эта ошибка первого рода приводит к материальному ущербу, если же принято неправильное решение "продолжать эксплуатацию шахты", несмотря на опасность взрыва газа, то эта ошибка второго рода может привести (помимо материального ущерба) еще и к катастрофическим последствиям.

3. Пусть вероятность ошибки первого рода α задана, а вероятность ошибки второго рода не должна превышать (например, по техническим соображениям) заданного значения β. Тогда минимальный объем выборки можно найти из решения системы

$$\begin{cases} P(U \in Y_k / H_0) = \alpha, \\ P(U \in (Y \setminus Y_k)/H_1) \leq \beta. \end{cases} \tag{83}$$

4. С помощью статистических критериев отвергают гипотезу более категорично, чем принимают. В самом деле, достаточно привести один пример, противоречащий некоторому общему утверждению, чтобы это утверждение отвергнуть.

§ 2. Вычисление теоретических частот

Случай нормально распределенной случайной величины

Пусть генеральная совокупность распределена нормально. Для вычисления теоретических частот необходимо выполнить следующие операции (табл. 5):

1. Весь интервал наблюдаемых значений случайной величины (выборки объема n) делят (см. пример 1) на g частичных интервалов (x_{i-1}, x_i) одинаковой длины. Находят срединные значения частичных интервалов $\tilde{x}_i = \frac{x_{i-1} + x_i}{2}$; в качестве частоты m_i для $\tilde{x}_i$ принимают число вариант, которые попали в i-й частичный интервал. Заметим, что в случае сгруппированного статистического материала в качестве частичных интервалов разбиения обычно берут классы, на которые группируется исходный статистический материал.

2. Вычисляют, например, методом произведений выборочную среднюю $\bar{x}_в$ и выборочное среднее квадратическое отклонение $\sigma_в$ (см. пример 3).

Таблица 5

Вычисление теоретических частот нормально распределенной случайной величины. Критерий Пирсона

Срединное значение интервала	Границы интервалов группировки	Частота m_i	$z_i = \frac{x_i - \bar{x}_в}{s}$	$\Phi(z_i)$	$P_i = \Delta\Phi_i$	Теоретическая частота $\tilde{m}_i = nP_i$	$\frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i}$
	115,5		-2,78	0,0027	–	–	–
117,5		2					
	119,5						
121,5		1					
	123,5	} 10	–	–	0,0532	11,4912	0,1935
125,5		5					
	127,5						
129,5		2					
	131,5		-1,59	0,0559			
133,5		11			0,0409	8,8344	0,5308
	135,5		-1,3	0,0968			
137,5		16			0,0619	13,3704	0,5172
	139,5		-1,0	0,1583			
141,5		16			0,0833	17,9928	0,2207
	143,5		-0,7	0,2420			
145,5		22			0,0989	21,3624	0,0190
	147,5		-0,41	0,3409			
149,5		21			0,1153	24,9048	0,6123
	151,5		-0,11	0,4562			
153,5		27			0,1191	25,7256	0,0631
	155,5		0,19	0,5753			
157,5		23			0,1091	23,5656	0,0135
	159,5		0,48	0,6844			
161,5		21			0,0979	21,1464	0,0010
	163,5		0,78	0,7823			
165,5		18			0,0754	16,2864	0,1803
	167,5		1,07	0,8577			
169,5		11			0,0570	12,312	0,1398
	171,5		1,37	0,9147			
173,5		11			0,0378	8,1648	0,9839
	175,5		1,67	0,9525			
177,5		5					
	179,5						
181,5		2 } 9	–	–	0,0423	9,1368	0,0020
	183,5						
185,5		2					
–	187,5		2,56	0,9948	–	–	–
Итого	–	216	–	–	0,9921 ≈ 1,0	214,2936 ≈ 216	$\chi^2 = 3,4771$

3. Нормируют случайную величину X, т.е. переходят к величине $Z = \frac{X - x_в}{\sigma_в}$ и вычисляют концы интервалов (z_{i-1}, z_i):

$$z_i = \frac{x_i - \bar{x}_в}{\sigma_в}; \quad z_{i-1} = \frac{x_{i-1} - \bar{x}_в}{\sigma_в}. \tag{84}$$

4. Для каждого интервала группировки вычисляется вероятность попадания случайной величины X в интервал (x_{i-1}, x_i). Она равна

$$P_i = \Delta\Phi_i = \Phi(z_i) - \Phi(z_{i-1}), \tag{85}$$

где

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-y^2/2}\, dy -$$

функция Лапласа, которая табулирована (см. табл. ПII).

5. Объем выборки n умножается на эту вероятность, что дает теоретическую частоту

$$\tilde{m}_i = n P_i. \tag{85'}$$

Пример I5. Пусть число циклов растяжения парашютной ткани распределено по нормальному закону (см. примеры I-3). Вычислить теоретические частоты.

Решение. Применим рассмотренный метод вычисления теоретических частот по этапам (см. табл. 5).

I. Используя метод моментов, параметры нормального распределения оценим как (см. примеры 3 и 8) $\bar{x}_в = 153{,}0$; $s = 13{,}5$.

2. Для каждой границы интервалов группировки определим

$$z_i = \frac{x_i - 153}{13{,}5}.$$

Заметим, что первые четыре интервала объединены, как и три последующих. Этим обеспечивается требование, чтобы в интервале частоты были не намного меньше десяти [7] .

3. Для каждого значения z_i четвертого столбца с помощью табл. ПII находим значения $\Phi(z_i)$ (табл. 5, пятый столбец), затем вычисляем разности $\Delta\Phi_i = \Phi(z_i) - \Phi(z_{i-1})$ (табл. 5, шестой столбец).

4. Теоретические частоты получим путем умножения значений P_i (см. табл. 5) на объем выборки $n = 216$. Результаты вычислений приведены в седьмом столбце табл. 5 и изображены графически на рис. 4.

Вычисление теоретических частот в произвольном случае

Пусть $x_1, \dots, x_n$ - выборка наблюдений случайной величины X некоторой генеральной совокупности. Пусть также выполнена первичная математическая обработка результатов эксперимента (гл. I), найдены выборочные числовые характеристики (гл. II). На основе этих результатов выдвигается и проверяется гипотеза H_0 (см. пример 6), утверждающая, что случайная величина X имеет функцию распределения $F_X(x/\vartheta)$ или соответствующую ей плотность $f_X(x/\vartheta)$; $\vartheta = (a_0, a_1, \dots)$ – параметры этого распределения. Дальнейшие расчеты таковы:

I. В зависимости от вида функции $f_X(x/\vartheta)$ (см. гл. IУ, § 4) оценивают параметры ϑ выбранной статистической модели. При этом стремятся, по возможности, получить состоятельные, несмещенные и эффективные оценки.

2. Пусть $\vartheta = \vartheta_0$. С помощью выбранной функции распределения $F_X(x/\vartheta_0)$ для каждого интервала группировки (x_{i-1}; x_i) по формуле

$$P_i = \Delta F_X(x_i/\vartheta_0) = F_X(x_i/\vartheta_0) - F_X(x_{i-1}/\vartheta_0) \tag{86}$$

вычисляют вероятность попадания случайной величины X в этот интервал.

3. По формулам (85') находят теоретические частоты $\tilde{m}_i$.

§ 3. Критерий Пирсона χ^2 (критерий согласия "хи-квадрат")

Как отмечалось в § I данной главы, частным случаем понятия статистического критерия является критерий согласия U, под которым понимают некоторое правило проверки гипотезы о предполагаемом законе (виде) неизвестного распределения генеральной совокупности. Естественно, это правило должно содержать в себе требование "наилучшего согласия" теоретического $F_X(x)$ и статистического (эмпирического) $F^*(x)$ распределений, т.е. содержать в себе некоторую меру их расхождения. По этой причине критерий согласия может быть выбран различными способами. Например, если в качестве критерия U взять максимальное расхождение (максимальное отклонение по модулю λ) между $F_X(x)$ и $F^*(x)$, то можно прийти к λ-критерию Колмогорова [I]. Если в качестве критерия U выбрать сумму квадратов отклонений теоретических и соответствующих им наблюдаемых частот с некоторыми "весами", то можно прийти к критерию Пирсона.

Заметим, что для любого критерия согласия справедливо следующее утверждение, теоретическое обоснование которого можно найти в полных курсах по математической статистике [6, 8]. Допустим, что критерий выбран тем или иным способом; критерий U является случайной величиной, закон распределения которой зависит от закона распределения изучаемой случайной величины $F(x)$ и от объема выборки n. Если нулевая гипотеза H_0 верна, то закон распределения для U полностью определяется функцией $F(x)$ и числом n. Оказывается, что при некоторых способах построения критерия U его закон распределения обладает весьма простыми свойствами и при достаточно большом объеме выборки практически не зависит от функции $F(x)$. Именно такими мерами расхождения между $F(x)$ и $F^*(x)$ и стараются пользоваться в математической статистике в качестве критериев согласия.

Рассмотрим один из наиболее часто применяемых на практике критериев согласия – так называемый "критерий χ^2 " Пирсона.

Предположим, что проведено n независимых экспериментов; выполнена первичная математическая обработка (см. гл. I, II), т.е. статистический материал сгруппирован по классам (их число равно g), найдены границы, срединные значения классов x_i и соответствующие им частоты m_i/n. Затем предположим, что из сопоставления гистограммы относительных частот и кривых наиболее часто употребляемых распределений (см. пример 6) подобран "на глаз" закон распределения, заданный плотностью $f(x)$.

Назовем этот закон распределения "теоретическим" (описывающим генеральную совокупность).

Требуется проверить, согласуются ли результаты эксперимента с нулевой гипотезой H_0 : "случайная величина X имеет выбранный теоретический закон $f(x)$ ".

В качестве критерия U – меры расхождения между теоретическим и статистическим распределениями – рассмотрим сумму квадратов отклонений ($p_i - m_i/n$), взятых с некоторыми "весами" C_i :

$$U = \sum_{i=1}^{g} C_i\left(\frac{m_i}{n} - P_i\right)^2, \quad \tilde{m}_i = nP_i . \tag{87}$$

Коэффициенты C_i здесь введены потому, что в общем случае отклонения, относящиеся к различным классам, нельзя считать равноправными по значимости. Действительно, одно и то же по абсолютной величине отклонение $m_i/n - P_i$ может быть малосущественным, если сама вероятность P_i велика, и очень заметным, если она мала. По этой причине "веса" C_i следует взять обратно пропорциональными вероятностям классов, т.е. $C_i = A/P_i$.

