

Министерство образования Республики Беларусь
БЕЛОРУССКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Кафедра «Электроснабжение»

В.Б.Козловская

МАТЕМАТИЧЕСКИЕ ЗАДАЧИ ЭНЕРГЕТИКИ

Учебно-методическое пособие к практическим занятиям
и курсовому проектированию для студентов
специальности 1 – 43 01 03 «Электроснабжение»
специализации 1– 43 01 03 01
«Электроснабжение промышленных предприятий»

Минск 2005

УДК 628.97 (075.8)

ББК 31.294 я7

К 59

Рецензенты:

В.А.Булат, В.М.Цыганков

Козловская В.Б.

К 59 Математические задачи энергетики: Учебно-метод. пособие к практическим занятиям и курсовому проектированию / В.Б.Козловская— Мн.: БНТУ, 2005. - с.

В пособии изложены основные положения теории вероятностей, математической статистики, необходимые для решения наиболее распространенных практических задач в области электроснабжения промышленных предприятий.

Предлагаемый материал иллюстрируется примерами. В приложении приводятся справочные материалы, необходимые для решения предлагаемых задач..

Издание предназначено для студентов специальности 43 01 03 «Электроснабжение» специализации 43 01 03 01 «Электроснабжение промышленных предприятий», а также может быть полезно студентам других специальностей, изучающих данную дисциплину.

УДК 628.97 (075.8)

ББК 31.294 я7

ISBN

© Козловская В.Б., 2005

СОДЕРЖАНИЕ

Введение.....	4
.....
1. Основные положения математической статистики.....	5
2. Определение точечных оценок числовых характеристик случайной величины.....	10
3. Предварительная обработка исходных данных.....	13
3.1. «Очистка» исходных данных по правилу трех сигм.....	14
3.2. «Очистка» исходных данных с применением гистограмм частот.....	15
4. Выравнивание статистических распределений.....	18
5. Проверка статистических гипотез.....	24
5.1. Проверка гипотезы о равенстве математических ожиданий двух случайных величин, имеющих нормальное распределение.....	26
5.2. Проверка гипотезы о равенстве дисперсий двух случайных величин, имеющих нормальное распределение.....	28
5.3. Проверка гипотезы о нормальном распределении генеральной совокупности.....	30
6. Определение интервальных оценок математического ожидания случайной величины.....	33
ПРИЛОЖЕНИЯ.....	39
Список использованных источников.....	45

ВВЕДЕНИЕ

В практической деятельности при проектировании, строительстве и эксплуатации предприятий для выполнения поставленной задачи инженеру постоянно приходится принимать решения. При проектировании это может быть выбор схемы или какого-то параметра системы электроснабжения (сечение проводов, мощность трансформатора и т.п.); при строительстве и монтаже – план и технология проведения отдельных работ и всего их комплекса; в эксплуатации это могут быть задачи выбора уставок защит, определение точности приборов, параметров системы регулирования напряжения и т.д. При этом принятое решение будет тем более оптимальным, чем точнее и полнее учтен характер рассматриваемых процессов и чем большая информация будет получена о них.

Явления и процессы, происходящие в электроэнергетических системах обычно не являются детерминированными, т.е. с точным результатом, а как правило, носят случайный характер, т.к. они формируются под воздействием большого количества разнообразных факторов. Однако при неоднократном повторении одного и того же опыта проявляются закономерности, которые можно изучить и использовать при исследовании.

Установление закономерностей, которым подчинены массовые случайные явления, основано на изучении методами теории вероятностей статистических данных – результатов наблюдений (измерений). Этими вопросами занимается математическая статистика, основная задача которой состоит в создании методов сбора и обработки статистических данных для анализа происходящих процессов и получения научных и практических выводов.