Пирсон теоретически обосновал [6], что если принять коэффициент пропорциональности A равным n, т.е. $C_i = n/P_i$, то при достаточно больших объемах выборки n закон распределения критерия

$$U \equiv \chi^2 = \sum_{i=1}^{g} \frac{n}{P_i}\left[\frac{m_i}{n} - P_i\right]^2 = \sum_{i=1}^{g} \frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i} \tag{87'}$$

практически не зависит от функции $F(x)$ и от объема выборки n, а зависит лишь от числа интервалов группировки g и при больших n приближается к так называемому распределению χ^2 ("хи-квадрат" распределения Пирсона) с одним параметром τ (см. гл. IУ, § 4):

$$f(u;\tau) = \begin{cases} \dfrac{1}{2^{\tau/2}\Gamma\left(\frac{\tau}{2}\right)} \cdot u^{\frac{\tau}{2}-1} e^{-u/2}, & u \geqslant 0, \\ 0, \quad \text{если} \quad u < 0, \end{cases} \tag{88}$$

где $\Gamma(y)$ – известная гамма-функция Эйлера /см. формулу (54)/. Параметр τ распределения (88), называемый числом "степеней свободы", связан с числом g интервалов группировки соотношением

$$\tau = g - \gamma - 1 . \tag{89}$$

Здесь γ – число независимых условий ("связей"), наложенных на частоты m_i при оценке параметров теоретического распределения. Например, если в качестве статистической модели выбрано нормальное распределение с двумя параметрами a и σ, то число таких связей равно двум ($\gamma = 2$) и $\tau = g - 3$. Вычитаемая единица в (89) отражает тот факт, что условие $\frac{1}{n}\sum_{i=1}^{g} m_i = 1$ является обязательным и вводится в рассмотрение во всех случаях.

Гипотеза H_0 согласуется с результатами наблюдений на уровне значимости α /см. формулу (8I) и рис. 6/, если

$$U_в \equiv \chi^2_в < \chi^2_{1-\alpha}(\tau), \qquad (90)$$

где $\chi^2_{1-\alpha}(\tau)$ – квантили уровня $1-\alpha$ χ^2-распределения с τ степенями свободы; если же $U_в \equiv \chi^2_в \geqslant \chi^2_{1-\alpha}(\tau)$, то гипотеза H_0 отвергается. Особо подчеркнем, что при теоретическом обосновании этого критерия Пирсон использовал тот факт, что случайная величина (87'), распределенная строго говоря по биномиальному закону [6], при больших n имеет распределение, близкое к нормальному $N(0,1)$. Чтобы это утверждение было достаточно точным, необходимо, чтобы для всех интервалов выполнялось условие $m_i \geqslant 5$, $i=1,\ldots,q$ (некоторые исследователи полагают $m_i \geqslant 10$). Если в некоторых интервалах это условие не выполняется, то их объединяют с соседними.

<u>Пример I6</u>. Используя критерий Пирсона, при уровне значимости $\alpha = 0{,}05$ установить, случайно или значимо расхождение между опытными и теоретическими частотами числа циклов растяжения при условии, что теоретические частоты вычислены, исходя из нулевой гипотезы о нормальном распределении числа циклов (см. примеры 6, I5).

<u>Решение</u>. В третьем и седьмом столбцах табл. 5 приведены соответственно наблюдаемые m_i и теоретические $\tilde{m}_i = n\,P_i$ частоты числа циклов растяжения, отнесенные к срединным значениям x_i каждого интервала. В восьмом столбце табл. 5 приведено наблюдаемое значение критерия

$$\chi^2 \equiv U = \sum_{i=1}^{g} \frac{(m_i - \tilde{m}_i)^2}{\tilde{m}_i} = 3{,}4771 \; .$$

Подсчитаем число степеней свободы τ : число интервалов группировки первичного статистического материала равно I8 (см. пример I), однако из-за малочисленности первые четыре и

последние три интервала объединены (см. пример I5) и, таким образом, q = I3. При оценке параметров нормального распределения были наложены две "связи" (см. пример 3), $\gamma = 2$. Значит, согласно (89), τ = I3-2-I = I0. По табл. ПІУ по заданному уровню значимости $\alpha = 0{,}05$ и числу степеней свободы τ = I0 находим $\chi^2_{0,95}(10) = 18{,}3$.

<u>Вывод</u>. Так как $\chi^2 < \chi^2_{кр}$ (3,477I < I8,3), то нет оснований отвергнуть гипотезу о нормальном распределении числа циклов растяжения. Иными словами, расхождение между эмпирическими частотами незначимо (случайно).

Критерий Пирсона можно применять аналогичным способом и для других распределений.

§ 4. Статистические модели, наиболее часто используемые в инженерных задачах

Распределение (статистическая модель) $f(x)$, график	Параметры	Математическое ожидание MX	Дисперсия DX	Коэффициенты	
				асимметрии A_x	эксцесса $Э_x$
I	2	3	4	5	6
Нормальное $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}$	$a\in R$, $\sigma>0$	a	σ^2	0	0

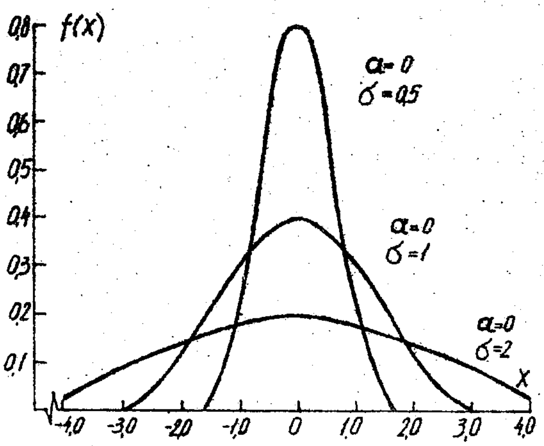


Рис. 7

I	2	3	4	5	6
Логарифмическое нормальное $f(x)=\frac{1}{x\sigma\sqrt{2\pi}}\,e^{-(\ln x-\mu)^2/2\sigma^2}$	$\mu\in R$, $\sigma>0$	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$	*)	**)

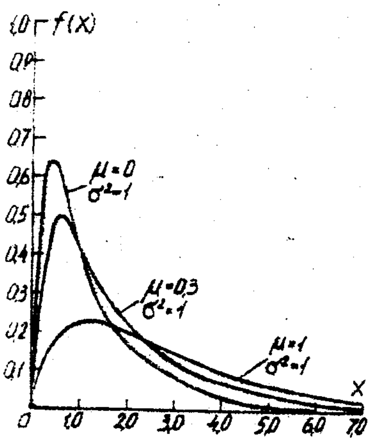


Рис. 8

I	2	3	4	5	6

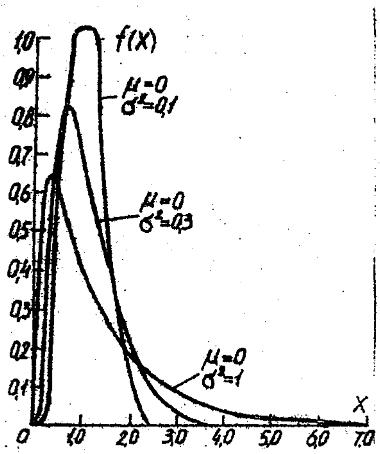


Рис. 9

*) $(e^{\sigma^2}-1)\quad(e^{\sigma^2}+2)$

**) $(w-1)(w^3+3w^2+6w+6)$;
$w=e^{\sigma^2}$

I	2	3	4	5	6
Равномерное $f(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b]\\ 0, & \end{cases}$	$a\in R$ $b\in R$ $a<b$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	0	-1,2

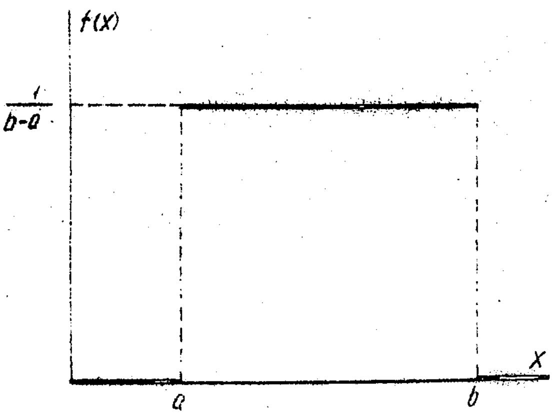


Рис. 10

I	2	3	4	5	6
Релея $f(x)=\begin{cases} \frac{x}{\sigma^2}\, e^{-x^2/2\sigma^2}, & x \geq 0 \\ 0, & \end{cases}$	$\sigma > 0$	$\left(\frac{\sigma^2 \pi}{2}\right)^{1/2}$	$0{,}429\sigma^2$	0,63	0,26

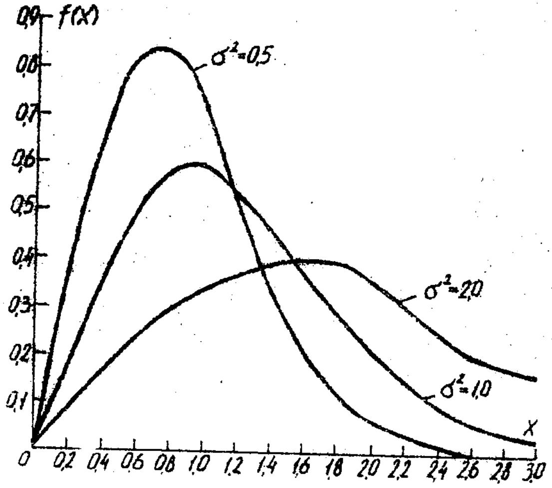


Рис. II

Экспоненциальное $f(x)=\begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \end{cases}$	$\lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	2,0	6,0

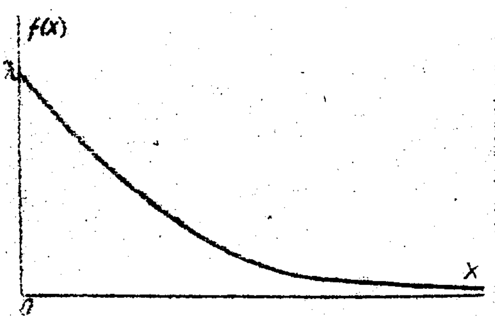


Рис. I2

I	2	3	4	5	6
Гамма-распределение $f(x)=\begin{cases}\dfrac{\lambda^{\eta}}{\Gamma(\eta)}x^{\eta-1}e^{-\lambda x}, & x\geq 0\\ 0, & \end{cases}$	$\lambda>0,$ $\eta>0$	$\dfrac{\eta}{\lambda}$	$\dfrac{\eta}{\lambda^2}$	$\dfrac{2}{\sqrt{\eta}}$	$\dfrac{6}{\eta}$

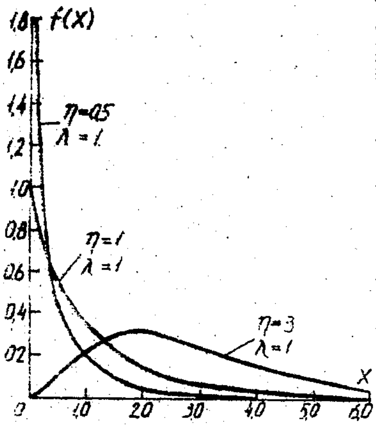


Рис. 13

I	2	3	4	5	6
$f(x)=\begin{cases}\dfrac{\lambda^{\eta}}{\Gamma(\eta)}x^{\eta-1}e^{-\lambda x}, & x\geq 0\\ 0, & \end{cases}$	$\lambda>0$ $\eta>0$	$\dfrac{\eta}{\lambda}$	$\dfrac{\eta}{\lambda^2}$	$\dfrac{2}{\sqrt{\eta}}$	$\dfrac{6}{\eta}$

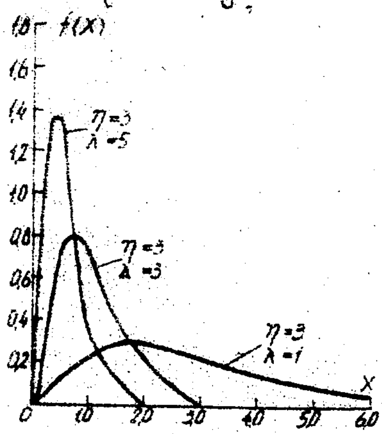
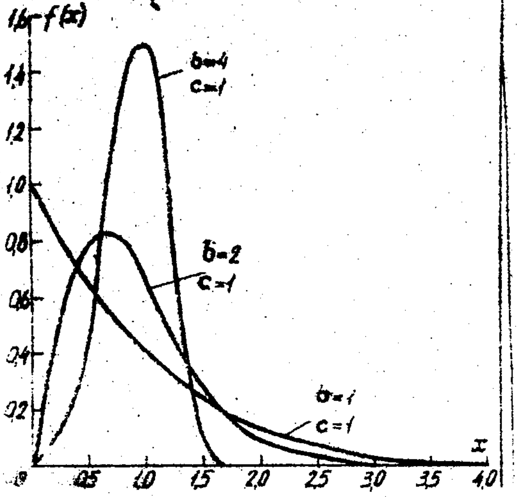
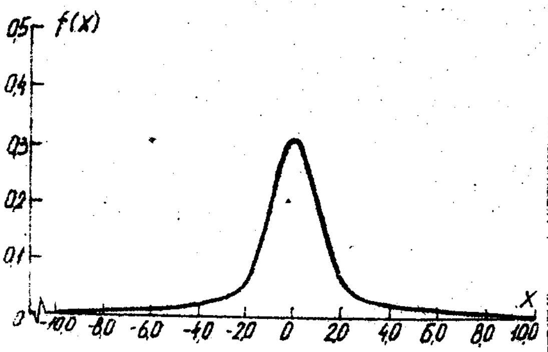


Рис. 14

I	2	3	4	5	6
Вейбулла $$f(x)=\begin{cases}\frac{b}{c}\left(\frac{x}{c}\right)^{b-1}e^{-\left(\frac{x}{c}\right)^{b}}, & x\geqslant 0\\ 0, & \end{cases}$$ 	$b>0$ $c>0$	$c\Gamma\left(1+\frac{1}{b}\right)$	$c^2\left\{\Gamma\left(1+\frac{2}{b}\right)-\left[\Gamma\left(1+\frac{1}{b}\right)\right]^2\right\}$	*)	**)

*) $$\frac{\Gamma\left(1+\frac{3}{b}\right)-3\Gamma\left(1+\frac{2}{b}\right)\Gamma\left(1+\frac{1}{b}\right)+2\left[\Gamma\left(1+\frac{1}{b}\right)\right]^3}{\left\{\Gamma\left(1+\frac{2}{b}\right)-\left[\Gamma\left(1+\frac{1}{b}\right)\right]^2\right\}^{3/2}};$$

**) $$\frac{\Gamma\left(1+\frac{4}{b}\right)-4\Gamma\left(1+\frac{3}{b}\right)\Gamma\left(1+\frac{1}{b}\right)+6\Gamma\left(1+\frac{2}{b}\right)\left[\Gamma\left(1+\frac{1}{b}\right)\right]^2-3\left[\Gamma\left(1+\frac{1}{b}\right)\right]^4}{\left\{\Gamma\left(1+\frac{2}{b}\right)-\left[\Gamma\left(1+\frac{1}{b}\right)\right]^2\right\}^{3/2}}$$

I	2	3	4	5	6
Коши $$f(x)=\frac{1}{\pi\sigma}\left[1+\frac{x-\mu}{\sigma^2}\right]^{-1},\quad x\in R$$  Рис. 16	$\mu\in R$, $\sigma>0$	Конечное значение отсутствует			

I	2	3	4	5	6
χ^2 - распределение $f(x) = \begin{cases} \dfrac{1}{2^{k/2}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k-2}{2}} e^{-\frac{x}{2}}, & x > 0 \\ 0 \end{cases}$	$k > 2$	k	$2k$	$2\sqrt{\frac{2}{k}}$	$\frac{12}{k}$

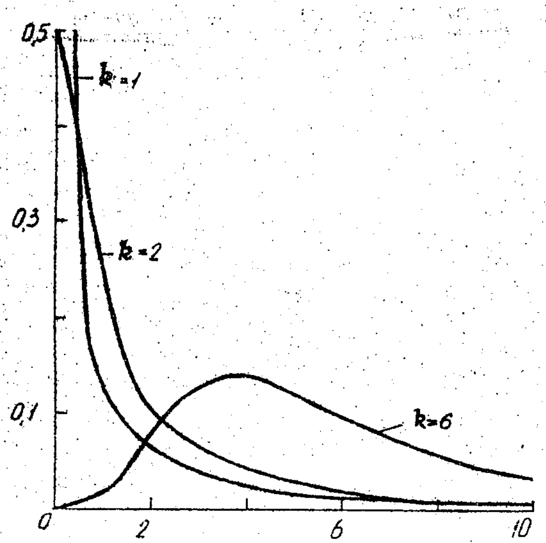


Рис. 17

I	2	3	4	5	6
Стьюдента (t -распределение) $f(x) = \dfrac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma(k/2)\sqrt{k\pi}} \left(1 + \dfrac{x^2}{k}\right)^{-\frac{k+1}{2}}$	$k > 5$	0	$\frac{k-1}{k-3}$	0	$\frac{3(k-3)}{k-5}$

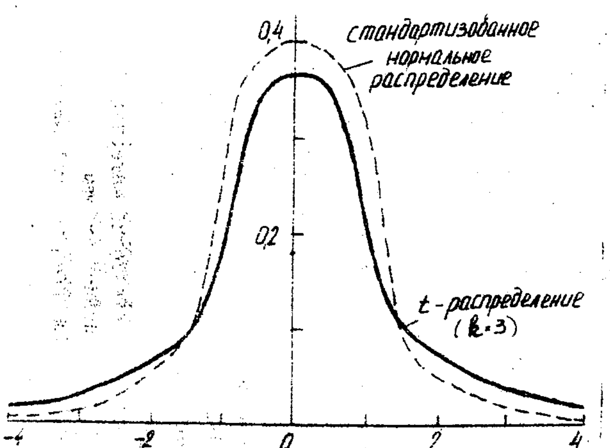


Рис. 18

Глава У. ЗАДАНИЕ. ВАРИАНТЫ ЗАДАЧ

Цель задания – на примере выборки, полученной при решении типичных задач из различных областей науки и техники, изучить основные положения теории статистического вывода, познакомиться с постановкой и решением основных задач математической статистики.

Задание состоит из пяти частей, в каждой из которых решается одна из основных задач математической статистики, имеющая практическое применение.

Задание

Часть I

I. Изучить теоретический материал [4 , гл. ХУ, § I–8] и гл. I данной работы.

2. Для заданной выборки найти оптимальную величину интервала группировки, сгруппировать с помощью штриховального листа статистический материал.

3. Найти частоту, относительную частоту, накопленную частоту и накопленную относительную частоту каждого интервала группировки. Выполнить графическую иллюстрацию.

4. Найти статистическую функцию распределения, построить ее график.

5. Вычислить относительную частоту попадания случайной величины в заданный интервал (интервал указывает преподаватель).

Часть II

6. Изучить теоретический материал [4 , гл. ХУI, § I–4, 8–I0; гл. ХУП, § I–4] и гл. II данной работы.

7. Вычислить выборочные медиану и моду.

8. Найти методом произведений выборочную среднюю и выборочную дисперсию.

9. Найти исправленную выборочную дисперсию. Выполнить поправки Шеппарда.

Часть Ш

I0. Изучить теоретический материал [4 , гл. ХУП, § I–5, 8] и гл. II данной работы.

II. Найти методом произведений выборочные центральные моменты третьего и четвертого порядков.

I2. Вычислить коэффициенты асимметрии и эксцесса.

Часть IУ

I3. Изучить теоретический материал гл. Ш данной работы.

I4. На основе анализа результатов исследования выборки произвести подбор распределения (статистической модели).

I5. Методом моментов и методом максимального правдоподобия найти точечные оценки параметров распределения.

I6. Найти интервальные оценки параметров распределения (выполняется по усмотрению преподавателя).

Часть У

I7. Изучить теоретический материал [4, гл. ХУП, § 6–7] и гл. IУ данной работы.

I8. Вычислить теоретические частоты.

I9. Проверить с помощью критерия Пирсона гипотезу о виде распределения.

Варианты задач

I. При производстве миниатюрных радиоламп их выводы устанавливаются автоматически; изогнутость выводов является важным показателем при сборке готовой продукции.

Данные измерения оптическим компаратором изогнутости выводов радиоламп, $I0^{-I}$ мм

20	3I	II6	32	I00	28	I30	97	II	27	I22	29
28	44	I2	46	47	52	3I	I5	2I	32	I4	I9
45	52	9I	35	53	92	38	03	06	37	I42	II7
07	57	46	66	63	5I	56	52	34	43	29	40
35	6I	7I	74	83	68	84	67	47	52	54	46
52	76	86	85	78	60	68	60	72	59	6I	57
I7	62	69	82	75	I9	62	69	83	67	70	50
I5	58	4I	44	53	02	54	42	35	75	36	I8
I24	30	52	39	34	23	36	2I	28	99	22	I6
I32	96	II6	27	96	30	25	98	I0	67	II8	90
67	75	65	66								

2. Данные о пределе текучести для 100 образцов из титанового сплава при 1000 фунт/кв.дюйм*

I52	I54	I47	I42	I32	I64	I54	I73	I64	I60
I66	I39	I6I	I63	I52	I50	I56	I54	I60	I35
I54	I50	I4I	I55	I53	I35	I44	I48	I50	I48
I48	I66	I48	I49	I54	I56	I50	I53	I5I	I38
I49	I58	I39	I46	I36	I55	I45	I5I	I54	I4I
I60	I38	I53	I56	I66	I42	I50	I44	I58	I45
I47	I7I	I52	I46	I58	I54	I56	I36	I69	I5I
I67	I58	I68	I57	I36	I47	I30	I4I	I47	I58
I64	I36	I53	I60	I43	I56	I37	I47	I52	I56
I50	I59	I25	I44	I39	I39	I34	I46	I55	I44

* Фунт (*pound*) – единица веса, равная ~ 4,54 Н; дюйм (*inch*) – единица длины, равная 2,54 см