1. ОСНОВНЫЕ ПОЛОЖЕНИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Проведение исследований в какой-либо области знаний, как правило, начинается со сбора исходной информации о предмете исследования – статистических данных, которые в дальнейшем подлежат обработке с использованием методов теории вероятностей и математической статистики для углубленного анализа и получения обоснованных выводов о предмете исследования. Математическая статистика решает две основных задачи: во-первых, необходимо указать способ сбора и группировки статистических данных, полученных в результате наблюдений, и, во-вторых, разработать метод их анализа в зависимости от целей исследования. При решении последней задачи осуществляется проверка статистических гипотез о виде неизвестного распределения и величине параметров распределения, вид которого известен.

Для описания и анализа полученных статистических данных, как правило, применяется выборочный метод, сущность которого заключается в следующем. Пусть требуется изучить совокупность объектов относительно некоторого качественного или количественного признака. Например, если имеется партия ламп накаливания, то качественным признаком может служить стандартность лампы, а количественным — её срок службы. Можно произвести сплошное обследование каждой лампы, но, во-первых, это сопряжено со значительными затратами, а во-вторых, проверка срока службы всех ламп накаливания свела бы к нулю полезный выпуск этой продукции. Поэтому обычно из всей совокупности объектов случайно отбирают некоторое их количество и подвергают изучению.

Выборочной совокупностью (выборкой) называют совокупность случайно отобранных объектов (результатов измерений). Объём выборки обозначается n .

Генеральной совокупностью называют совокупность объектов, из которых производится выборка. Объём генеральной совокупности обозначается N .

Рассмотрим основные понятия математической статистики.

Случайной величиной (СВ) считается величина, которая в результате опыта со случайным исходом принимает то или иное значение. Случайными величинами являются результаты измерений напряжения, активной и реактивной мощности, частоты, коэффициента мощности, количества повреждённого оборудования, погрешностей измерений и т.п. Случайные величины бывают дискретными и недискретными (непрерывными).

Дискретной называется СВ, множество возможных значений которой конечно или счётно.

Непрерывной называется СВ, множество значений которой нечётно и возможные значения СВ заполняют целиком некоторый интервал.

Для визуальной оценки распределения исследуемой СВ X производят группировку исходных данных [2]. Если изучается дискретная СВ, то наблюдаемые значения располагаются в порядке возрастания и подсчитываются частоты m_i или относительные частоты $\frac{m_i}{n}$ появления одинаковых значений СВ X . В результате получается статистический ряд распределения СВ:

Таблица 1

x_i – исход эксперимента	x_1	x_2	...	x_r
m_i	m_1	m_2	...	m_r
$\frac{m_i}{n}$	$\frac{m_1}{n}$	$\frac{m_2}{n}$...	$\frac{m_r}{n}$

При этом должны выполняться соотношения:

$$\sum_{i=1}^r m_i = n; \sum_{i=1}^r \frac{m_i}{n} = 1. \quad (1)$$

Если изучается непрерывная СВ X , то группировка заключается в разбиении интервала наблюдаемых значений СВ на r разрядов обычно равной длины

$h : [x_1; x_2[; [x_2; x_3[; \dots [x_{r-1}; x_r[$ и подсчёте частоты или относительной частоты попадания наблюдаемых значений СВ в соответствующие разряды.

Количество разрядов выбирается произвольно, но обычно не менее 5 и не более 15. При этом необходимо, чтобы в каждый разряд (кроме крайних) попало не менее 5 значений СВ. В результате получают статистический ряд распределения СВ X вида :

Таблица 2

Разряды, X	$[x_1; x_2[$	$[x_2; x_3[$...	$[x_{r-1}; x_r[$
m_i	m_1	m_2	...	m_r
$\frac{m_i}{n}$	$\frac{m_1}{n}$	$\frac{m_2}{n}$...	$\frac{m_r}{n}$

В этом случае также должны выполняться равенства (1).

Относительная частота

$$\frac{m_i}{n} = P_i^* \quad (1)$$

является приближённым значением вероятности P_i и при увеличении числа наблюдений (объёма выборки n) стремится к его значению.