3. Чувствительность канала изображения телевизора в метровом диапазоне, мкВ

20,5	I5,0	2I,5	20,0	I9,0	2I,5	I9,0	I9,0	24,0	28,0	24,0	28,0	24,0	25,0	29,0	25,0
28,0	37,5	26,0	29,0	23,4	I2,6	20,6	27,0	23,2	22,6	28,5	23,0	27,2	25,2	2I,0	24,2
24,2	24,2	25,2	2I,6	2I,0	2I,6	20,8	22,2	30,2	25,0	28,0	25,0	27,0	I7,4	25,8	24,2
23,2	2I,2	26,6	27,0	3I,0	33,4	26,0	27,0	2I,6	30,2	22,8	26,4	25,8	25,2	29,0	25,0
25,2	25,2	25,0	27,3	20,4	22,7	2I,0	26,0	20,0	2I,6	24,0	22,0	27,0	24,2	25,8	26,2
30,0	3I,0	25,0	26,2	20,6	25,2	23,0	25,0	27,0	25,I	22,0	29,2	24,0	30,0	24,5	2I,5
29,0	23,4	23,5	25,9	22,6	25,0	30,0	30,2	32,6	23,8	39,2	25,0	27,2	25,6	23,4	26,2
2I,9	26,9	23,6	26,9	23,I	I9,9	23,4	I9,2	I4,4	20,7	29,2	2I,9	2I,0	2I,9	30,0	22,6
24,6	24,I	20,6	27,8	22,7	23,4	2I,6	24,6	2I,9	23,8	27,2	34,0	25,4	23,2	27,7	23,0
30,0	25,I	22,7	27,8	27,0	22,6	20,7	I9,4	2I,4	23,0	2I,0	24,3	23,0	23,2	29,2	24,4
24,4	2I,8	29,4	30,0	29,7	29,2	23,0	23,4	23,0	25,9	24,6	22,6	29,2	23,4	28,8	25,4
23,8	30,0	27,8	2I,0	28,6	27,2	23,I	26,9	25,9	24,2	3I,2	25,9	23,I	27,6	26,2	22,2
25,9	27,8	20,0	27,0												

3. Чувствительность канала изображения телевизора в метровом диапазоне, мкВ

20,5	I5,0	2I,5	20,0	I9,0	2I,5	I9,0	I9,0	24,0	28,0	24,0	28,0	24,0	25,0	29,0	25,0
28,0	37,5	26,0	29,0	23,4	I2,6	20,6	27,0	23,2	22,6	28,5	23,0	27,2	25,2	2I,0	24,2
24,2	24,2	25,2	2I,6	2I,0	2I,6	20,8	22,2	30,2	25,0	28,0	25,0	27,0	I7,4	25,8	24,2
23,2	2I,2	26,6	27,0	3I,0	33,4	26,[illegible]	27,0	2I,6	30,2	22,8	26,4	25,8	25,2	29,0	25,0
25,2	25,2	25,0	27,3	20,4	22,7	2I,0	26,0	20,0	2I,6	24,0	22,0	27,0	24,2	25,8	26,2
30,0	3I,0	25,0	26,2	20,6	25,2	23,0	25,0	27,0	25,I	22,0	29,2	24,0	30,0	24,5	2I,5
29,0	23,4	23,5	25,9	22,6	25,0	30,0	30,2	32,6	23,8	39,2	25,0	27,2	25,6	23,4	26,2
2I,9	26,9	23,6	26,9	23,I	I9,9	23,4	I9,2	I4,4	20,7	29,2	2I,9	2I,0	2I,9	30,0	22,6
24,6	24,I	20,6	27,8	22,7	23,4	2I,6	24,6	2I,9	23,8	27,2	34,0	25,4	23,2	27,7	23,0
30,0	25,I	22,7	27,8	27,0	22,6	20,7	I9,4	2I,4	23,0	2I,0	24,3	23,0	23,2	29,2	24,4
24,4	2I,8	29,4	30,0	29,7	29,2	23,0	23,4	23,0	25,9	24,6	22,6	29,2	23,4	28,8	25,4
23,8	30,0	27,8	2I,0	28,6	27,2	23,I	26,9	25,9	24,2	3I,2	25,9	23,I	27,6	26,2	22,2
25,9	27,8	20,0	27,0												

4. Точность измерительного прибора, систематическая ошибка которого практически равна нулю, м

38I	42I	372	4I8	392	427	385	358	370
4I2	4II	385	395	382	376	380	383	395
39I	430	39I	377	372	406	429	429	376
43I	405	430	382	429	4I3	42I	395	4I3
430	373	393	375	364	449	382	375	37I
4II	427	362	388	409	400	392	378	42I
399	395	384	373	39I	340	4I0	428	382
397	389	403	440	4I8	4I2	378	398	4I8
365	399	4I8	400	402	405	4I0	423	373
399	389	440	429	369	394	432	390	409
35I	384	425	407	383	4I5	4I8	456	303
398	420	4I8	404	400	383	425	422	388
388	42I	437	4I8	379	383	347	428	388
395	429	363	4I0	384	4I6	380	433	398

5. Расстояние безотказной работы тепловозов (расстояние, пройденное тепловозами до выхода из строя одного из его контрольных приборов), тыс. км

46,0	I20,0	I22,5	93,5	69,5	I02,5	76,5	37,5	22,5	77,0
I07,0	I23,0	48,5	78,5	I08,5	I27,5	5I,5	80,0	II2,5	I3I,5
53,0	8I,5	II3,5	I32,0	54,5	82,0	II6,0	I34,0	57,5	83,0
II7,0	66,5	84,0	II8,5	68,0	9I,5	II9,0	38,5	66,0	43,5
60,5	9I,5	39,0	65,5	I37,5	40,5	99,5	52,5	I43,0	89,5
94,5	80,5	79,0	62,0	87,5	97,5	62,5	64,0	23,5	78,5
6I,0	98,0	62,5	97,5	70,0	65,5	7I,5	99,0	72,5	63,5
47,0	77,0	76,5	64,0	63,5	56,5	77,0	63,5	72,0	66,0
87,5	65,5	55,0	I08,5	99,0	II0,0	86,5	88,0	66,0	I05,5

Указание. В качестве гипотезы H_0 рассмотреть возможность применения статистических моделей: нормальное распределение, гамма-распределение, логарифмически нормальное распределение.

6. Процентное содержание триоксида серы в горной породе некоторого региона, %

I5,6	I5,8	I5,7	I5,8	I5,7	I6,0	I5,7	I5,9	I5,7	I5,8	I5,7
I5,8	I5,4	I5,8	I5,7	I5,7	I5,9	I6,0	I5,7	I6,0	I5,7	I6,0
I5,9	I5,8	I5,5	I6,0	I5,7	I5,7	I5,7	I5,9	I5,7	I5,8	I5,8
I5,I	I5,8	I6,0	I6,2	I5,7	I5,5	I5,9	I5,7	I5,7	I5,3	I5,6
I6,I	I5,7	I6,I	I5,9	I5,8	I6,0	I5,0	I5,7	I5,6	I5,5	I5,8
I5,6	I5,8	I5,8	I5,5	I5,6	I5,6	I5,9	I5,8	I5,9	I5,8	I5,7
I5,5	I5,7	I5,8	I5,9	I5,4	I5,8	I5,3	I5,4	I5,5	I5,7	I5,6
I5,8	I5,9	I5,4	I5,9	I5,6	I5,7	I5,6	I5,7	I5,7	I5,7	I5,7
I5,3	I6,I	I5,6	I6,0	I6,I	I5,6	I5,5	I5,6	I5,7	I5,5	I6,I
I5,8	I5,7	I5,4	I6,3	I5,7	I5,6	I6,2	I5,6	I5,6	I5,3	I5,5
I5,4	I5,9	I5,6	I6,0	I5,7	I5,8	I5,9	I6,0	I6,I	I5,8	I5,9
I5,7	I5,6	I5,7	I5,9	I6,0	I6,I	I5,5				

7. Результаты измерения обхвата груди I20 женщин, см

95	93	89	I00	94	95	94	I0I	90	95
I03	98	99	9I	95	94	95	94	89	93
98	95	93	89	I00	I07	I00	98	I0I	97
90	95	I03	98	99	9I	94	95	94	89
93	98	93	96	I0I	97	I02	97	I06	I0I
96	96	94	I00	95	92	93	96	97	98
99	97	I04	I0I	98	I09	98	I04	95	I00
I02	92	95	99	93	92	97	99	98	I02
98	94	98	97	94	90	95	97	I03	I00
97	9I	96	I08	I00	9I	93	I06	93	97
93	90	95	97	97	99	93	96	I0I	96
I00	I06	I05	94	I02	9I	94	I06	98	I00

8. Результаты измерения обхвата груди I24 мужчин, см

98	92	I0I	I02	99	I09	I0I	I04	94	96	I04	I00	I00	97	I06
I0I	I0I	I02	99	I09	I0I	I04	93	96	I04	I00	II0	97	I06	I0I
I0I	99	I03	I0I	99	93	I00	I03	98	I08	I02	I03	88	97	II6
97	I05	I03	II0	I02	96	I09	I04	II2	97	98	II4	I05	II6	I02
I0I	I09	98	I09	98	I05	I03	I0I	97	92	I06	I09	98	I03	I04
I00	I0I	9I	99	I0I	I0I	I05	97	II0	99	93	I07	88	I03	94
III	98	90	I00	II6	97	I08	I04	II2	96	92	II0	I03	I05	87
98	I09	98	I09	I0I	I02	II0	I05	I09	I03	98	I08	I06	92	97
I0I	I03	I05												

9. Результаты определения выносливости шерстяной ткани при многок атном растяжении при заданной циклической деформации 8 %, число циклов

I02	99	I02	II3	9I	I0I	I07	94	I09	III	I06	95	I06	87	97
I05	I0I	93	98	95	I05	98	I0I	88	99	I00	I07	I08	97	92
I04	I02	97	II4	I0I	97	III	I0I	I04	III	I0I	I03	I0I	92	I02
II0	I06	I05	95	96	I03	I08	93	II2	96	99	II6	I00	II2	I0I
I03	II2	I02	97	95	94	I00	I07	I03	99	I05	I04	II0	I08	98
97	I03	I02	89	92	99	89	I09	98	I0I	I06	I02	99	II0	86
97	I06	I05	97	I0I	I09	96	I04	I03	I09	I03	85	I05	I00	[illegible]
I00	I00	98	I03	I00	II0	99	96	94	I03	II0	I03	I09	99	I02
9I	I00	97	93	II0	I09	I04	I03	I0I	I03	I06	87	I05	96	I0I
I0I	93	98	I06	III	I02	92	98	I09	I04	II4	I08	I03	I0I	70
I08	99	I02	I03	I06	I0I	I05	97	II6	I02	I09	98	97	I00	95

I0. Сформировать до объема n = 200 массив случайных чисел мультипликативным способом по формуле

$$\xi_{n+1} = Д(k\xi_n),$$

где $Д(k\xi_n)$ – дробная часть числа $k\xi_n$; $k = 8t \pm 3$.

Параметр t принять равным 5; значение начального случайного числа ξ_0 – равным 0,I234567.

<u>Указание</u>. Вычисление случайных чисел провести до третьего знака после запятой, округление случайных чисел не производить.

II. Результаты измерения стойкости резца из TI5K6 при скорости резания 0,33 м/с и подаче 0,I2 мм/об, мин

I52	I43	I70	I62	I63	I5I	I64	I6I	I63	I65	I59	I63	I70	I66	I68
I55	I64	I65	I74	I59	I65	I70	I58	I59	I60	I58	I60	I62	I66	I63
I64	I65	I65	I58	I58	I60	I63	I64	I70	I69	I70	I72	I70	I65	I58
I64	I7I	I76	I70	I58	I65	I60	I64	I67	I70	I6I	I60	I65	I65	I58
I70	I68	I68	I60	I64	I58	I60	I62	I56	I70	I63	I60	I63	I68	I62
I65	I63	I63	I65	I58	I68	I64	I7I	I66	I60	I60	I62	I64	I55	I69
I65	I65	I65	I65	I66	I64	I64	I50	I65	I70	I75	I60	I65	I66	I62
I68	I64	I64	I70	I64	I67	I60	I68	I58	I70	I65				

I2. Результаты измерений максимальной скорости испытаний спортивного самолета, м/с

43I	398	423	40I	423	404	389	428	402	404
427	[illegible]	422	409	420	422	397	458	403	4II
398	408	438	4I4	4I3	404	426	434	430	397
383	4I5	4I8	438	394	4I7	4I2	404	389	398
43I	423	40I	423	435	427	428	405	4I4	4I5
439	409	39I	4I6	4I9	40I	372	395	4I8	4I3
407	445	428	420	429	395	433	406	402	398
399	432	405	4I2	425	4I7	424	4I6	396	403
432	402	43I	4I9	423	44I	424	4I0	424	4I3
393	4I2	[illegible]	408	437	4I6	436	4I5	42I	407
404	404	403	434	4I2	4I9	405	402	394	423
398	4I5	40I	398	428	4I6	453	37I	424	4I7

I3. Результаты измерения роста I54 девушек некоторого региона, см

I68	I63	I60	I70	I60	I55	I58	I57	I57	I59	I55	I55	I60	I63
I64	I68	I73	I70	I63	I60	I56	I58	I63	I64	I65	I64	I7I	I63
I72	I68	I65	I68	I70	I68	I59	I72	I66	I54	I65	I64	I64	I68
I65	I54	I67	I59	I60	I64	I65	I64	I69	I58	I63	I56	I70	I74
I79	I72	I63	I62	I60	I64	I70	I74	I67	I67	I54	I64	I70	I60
I67	I67	I65	I68	I58	I56	I67	I55	I62	I70	I70	I70	I64	I68
I60	I66	I62	I64	I62	I65	I57	I66	I55	I58	I60	I62	I63	I67
I57	I64	I63	I58	I68	I58	I64	I62	I64	I66	I70	I62	I68	I69
I67	I74	I69	I75	I68	I66	I68	I68	I68	I66	I70	I60	I65	I70
I68	I62	I55	I68	I64	I63	I66	I68	I64	I65	I66	I56	I65	I64
I59	I56	I63	I64	I65	I65	I57	I70	I66					

I4. Масса одного колоса пшеницы сорта *Sonnora* (Япония) при плотности посева I5x2,5 см, г

I,80	I,40	I,I2	2,30	2,70	3,30	I,30	I,I3	I,70	I,40
I,25	I,90	I,64	I,47	I,65	I,50	I,85	I,68	I,5I	I,48
I,95	0,80	2,80	2,40	2,95	2,50	2,30	2,90	I,84	2,20
I,68	2,50	2,52	I,29	3,30	I,85	2,I0	3,60	2,40	2,55
I,50	I,29	I,85	I,58	I,3I	I,69	I,28	I,90	I,87	I,70
I,49	2,I0	I,90	I,49	I,80	2,45	2,30	3,00	3,I0	3,I0
I,60	I,88	2,20	I,63	0,80	I,63	I,45	I,29	I,47	2,55
I,49	2,40	2,55	I,26	0,80	I,25	2,I0	0,70	2,00	I,85
0,90	I,90	2,I0	2,55	2,55	2,40	0,60	2,I0	0,40	2,50
I,50	I,69	2,70	I,48	I,50	I,69	I,46	I,48	I,52	I,30

I5. Масса одного колоса пшеницы сорта *Sonnora* (Япония) при плотности посева I5x5 см, г

3,9I	4,2I	I,73	2,70	I,57	2,00	4,00	I,I0	I,62	I,30
2,50	I,I0	2,60	3,90	0,70	I,45	I,5I	I,97	I,46	3,82
I,42	I,62	2,45	0,78	3,50	3,75	I,39	2,40	3,80	2,48
I,I0	2,03	I,47	5,40	0,7I	I,4I	I,40	I,48	I,49	5,20
2,35	I,49	I,6I	I,44	2,40	0,75	2,60	2,95	3,00	2,08
I,49	2,85	I,58	3,90	I,59	I,98	0,80	2,80	I,49	I,90
5,I0	I,49	2,0I	3,65	2,08	I,48	3,25	I,50	4,I9	0,94
I,95	2,03	0,80	I,58	I,90	2,02	I,53	0,84	I,85	2,0I
2,02	2,38	I,96	2,I0	2,47	I,4I	2,07	I,50	0,80	I,45
3,80	I,50	I,49	3,98	I,98	2,78	3,95	2,9I	2,50	I,90
I,35	2,I0	0,74	I,28	0,75	I,59	I,50			

Указание. В качестве гипотезы H_0 рассмотреть возможность применения моделей: распределений Реллея, экспоненциальное распределение, гамма-распределение.

I6. Процентное содержание лавсанового волокна в хлопко-лавсановой пряже (данные чулочно-носочной фабрики им. В.П.Ногина), %

I3,39	I3,43	I3,54	I3,64	I3,40	I3,55	I3,40	I3,26	I3,42	I3,50
I3,32	I3,3I	I3,28	I3,52	I3,46	I3,63	I3,38	I3,44	I3,52	I3,53
I3,37	I3,33	I3,24	I3,I3	I3,53	I3,53	I3,39	I3,57	I3,5I	I3,34
I3,39	I3,47	I3,5I	I3,48	I3,62	I3,58	I3,57	I3,33	I3,5I	I3,40
I3,30	I3,48	I3,40	I3,57	I3,5I	I3,40	I3,52	I3,56	I3,40	I3,34
I3,23	I3,37	I3,48	I3,48	I3,62	I3,35	I3,40	I3,36	I3,45	I3,48
I3,29	I3,58	I3,44	I3,56	I3,38	I3,20	I3,54	I3,62	I3,46	I3,47
I3,59	I3,29	I3,43	I3,30	I3,56	I3,5I	I3,47	I3,40	I3,29	I3,20
I3,46	I3,44	I3,42	I3,29	I3,4I	I3,39	I3,50	I3,48	I3,53	I3,34
I3,45	I3,42	I3,29	I3,38	I3,45	I3,50	I3,55	I3,33	I3,32	I3,69
I3,46	I3,32	I3,48	I3,29						

I7. Глубина вдавливания (глубокий отпуск) стальных образцов, мм

9,57	I0,07	I0,77	I0,24	9,98	9,65	9,30	I0,33	II,5I	9,23
I0,32	9,I2	I0,33	9,28	I0,57	I0,24	I0,62	I0,I8	I0,85	II,02
9,78	I0,42	I0,90	I0,23	9,45	I0,50	I0,48	II,II	9,53	I0,05
II,58	9,72	I0,59	9,68	I0,92	9,87	I0,27	I0,22	I0,97	I0,82
I0,66	I0,69	I0,80	9,42	I0,69	I0,54	I0,85	I0,24	I0,48	I0,35
II,07	9,54	II,I8	9,67	II,43	9,80	I0,86	II,25	I0,23	I0,08
9,75	II,05	I0,07	I0,03	I0,57	I0,27	9,97	9,92	I0,62	I0,87
I0,47	I0,I2	I0,08	9,99	9,96	9,85	9,85	I0,63	I0,22	9,30
9,83	I0,75	I0,65	I0,20	9,57	9,89	I0,I7	I0,05	I0,02	I0,35
I0,34	I0,22	9,75	I0,00	9,85	I0,77	II,23	I0,05	I0,30	I0,03
I0,73	9,79	I0,88	I0,03	I0,I7	I0,22	9,I0	I0,02	II,53	II,40
9,80	9,80	9,83	I0,I3	I0,23	I0,50	II,45	I0,5I	I0,67	I0,45
I0,77	9,97	I0,72	I0,55	I0,42	II,66	9,3I	9,46	I0,00	II,35
9,33	I0,05	I0,27	I0,38	I0,24	I0,43	I0,30	II,6I	I0,22	9,08
I0,34	I0,4I	II,22	II,28	9,85	9,63	I0,03	I0,40	I0,93	I0,46

I8. Содержание влаги в 80 кирпичах, используемых для футеровки печи, после хранения их в течение месяца, %

7,I	6,7	7,0	7,3	7,2	7,I	6,9	6,8	7,5	7,0
7,0	7,I	7,I	6,8	7,2	7,0	7,2	6,9	6,7	6,9
6,9	7,0	7,0	6,8	6,9	7,0	7,0	7,I	6,8	7,I
7,2	7,I	6,9	6,7	7,I	6,9	6,9	7,I	7,0	7,3
6,8	7,3	7,4	6,8	7,2	7,2	6,8	6,7	7,3	7,I
6,9	7,5	7,0	6,5	7,I	7,2	7,0	7,0	6,9	7,0
6,7	6,8	7,I	7,2	7,I	7,5	7,I	6,8	6,9	7,2
7,2	6,9	7,I	7,5	7,0	7,I	7,0	7,I	6,8	7,0

I9. Результаты определения линейной плотности стальной проволоки, г/м

38I	388	384	4I8	373	364	376	383	432	428	4I3	4I2	395	420
440	440	409	406	4I6	4I8	398	37I	39I	42I	42I	425	400	39I
4I3	385	425	423	42I	43I	429	4II	4I8	429	4I8	449	380	347
390	382	430	372	430	437	407	402	400	429	380	456	4I8	4II
385	405	363	404	369	340	42I	358	422	373	399	39I	373	4I8
4I8	383	4I2	382	383	428	409	397	427	430	395	4I0	400	405
392	376	433	363	365	395	393	377	392	379	394	4I0	385	370
388	399	389	362	382	382	384	4I5	378	375	395	388	36I	399
384	375	372	427	385	4I0	378	392	398	398	389	403	388	429

20. Время безотказной работы некоторого прибора, тыс. ч

26,7	94,2	74,8	88,7	93,2	78,7	90,5	73,3	76,3	7I,9	80,3	27,3
73,3	69,8	69,I	8I,9	67,7	57,7	68,4	96,I	67,0	64,4	92,3	67,0
39,9	53,8	79,5	74,I	63,8	77,I	86,9	87,8	8I,I	6I,3	97,0	5,5
4I,5	48,7	95,I	7I,2	58,3	53,3	49,2	55,4	50,7	47,7	52,7	60,0
I3,5	50,2	77,9	60,6	45,4	98,0	I00	72,6	44,9	59,5	56,5	56,0
I6,5	42,7	70,5	43,2	4I,9	85,2	38,7	48,2	39,I	44,5	9,5	39,5
26,I	49,7	99,0	45,8	40,3	82,7	86,I	5I,7	83,5	43,6	52,2	5I,2
22,3	30,2	89,6	39,9	33,3	9I,4	38,3	26,2	37,5	36,8	28,3	37,9
65,0	I3,5	84,4	27,3	24,7	66,4	58,9	54,9	46,8	6I,9	47,2	65,7
30,0	42,3	75,6	63,I	62,5	40,7	4I,I	46,3	44,0	37,2	57,I	54,9

Указание. В качестве гипотезы H_0 рассмотреть возможность применения статистических моделей: экспоненциальное распределение, распределение Релея и гамма-распределение.