Теория вероятности оперирует понятием **функции распределения** – это функция $F(x)$, определяющая вероятность того, что СВ X в результате эксперимента примет значение, меньше некоторого фиксированного x :

$$F(x) = F(X < x) = P(X \in]-\infty; x]). \quad (3)$$

Свойства $F(x)$:

- $\lim_{x \rightarrow -\infty} F(x) = F(-\infty) = 0, \quad \lim_{x \rightarrow +\infty} F(x) = F(+\infty) = 1;$
- $F(x)$ – неубывающая функция, т.е. $F(x_2) \geq F(x_1)$ при $x_2 > x_1$;
- $\lim_{x \rightarrow x_0 - 0} F(x) = F(x_0)$ – т.е. непрерывна слева.

С помощью функции распределения можно определить вероятность попадания СВ X в некоторый интервал [4]:

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = F(b) - F(a) . \quad (4)$$

Функция распределения строится как для дискретной, так и для непрерывной случайной величины. Это универсальный вид закона распределения.

Аналогом (приближённым значением) функции распределения в математической статистике является *эмпирическая (статистическая) функция распределения* – это функция $F^*(x)$, определяющая относительную частоту события, когда СВ X не превышает заданного x :

$$F^*(x) = P^*(X < x), \quad (5)$$

$$F^*(x) = \frac{n_x}{n}, \quad (6)$$

где n_x – число значений СВ X , меньших x ; n – объём выборки.

Эмпирическая функция распределения $F^*(x)$ используется в качестве оценки функции распределения $F(x)$ и при достаточно большом объёме выборки мало отличается от неё. Функция $F^*(x)$ обладает теми же свойствами, что и $F(x)$.

Графически эмпирическая функция распределения представляет собой ступенчатую разрывную функцию (рис.1), непрерывную слева; $F(x) = 0$ левее наименьшего значения x_1 и $F(x) = 1$ правее наибольшего значения x_n .

Пример 1. Построить эмпирическую функцию распределения по статистическому распределению СВ отклонения напряжения U , В, при оценке погрешности вольтметра:

Наблюдаемое значение СВ U, B	2	3	5
P_i^*	0,75	0,2	0,05

Решение. Относительная частота события ($U < u$) равна $F^*(u)$, следовательно:

$$F^*(u) = \begin{cases} 0, & \text{при } u \leq 2; \\ 0,75, & \text{при } 2 < u \leq 3; \\ 0,95, & \text{при } 3 < u \leq 5; \\ 1, & \text{при } u > 5. \end{cases}$$

График этой эмпирической функции распределения:

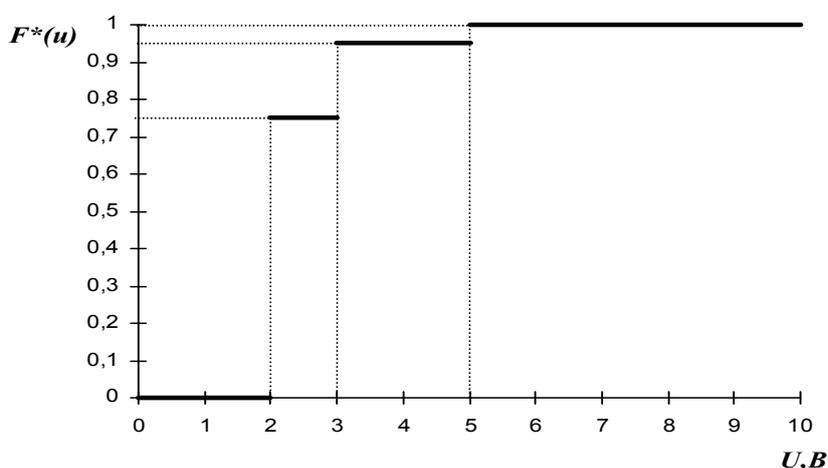


Рис. 1

При решении многих практических задач нет необходимости всегда строить математическую модель вероятностного эксперимента. Иногда достаточно иметь о случайной величине общее представление, знать характерные черты математической модели. Существенные особенности математической модели можно в сжатой форме выразить с помощью числовых характеристик. Основными из них являются математическое ожидание $M(x)$, дисперсия $D(x)$ и среднеквадратическое отклонение $\sigma(x)$. В математической статистике чаще используют не сами эти характеристики, а их приближённые значения – оценки.