2I. Результаты определения плотности в петлях трикотажного полотна, петл./5 см

67	65	65	62	63	66	68	7I	68	64	6I	63	60
7I	64	64	69	59	65	64	64	65	64	66	64	62
64	68	65	67	67	67	67	7I	68	7I	69	65	67
62	68	70	67	64	65	65	64	6I	66	67	6I	65
64	70	64	68	60	6I	68	65	60	67	65	63	65
65	63	64	66	62	65	65	68	6I	65	6I	64	62
68	69	70	7I	70	69	70	7I	65	7I	70	7I	69
70	64	7I	70	70	68	70	62	66	69	70	7I	69
72	73	74	73	70	63	67	65	63	68	70		

22. Результаты определения поверхностной плотности асбестового полотна, г/м^2

43I	470	43I	432	434	450	449	437	448	445	35I	393
370	26I	360	362	368	36I	369	4II	4I2	4I3	4II	430
429	425	424	427	402	429	4II	4I9	4I4	4I7	429	4I5
42I	420	4I9	429	427	424	430	420	42I	42I	429	4I7
4I5	4I4	4I3	4II	39I	392	398	400	4I0	409	406	400
399	397	396	409	408	4I0	400	405	407	406	400	403
404	405	4I0	4I0	405	40I	402	407	406	39I	392	399
405	407	407	402	37I	372	390	385	380	38I	382	383
380	375	375	374	380	379	379	372	374	377	376	37I
373	374	376	378	375	376	378	379	380	38I	382	383
383	383	37I	372	372	390	378	400	399	390	387	40I

23. Горизонтальное отклонение от цели при испытаниях I90 ракет, м

4,3	-29,3	20,5	27,3	-20,8	-28,7	26,4	-30,I	20,8	-27,3
II,2	9,5	-5,3	I9,2	5,2	-6,0	2,6	4,9	-0,8	0,2
7,5	I5,I	8,0	I7,9	I0,3	II,4	5,I	I4,8	I7,8	-8,3
2,5	-5,8	56,9	9,0	-5,9	I,2	I9,2	-22,4	I9,4	-I9,5
2I,3	I9,8	-32,2	48,I	-2I,I	-2I,3	-8,8	I0,2	-37,2	-0,3
I4,5	26,3	-I,9	26,3	-I,9	I2,4	I4,9	I8,2	I,5	I,6
I,7	-I0,5	I,7	2,7	I6,I	I,8	3,2	32,I	-50,8	6,9
5I,2	3I,3	-47,9	53,4	30,2	-56,I	I4,0	II,8	-7,5	I8,4
II,5	-5,0	-6,2	-II,2	I8,6	I6,7	-I2,3	I7,I	-I2,3	25,3
I,9	-I6,3	-54,3	-32,7	-I9,3	3,7	2,0	3,8	0,I	0
I3,5	0,3	6,8	46,2	42,3	-40,I	22,3	27,I	-23,0	2I,8
0	-2,5	0,8	-5,2	2,9	6,0	I8,8	-8,I	-20,0	-23,7
23,4	5,4	4,2	-9,0	23,8	4,4	-I8,3	I5,7	5,0	-3,2
I0,8	7,2	I2,8	I3,0	-7,3	7,8	I7,3	7,9	I3,9	I2,0
7,8	-I3,2	8,I	24,3	-I6,5	-I4,2	-I2,3	-I5,2	8,8	-6,8
I3,8	-20,8	I5,5	8,9	I5,3	8,7	-6,5	9,3	I8,8	-I7,7
I0,0	24,8	-8,I	I9,9	0	9,8	-I0,0	I6,9	25,8	-7,2
I6,5	-I4,8	-I3,5	-6,9	I2,4	-26,2	27,8	28,5	29,5	-27,3
29,8	30,0	-24,8	-46,3	-25,2	-34,5	38,3	-37,5	37,4	42,3

24. Распределение скорости автомобилей на одном из участков шоссе, км/ч

65	85	78	73	80	76	8I	70	80	80	77	90	75	69	77	87	78	84
79	75	79	67	80	95	83	68	72	76	83	89	76	84	79	85	74	86
79	74	78	8I	92	8I	66	8I	82	59	87	58	75	88	77	79	80	77
73	69	79	72	80	78	75	73	I0I	73	83	89	97	83	I03	73	94	79
74	9I	79	76	63	74	92	78	84	80	83	99	78	82	59	79	6I	78
94	92	79	85	82	84	68	76	7I	79	73							

25. Результаты определения разрывной нагрузки асбестовых нитей, сН

780	860	820	860	600	720	720	600	800	820
980	I020	600	760	I220	I060	I240	I020	860	740
660	600	580	780	500	800	680	600	760	II60
880	I040	960	800	760	980	840	840	700	I000
640	620	I000	I000	I040	740	640	860	840	I000
I040	820	920	900	880	840	700	II20	900	660
860	680	I080	920	780	700	660	640	580	640
720	720	580	840	840	920	940	900	500	980
760	620	580	I040	I080	840	920	900	660	I040
520	900	860	I060	980	900	860	980	I300	II60
880	780	580	880	900	880	900	720	640	660
820	930	680	500	780	9I0	700	760	780	660
740	300	760	780	860	780	560	560	900	700
740	740	I300	740	940	940	740	900	900	I220

Указание. В качестве гипотезы H_0 рассмотреть возможность применения статистических моделей: нормальное распределение, гамма-распределение и логарифмически нормальное распределение.

26. Результаты определения долговечности шерстяной пряжи при самоистирании в петле на приборе ИШ, число циклов

288	284	29I	268	265	280	382	290	335	353	440	353	400
366	338	3I5	384	367	328	388	348	360	409	3II	336	280
290	335	353	400	335	300	36I	360	325	345	349	307	344
323	360	397	379	334	399	352	349	36I	385	333	377	347
32I	359	449	356	343	39I	332	375	345	358	320	342	420
352	368	33I	373	357	339	3I9	309	34I	335	367	375	37I
292	356	3I7	340	329	334	366	389	332	354	3I3	328	425
295	355	345	339	334	365	379	349	40I	367	364	386	3I8
407	38I	337	289	366	369	384	347	405	360	344	336	305
350	369	403	346	362	326	346	340	385	4I9	35I	356	377

27. Результаты измерений геометрического размера изделий, мм

14,12	14,55	14,26	14,43	14,50	14,46	14,15	14,40	14,22	14,61
14,24	14,42	14,03	14,35	14,18	14,48	14,51	14,52	14,62	14,45
14,32	14,14	14,59	14,51	14,54	14,38	14,27	14,53	14,54	14,64
14,37	14,58	14,56	14,80	14,60	14,48	14,44	14,50	14,38	14,63
14,45	14,46	14,36	14,52	14,33	14,65	14,82	14,61	14,49	14,78
14,81	14,40	14,88	14,47	14,57	14,94	14,60	14,59	14,64	14,70
14,80	14,62	14,43	14,96	14,53	14,58	14,85	14,44	14,41	14,79
14,92	14,55	15,84	14,67	14,57	14,95	14,50	15,06	14,66	14,65
14,71	14,51	14,66	14,94	14,67	15,14	14,56	14,86	14,69	14,77
15,04	14,71	14,79	14,73	14,68	14,78	14,93	14,68	14,75	14,70

28. Результаты измерений боковой ошибки наводки при стрельбе с самолета по наземной цели, тыс. доли радиана

–0,5	–1,8	0,2	1,5	–2,4	–1,0	–1,8	–0,6	0,3	0,7	–0,5
1,8	–1,9	–0,9	–0,7	0,2	3,4	1,9	0,7	2,7	1,6	1,0
–0,2	–2,7	1,7	2,7	–0,2	–1,5	1,5	1,7	–1,9	–2,5	3,0
0,8	2,5	–1,8	–0,1	–1,7	–0,9	–0,9	3,3	–2,5	–2,9	–0,9
–1,9	–2,6	0,9	1,8	–2,0	–2,6	–0,8	0,2	0,4	1,9	2,0
–1,2	–1,4	–2,4	2,9	–1,6	–1,4	2,3	–1,7	–2,4	–2,4	–1,8
–2,3	–0,7	2,9	–3,8	1,6	–1,9	–1,4	–0,8	–1,5	2,8	–1,5
–2,2	–1,5	–3,4	1,9	–0,1	–0,6	–0,1	1,4	–0,9	–1,3	2,6
–1,6	–0,8	0,2	0,4	–1,7	1,9	0,2	1,7	0,3	1,5	–1,9
–0,7	1,5	1,7	0,7	1,8	–0,8	–0,9	–0,7	1,6	–0,9	–1,0
0,9	0,8	0,5	–0,7	0,3	0,7	–0,8	0,7	–0,6	–0,8	0,8
–0,5	–0,6	–0,5	–0,4	–0,9	1,5	1,8	–0,4	1,9	–0,3	–0,6
0,5	–0,3	–0,5	1,9	0,2	–0,4	–0,6	–0,8	0,7	–0,7	–0,9
–0,2	0,8	–0,6	1,2	0,3	1,8	–0,8	–0,6	–0,7	1,7	1,8
0,7	–0,7	0,6	–0,3	0,6	0,3	–0,2	0,3	–0,5	–0,4	0,5

29. Предел прочности образцов сварного шва, Н/мм^2

34,0	39,4	36,3	34,1	39,1	33,1	40,1	35,3	39,2	38,7	38,4
41,5	34,9	38,8	36,9	41,1	33,8	38,0	37,8	42,3	35,2	35,4
35,4	36,4	32,9	37,3	36,5	30,2	30,0	30,4	30,1	40,7	35,9
37,0	40,9	35,8	37,2	31,1	36,9	36,9	37,4	40,8	38,1	33,5
30,8	38,2	32,5	41,1	33,2	38,9	39,9	38,9	38,3	35,3	37,1
35,5	37,1	43,9	35,0	32,6	28,9	34,4	29,0	33,9	32,8	40,4
28,1	31,8	39,5	33,4	42,3	35,5	39,6	37,8	39,9	37,6	29,4
32,4	40,0	34,6	28,3	32,3	38,7	28,7	29,8	34,8	38,6	41,8
31,9	43,1	30,4	41,9	30,6	38,8	32,7	42,8	39,7	33,3	34,5
40,0	31,6	36,8	31,3	39,8	37,2					

30. Сформировать до объема $n = 120$ массив случайных чисел по формуле

$$\xi_{n-1} = Д(11\xi_n + \pi),$$

где $Д(y)$ – дробная часть числа $y = 11\xi_n + \pi$. Значение начального случайного числа ξ_0 принять равным 0,002. Вычисление случайных чисел провести до третьего знака после запятой, округление случайных чисел не производить.

Литература

1. Вентцель Е.С. Теория вероятностей. М.: Наука, 1970.
2. Карасев А.И. Теория вероятностей и математическая статистика. М.: Статистика, 1979.
3. Сборник задач по математике для втузов: В 4 т. М.: Наука, 1990. Т. 3: Теория вероятностей и математическая статистика / Под ред. А.В.Ефимова.
4. Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1972.
5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. М.: Высшая школа, 1975.
6. Крамер Г. Математические методы статистики / Пер. с англ. М.: Мир, 1975.
7. Ван-дер-Варден Б.Л. Математическая статистика / Пер. с англ. М.: Изд-во иностр. лит., 1960.
8. Пугачев В.С. Теория вероятностей и математическая статистика. М.: Наука, 1979.

9. Селиверстов В.В., Тищенко А.А. Практика математической обработки результатов эксперимента в дипломных и курсовых работах. М.: Изд-во УДН, I98I.

I0. Сборник задач по теории вероятностей, математической статистике, теории случайных функций / Под ред. А.В.Свешникова. М.: Наука, I965.

II. Кобляков А.И., Селиверстов В.В. О вероятностных моделях для текстильных материалов при многократных механических воздействиях. Изв. высш. учеб. заведений. Сер. Технология текстильной промышленности. I990. № 4. С. 5–7.

I2. Ярослав Янко. Математико-статистические таблицы. М.: Госстатиздат, I96I.

ПРИЛОЖЕНИЕ

Таблица III

Значения функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

x	0,00	0,0I	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
I	2	3	4	5	6	7	8	9	I0	II
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,I	3970	3965	396I	3956	395I	3945	3939	3932	3925	39I8
0,2	39I0	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	38I4	3802	3790	3778	3765	3752	3739	3726	37I2	3697
0,4	3683	3668	3653	3637	362I	3605	3589	3572	3555	3538
0,5	352I	3503	3485	3467	3448	3429	34I0	339I	3372	3352
0,6	3332	33I2	3292	327I	325I	3230	3209	3I87	3I66	3I44
0,7	3I23	3I0I	3079	3056	3034	30II	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	266I	2637	26I3	2589	2565	254I	25I6	2492	2468	2444
I,0	0,2420	2396	237I	2347	2323	2299	2275	225I	2227	2203
I,I	2I79	2I55	2I3I	2I07	2083	2059	2036	20I2	I989	I965
I,2	I942	I9I9	I895	I872	I849	I826	I804	I78I	I758	I736
I,3	I7I4	I69I	I669	I647	I626	I604	I582	I56I	I539	I5I8
I,4	I497	I476	I456	I435	I4I5	I394	I374	I354	I334	I3I5
I,5	I295	I276	I257	I238	I2I9	I200	II82	II63	II45	II27

ПРИЛОЖЕНИЯ

Таблица II

Значения функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

x	0,00	0,0I	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
I	2	3	4	5	6	7	8	9	I0	II
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,I	3970	3965	396I	3956	395I	3945	3939	3932	3925	39I8
0,2	39I0	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	38I4	3802	3790	3778	3765	3752	3739	3726	37I2	3697
0,4	3683	3668	3653	3637	362I	3605	3589	3572	3555	3538
0,5	352I	3503	3485	3467	3448	3429	34I0	339I	3372	3352
0,6	3332	33I2	3292	327I	325I	3230	3209	3I87	3I66	3I44
0,7	3I23	3I0I	3079	3056	3034	30II	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	266I	2637	26I3	2589	2565	254I	25I6	2492	2468	2444
I,0	0,2420	2396	237I	2347	2323	2299	2275	225I	2227	2203
I,I	2I79	2I55	2I3I	2I07	2083	2059	2036	20I2	I989	I965
I,2	I942	I9I9	I895	I872	I849	I826	I804	I78I	I758	I736
I,3	I7I4	I69I	I669	I647	I626	I604	I582	I56I	I539	I5I8
I,4	I497	I476	I456	I435	I4I5	I394	I374	I354	I334	I3I5
I,5	I295	I276	I257	I238	I2I9	I200	II82	II63	II45	II27

Продолжение табл. III

I	2	3	4	5	6	7	8	9	I0	II
I,6	II09	I092	I074	I057	I040	I023	I006	0989	0973	0957
I,7	0940	0925	0909	0893	0878	0863	0848	0833	08I8	0804
I,8	0790	0775	076I	0748	0734	072I	0707	0694	068I	0669
I,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	055I
2,0	0,0540	0529	05I9	0508	0498	0488	0478	0468	0459	0449
2,I	0440	043I	0422	04I3	0404	0396	0387	0379	037I	0363
2,2	0355	0347	0339	0332	0325	03I7	03I0	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	024I	0235	0229
2,4	0224	02I9	02I3	0208	0203	0I98	0I94	0I89	0I84	0I80
2,5	0I75	0I7I	0I67	0I63	0I58	0I54	0I5I	0I47	0I43	0I39
2,6	0I36	0I32	0I29	0I26	0I22	0II9	0II6	0II3	0II0	0I07
2,7	0I04	0I0I	0099	0096	0093	009I	0088	0086	0084	008I
2,8	0079	0077	0075	0073	007I	0069	0067	0065	0063	006I
2,9	0060	0058	0056	0055	0053	005I	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,I	0033	0032	003I	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	002I	0020	0020	00I9	00I8	00I8
3,3	00I7	00I7	00I6	00I6	00I5	00I5	00I4	00I4	00I3	00I3
3,4	00I2	00I2	00I2	00II	00II	00I0	00I0	00I0	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006

Окончание табл. III

I	2	3	4	5	6	7	8	9	I0	II
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	000I	000I

Таблица III

Функция распределения нормального закона

$N(0,1)$; $\Phi(-x) = 1 - \Phi(x)$, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$

x	0,00	0,0I	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
I	2	3	4	5	6	7	8	9	I0	II
0,0	0,5000	0,5040	0,5080	0,5I20	0,5I60	0,5I99	0,5239	0,5279	0,53I9	0,5359
0,I	0,5398	0,5438	0,5478	0,55I7	0,5557	0,5596	0,5636	0,5675	0,57I4	0,5753
0,2	0,5793	0,5832	0,587I	0,59I0	0,5948	0,5987	0,6026	0,6064	0,6I03	0,6I4I
0,3	0,6I79	0,62I7	0,6255	0,6293	0,633I	0,6368	0,6406	0,6443	0,6480	0,65I7
0,4	0,6554	0,659I	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,69I5	0,6950	0,6985	0,70I9	0,7054	0,7088	0,7I23	0,7I57	0,7I90	0,7224
0,6	0,7257	0,729I	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,75I7	0,7549
0,7	0,7580	0,76II	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852

Продолжение табл. III

I	2	3	4	5	6	7	8	9	I0	II
0,8	0,788I	0,79I0	0,7939	0,7967	0,7995	0,8023	0,805I	0,8078	0,8I06	0,8I33
0,9	0,8I59	0,8I86	0,82I2	0,8238	0,8264	0,8289	0,83I5	0,8340	0,8365	0,8389
I,0	0,84I3	0,8438	0,846I	0,8485	0,8508	0,853I	0,8554	0,8577	0,8599	0,862I
I,I	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,88I0	0,8830
I,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,90I5
I,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9II5	0,9I3I	0,9I47	0,9I62	0,9I77
I,4	0,9I92	0,9207	0,9222	0,9236	0,925I	0,9265	0,9279	0,9292	0,9306	0,93I9
I,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9294	0,9406	0,94I8	0,9429	0,944I
I,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,95I5	0,9525	0,9535	0,9545
I,7	0,9554	0,9564	0,9573	0,9582	0,959I	0,9599	0,9608	0,96I6	0,9625	0,9633
I,8	0,964I	0,9649	0,9656	0,9664	0,967I	0,9678	0,9686	0,9693	0,9699	0,9706
I,9	0,97I3	0,97I9	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,976I	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,98I2	0,98I7
2,I	0,982I	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,986I	0,9864	0,9868	0,987I	0,9875	0,9878	0,988I	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,990I	0,9904	0,9906	0,9909	0,99II	0,99I3	0,99I6
2,4	0,99I8	0,9920	0,9922	0,9925	0,9927	0,9929	0,993I	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,994I	0,9943	0,9945	0,9946	0,9948	0,9949	0,995I	0,9952
2,6	0,9953	0,9955	0,9957	0,9959	0,9960	0,996I	0,9962	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,997I	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,998I
2,9	0,998I	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Окончание табл. III

I	2	3	4	5	6	7	8	9	I0	II
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,I	0,9990	0,999I	0,999I	0,999I	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Квантили u_p нормального распределения $N(0,1)$; $u_p = -u_{1-p}$:

p	0,90	0,95	0,975	0,99	0,995	0,999	0,9995
u_p	I,282	I,645	I,960	2,326	2,576	3,090	3,29I

Значения гамма-функции $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\,dt$

x	$\Gamma(x)$	x	$\Gamma(x)$	x	$\Gamma(x)$	x	$\Gamma(x)$
1,00	1,00000	1,25	0,90640	1,50	0,88623	1,75	0,91906
01	0,99443	26	0,90440	51	0,88569	76	0,92137
02	0,98884	27	0,90250	52	0,88704	77	0,92376
03	0,98355	28	0,90072	53	0,88757	78	0,92623
04	0,97844	29	0,89904	54	0,88818	79	0,92877
1,05	0,97350	1,30	0,89747	1,55	0,88887	1,80	0,93138
06	0,96874	31	0,89600	56	0,88964	81	0,93408
07	0,96415	32	0,89464	57	0,89049	82	0,93685
08	0,95973	33	0,89338	58	0,89142	83	0,93969
09	0,95546	34	0,89222	59	0,98243	84	0,94261
1,10	0,95135	1,35	0,89115	1,60	0,89352	1,85	0,94561
11	0,94740	36	0,89018	61	0,89468	86	0,94869
12	0,94359	37	0,88931	62	0,89592	87	0,95181
13	0,93993	38	0,88854	63	0,89724	88	0,95507
14	0,93642	39	0,88785	64	0,89864	89	0,95838
1,15	0,93304	1,40	0,88726	1,65	0,90012	1,90	0,96177
16	0,92980	41	0,88676	66	0,90167	91	0,96523
17	0,92670	42	0,88636	67	0,90330	92	0,96877
18	0,92373	43	0,88604	68	0,90500	93	0,97240
19	0,92089	44	0,88581	69	0,90678	94	0,97610
1,20	0,91817	1,46	0,88566	1,70	0,90864	1,95	0,97988
21	0,91558	46	0,88560	71	0,91057	96	0,98374
22	0,91311	47	0,88563	72	0,91258	97	0,98768
23	0,91075	48	0,88575	73	0,91467	98	0,99171
24	0,90852	49	0,88595	74	0,91683	99	0,99581
1,25	0,90640	1,50	0,88623	1,75	0,91906	2,00	1,00000

Значения гамма-функции для $x<1$ и для $x>2$ могут быть вычислены с помощью формул: $\Gamma(x)=\frac{\Gamma(x+1)}{x}$, $\Gamma(x)=(x-1)\Gamma(x-1)$.