2. ОПРЕДЕЛЕНИЕ ТОЧЕЧНЫХ ОЦЕНОК ЧИСЛОВЫХ ХАРАКТЕРИСТИК СЛУЧАЙНОЙ ВЕЛИЧИНЫ

На практике нередко приходится иметь дело со статистическим материалом небольшого объема, и в этом случае приходится вместо истинных значений числовых характеристик СВ использовать их приближенные значения – оценки.

Статистической оценкой неизвестного параметра Θ теоретического распределения называют функцию наблюдаемых случайных величин. Оценку обозначают $\hat{\Theta}$.

Оценки подразделяются на точечные и интервальные.

Точечная оценка параметра Θ определяется одним числом $\hat{\Theta} = u(x_1, x_2, \dots, x_n)$. Однако, не каждую функцию $u(x_1, x_2, \dots, x_n)$ можно использовать в качестве оценочной функции неизвестного параметра Θ , а только определенные классы функций, близкие к оцениваемому параметру Θ . Имеется специальный раздел математической статистики – теория оценивания, которая занимается выработкой правил конструирования функций $\hat{\Theta} = u(x_1, x_2, \dots, x_n)$ для нахождения точечных оценок неизвестных параметров.

Для того чтобы статистические оценки давали “хорошие” приближения оцениваемых параметров, они должны удовлетворять определенным требованиям. Во-первых, оценки должны быть несмещёнными, т.е. при использовании оценки $\hat{\Theta}$ вместо параметра Θ не делается систематической ошибки в сторону завышения или занижения:

$$M(\hat{\Theta}) = \Theta. \quad (7)$$

Однако несмещенная оценка не всегда дает хорошее приближение оцениваемого параметра. Возможные значения $\hat{\Theta}$ могут быть сильно рассеяны вокруг своего среднего значения, т.е. дисперсия оценки может быть значительной. Поэтому к статистической оценке предъявляется требование эффективности, т.е. выбранная несмещенная оценка должна обладать минимальной дисперсией:

$$D(\Theta) = \min . \quad (8)$$

При рассмотрении выборок большого объема к статистической оценке предъявляется также требование состоятельности, т.е. при увеличении объема выборки n оценка должна приближаться (сходится по вероятности) к искомому параметру:

$$\Theta \xrightarrow[n \rightarrow \infty]{p} \Theta . \quad (9)$$

На практике не всегда удается удовлетворить всем этим требованиям, но при выборе оценки любого параметра желательно ее критическое рассмотрение со всех этих точек зрения.

Для определения точечных оценок, существуют различные методы, дающие возможность получить оценки, в разной степени удовлетворяющие перечисленным требованиям. Точечные оценки основных числовых характеристик СВ наиболее просто могут быть рассчитаны методом моментов [2]. Если для изучения генеральных совокупностей СВ X извлечена выборка объемом n , то на основании неё определяются выборочные оценки числовых характеристик СВХ.

Математическое ожидание является характеристикой положения СВ и описывает ее центр распределения:

$$M(x) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^r m_i \cdot x_i = a . \quad (10)$$

Дисперсия и среднеквадратическое отклонение характеризуют разброс СВ около её математического ожидания. Их выборочные оценки рассчитываются по формулам:

$$\bar{D}(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^r m_i \cdot (x_i - \bar{x})^2 ; \quad (11)$$

$$\tilde{\sigma}(x) = \sqrt{\bar{D}(x)} . \quad (12)$$

Дисперсия и среднее квадратическое отклонение, полученные по формулам (11), (12) дают смещённые оценки, которые приводят к статистическим ошибкам, давая заниженные значения дисперсии генеральной совокупности. Для получения их несмещённых оценок дисперсию “исправляют”. Исправленную дисперсию обозначают через S^2 :

$$S^2(x) = \frac{1}{n-1} \cdot \bar{D}(x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^k m_i \cdot (x_i - \bar{x})^2 ; \quad (13)$$