Квантили "хи-квадрат" распределения $\chi^2_p(k)$

k	p: 0,010	0,025	0,05	0,10	0,20	0,30	0,70	0,80	0,90	0,95	0,975	0,990	0,995
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	$0,0^3157$	$0,0^3982$	$0,0^2393$	0,0158	0,0642	0,148	1,07	1,64	2,71	3,84	5,02	6,63	7,88
2	0,0201	0,0506	0,103	0,211	0,446	0,713	2,41	3,22	4,61	5,99	7,38	9,21	10,6
3	0,115	0,216	0,352	0,584	1,00	1,42	3,67	4,64	6,25	7,81	9,35	11,3	12,8
4	0,297	0,484	0,711	1,06	1,65	2,19	4,88	5,99	7,78	9,49	11,1	13,3	14,9
5	0,554	0,831	1,15	1,61	2,34	3,00	6,06	7,29	9,24	11,1	12,8	15,1	16,7
6	0,872	1,24	1,64	2,20	3,07	3,83	7,23	8,56	10,6	12,6	14,4	16,8	18,5
7	1,24	1,69	2,17	2,83	3,82	4,67	8,38	9,80	12,0	14,1	16,0	18,5	20,3
8	1,65	2,18	2,73	3,49	4,59	5,53	9,52	11,0	13,4	15,5	17,5	20,1	22,0
9	2,09	2,70	3,33	4,17	5,38	6,39	10,7	12,2	14,7	16,9	19,0	21,7	23,6
10	2,56	3,25	3,94	4,87	6,18	7,27	11,8	13,4	16,0	18,3	20,5	23,2	25,2
11	3,05	3,82	4,57	5,58	6,99	8,15	12,9	14,6	17,3	19,7	21,9	24,7	26,8
12	3,57	4,40	5,23	6,30	7,81	9,03	14,0	15,8	18,5	21,0	23,3	26,2	28,3
13	4,11	5,01	5,89	7,04	8,63	9,93	15,1	17,0	19,8	22,4	24,7	27,7	29,8
14	4,66	5,63	6,57	7,79	9,47	10,8	16,2	18,2	21,1	23,7	26,1	29,1	31,3
15	5,23	6,26	7,26	8,55	10,3	11,7	17,3	19,3	22,3	25,0	27,5	30,6	32,8

Квантили "хи-квадрат" распределения $\chi^2_p(k)$

k	p												
	0.0I0	0,025	0,05	0,I0	0,20	0,30	0,70	0,80	0,90	0,95	0,975	0,990	0,995
I	2	3	4	5	6	7	8	9	I0	II	I2	I3	I4
I	$0.0^{3}157$	$0.0^{3}982$	$0.0^{2}393$	0.0I58	0,0642	0,I48	I,07	I,64	2,7I	3,84	5,02	6,63	7.88
2	0.020I	0,0506	0.I03	0,2II	0,446	0,7I3	2,4I	3,22	4,6I	5,99	7,38	9,2I	I0,6
3	0.II5	0,2I6	0,352	0,584	I,00	I,42	3,67	4,64	6,25	7,8I	9,35	II,3	I2,8
4	0,297	0,484	0,7II	I,06	I,65	2,I9	4,88	5,99	7,78	9,49	II,I	I3,3	I4,9
5	0,554	0,83I	I,I5	I,6I	2,34	3,00	6,06	7,29	9,24	II,I	I2,8	I5,I	I6,7
6	0,872	I,24	I,64	2,20	3,07	3,83	7,23	8,56	I0,6	I2,6	I4,4	I6,8	I8,5
7	I,24	I,69	2,I7	2,83	3,82	4,67	8,38	9,80	I2,0	I4,I	I6,0	I8,5	20,3
8	I,65	2,I8	2,73	3,49	4,59	5,53	9,52	II,0	I3,4	I5,5	I7,5	20,I	22,0
9	2,09	2,70	3,33	4,I7	5,38	6,39	I0,7	I2,2	I4,7	I6,9	I9,0	2I,7	23,6
I0	2,56	3,25	3,94	4,87	6,I8	7,27	II,8	I3,4	I6,0	I8,3	20,5	23,2	25,2
II	3,05	3,82	4,57	5,58	6,99	8,I5	I2,9	I4,6	I7,3	I9,7	2I,9	24,7	26,8
I2	3,57	4,40	5,23	6,30	7,8I	9,03	I4,0	I5,8	I8,5	2I,0	23,3	26,2	28,3
I3	4,II	5,0I	5,89	7,04	8,63	9,93	I5,I	I7,0	I9,8	22,4	24,7	27,7	29,8
I4	4,66	5,63	6,57	7,79	9,47	I0,8	I6,2	I8,2	2I,I	23,7	26,I	29,I	3I,3
I5	5,23	6,26	7,26	8,55	I0,3	II,7	I7,3	I9,3	22,3	25,0	27,5	30,6	32,8

Окончание табл. ПIУ

I	2	3	4	5	6	7	8	9	I0	II	I2	I3	I4
I6	5,8I	6,9I	7,96	9,3I	II,2	I2,6	I8,4	20,5	23,5	26,3	28,8	32,0	34,3
I7	6,4I	7,56	8,67	I0,I	I2,0	I3,5	I9,5	2I,6	24,8	27,6	30,2	33,4	35,7
I8	7,0I	8,23	9,39	I0,9	I2,9	I4,4	20,6	22,8	26,0	28,9	3I,5	34,8	37,2
I9	7,63	8,9I	I0,I	II,7	I3,7	I5,4	2I,7	23,9	27,2	30,I	32,9	36,2	38,6
20	8,26	9,59	I0,9	I2,4	I4,6	I6,3	22,8	25,0	28,4	3I,4	34,2	37,6	40,0
2I	8,90	I0,3	II,6	I3,2	I5,4	I7,2	23,9	26,9	29,6	32,7	35,5	38,9	4I,4
22	9,54	II,0	I2,3	I4,0	I6,3	I8,I	24,9	27,3	30,8	33,9	36,8	40,3	42,8
23	I0,2	II,7	I3,I	I4,8	I7,2	I9,0	26,0	28,4	32,0	35,2	38,I	4I,6	44,2
24	I0,9	I2,4	I3,8	I5,7	I8,I	I9,9	27,I	29,6	33,2	36,4	39,4	43,0	45,6
25	II,5	I3,I	I4,6	I6,5	I8,9	20,9	28,2	30,7	34,4	37,7	40,6	44,3	46,9
26	I2,2	I3,8	I5,4	I7,3	I9,8	2I,8	29,2	3I,8	35,6	38,9	4I,9	45,6	48,3
27	I2,9	I4,6	I6,2	I8,I	20,7	22,7	30,3	32,9	36,7	40,I	43,2	47,0	49,6
28	I3,6	I5,3	I6,9	I8,9	2I,6	23,6	3I,4	34,0	37,9	4I,3	44,5	48,3	5I,0
29	I4,3	I6,0	I7,7	I9,8	22,5	24,6	32,5	35,I	39,I	42,6	45,7	49,6	52,3
30	I5,0	I6,8	I8,5	20,6	23,4	25,5	33,5	36,3	40,3	43,8	47,0	50,9	53,7
35	I8,5	20,6	22,5	24,8	27,8	30,2	38,9	4I,8	46,I	49,8	53,2	57,3	60,3
40	22,2	24,4	26,5	29,I	32,3	34,9	44,2	47,3	5I,8	55,8	59,3	63,7	66,8
45	25,9	28,4	30,6	33,4	36,9	39,6	49,5	52,7	57,5	6I,7	65,4	70,0	73,2
50	29,7	32,4	34,8	37,7	4I,4	44,3	54,7	58,2	63,2	67,5	7I,4	76,2	79,5
75	49,5	52,9	56,I	59,8	64,5	68,I	80,9	85,I	9I,I	96,2	I00,8	I06,4	II0,3
I00	70,I	74,2	77,9	82,4	87,9	92,I	I06,9	III,7	II8,5	I24.3	I29.6	I35.6	I40.2

Окончание табл. ПШ

1	2	3	4	5	6	7	8	9	10	11	12	13	14
16	5,81	6,91	7,96	9,31	11,2	12,6	18,4	20,5	23,5	26,3	28,8	32,0	34,3
17	6,41	7,56	8,67	10,1	12,0	13,5	19,5	21,6	24,8	27,6	30,2	33,4	35,7
18	7,01	8,23	9,39	10,9	12,9	14,4	20,6	22,8	26,0	28,9	31,5	34,8	37,2
19	7,63	8,91	10,1	11,7	13,7	15,4	21,7	23,9	27,2	30,1	32,9	36,2	38,6
20	8,26	9,59	10,9	12,4	14,6	16,3	22,8	25,0	28,4	31,4	34,2	37,6	40,0
21	8,90	10,3	11,6	13,2	15,4	17,2	23,9	26,9	29,6	32,7	35,5	38,9	41,4
22	9,54	11,0	12,3	14,0	16,3	18,1	24,9	27,3	30,8	33,9	36,8	40,3	42,8
23	10,2	11,7	13,1	14,8	17,2	19,0	26,0	28,4	32,0	35,2	38,1	41,6	44,2
24	10,9	12,4	13,8	15,7	18,1	19,9	27,1	29,6	33,2	36,4	39,4	43,0	45,6
25	11,5	13,1	14,6	16,5	18,9	20,9	28,2	30,7	34,4	37,7	40,6	44,3	46,9
26	12,2	13,8	15,4	17,3	19,8	21,8	29,2	31,8	35,6	38,9	41,9	45,6	48,3
27	12,9	14,6	16,2	18,1	20,7	22,7	30,3	32,9	36,7	40,1	43,2	47,0	49,6
28	13,6	15,3	16,9	18,9	21,6	23,6	31,4	34,0	37,9	41,3	44,5	48,3	51,0
29	14,3	16,0	17,7	19,8	22,5	24,6	32,5	35,1	39,1	42,6	45,7	49,6	52,3
30	15,0	16,8	18,5	20,6	23,4	25,5	33,5	36,3	40,3	43,8	47,0	50,9	53,7
35	18,5	20,6	22,5	24,8	27,8	30,2	38,9	41,8	46,1	49,8	53,2	57,3	60,3
40	22,2	24,4	26,5	29,1	32,3	34,9	44,2	47,3	51,8	55,8	59,3	63,7	66,8
45	25,9	28,4	30,6	33,4	36,9	39,6	49,5	52,7	57,5	61,7	65,4	70,0	73,2
50	29,7	32,4	34,8	37,7	41,4	44,3	54,7	58,2	63,2	67,5	71,4	76,2	79,5
75	49,5	52,9	56,1	59,8	64,5	68,1	80,9	85,1	91,1	96,2	100,8	106,4	110,3
100	70,1	74,2	77,9	82,4	87,9	92,1	106,9	111,7	118,5	124,3	129,6	135,6	140,2

Таблица ПУ

Квантили распределения Стьюдента

k	p						
	0,750	0,900	0,950	0,975	0,990	0,995	0,999
1	1,000	3,078	6,314	12,706	31,821	61,657	318,00
2	0,816	1,886	2,920	4,303	6,965	9,925	22,30
3	0,765	1,638	2,353	3,182	4,541	5,841	10,20
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408
29	0,683	1,311	1,699	2,045	2,462	2,756	3,393
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160
∞	0,674	1,282	1,645	1,960	2,326	2,576	3,090

ОГЛАВЛЕНИЕ

Редакция заказной литературы

Виктор Владимирович Селиверстов
Виктор Павлович Обухов

Статистика эксперимента в технике и науке.
Обработка выборки

Редактор Н.Г.Ковалевская
Корректор Л.И.Малютина

Подписано в печать 09.03.93. Формат 60х84/I6. Бумага тип. № 2.
Печ.л. 6,5. Усл.печ.л. 6,05. Уч-изд.л. 5,2. Тираж 800 экз.
Изд. № 48. Заказ № 577 С-295

Издательство МГТУ, типография МГТУ.
I07005, Москва, Б-5, 2-я Бауманская, 5.