тогда несмещённая оценка среднее квадратическое отклонения:

$$S(x) = \sqrt{S^2(x)} . \quad (14)$$

При небольшом объеме выборки ($n \leq 30$ значений) расхождение между $\sigma(x)$ и $S(x)$ оказывается существенным [1], поэтому для расчета дисперсии и среднее квадратическое отклонения в этом случае следует пользоваться формулами (13), (14). При большом объеме выборки ($n > 30$ значений) расхождение между $\sigma(x)$ и $S(x)$ незначительно и практически пропадает с увеличением числа статистических данных, поэтому дисперсию и среднее квадратическое отклонение можно определять по формулам (11), (12).

3. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИСХОДНЫХ ДАННЫХ

Исходные данные, полученные в результате измерений величины напряжения, активной и реактивной мощности, тока, погрешности приборов и т.п., могут содержать значения этих величин, попавшие в данную выборку случайно, в результате воздействия каких-либо не характерных для изучаемого процесса факторов. При проведении следующей серии подобных измерений эти случайные факторы не будут иметь место, но могут появиться другие. Например, при измерении активной мощности, потребляемой предприятием в часы максимальных нагрузок энергосистемы, оказалось, что в результате выхода из строя технологической линии был отключён один из цехов предприятия. Таким образом, в исследуемый период времени производилось не измерение активной мощности предприятия в нормальном режиме работы, а мощности, соответствующей работе предприятия без учёта одного из цехов. Такие исходные данные не являются характерными для производимого исследования, т.е. недостоверными, и для того, чтобы они не исказили полученных результатов, производится предварительная обработка этих исходных данных.

Предварительная обработка исходных данных преследует две основные цели:

- “очистка” исходных данных от недостоверных значений;
- предварительная оценка вида теоретического распределения, с которым согласуется данное эмпирическое.

“Очистка” исходных данных — это исключение из исходных данных тех выборочных значений, которые соответствуют грубым ошибкам и локальным выбросам.

Грубые ошибки в исходных данных могут быть вызваны неисправностью измерительных приборов, трансформаторов тока или напряжения, ошибкой при снятии показаний приборов и т.д.

Локальные выбросы не связаны с ошибкой значения исходных данных. Они столь сильно отличаются от большинства значений в выборке, что их можно считать исходными данными другой выборки и отбрасывать.

Для выявления грубых ошибок и локальных выбросов применяют два основных способа очистки:

- с использованием правила трёх сигм;
- с использованием гистограмм частот.

“ОЧИСТКА” ИСХОДНЫХ ДАННЫХ ПО ПРАВИЛУ ТРЁХ СИГМ

Сущность правила трёх сигм состоит в следующем: если случайная величина X распределена по нормальному закону, то абсолютная величина её отклонения от математического ожидания не превосходит утроенного среднеквадратического отклонения.

В [1] показано, что вероятность того, что отклонение по абсолютной величине будет меньше утроенного среднеквадратического отклонения, равна 0,9973. Другими словами, вероятность того, что абсолютная величина отклонения превысит утроенное среднеквадратическое отклонение, очень мала, а именно равна $1 - 0,9973 = 0,0027$. В теории вероятности и математической статистике такие события, вероятность наступления которых не превышает 0,05, называют практически невозможными.

Используя это правило, можно отстроиться от недостоверных исходных данных. Для этого необходимо определить оценки числовых характеристик исследуемой СВ X по формулам (9),(12),(13) и построить указанные интервалы

$$a(x) = \pm 3S(x). \quad (14)$$

Те из значений исходных данных, которые не попали в указанный интервал, можно считать полученными в результате воздействия случайных факторов, т.е. недостоверными, и отбрасывать.

“ОЧИСТКА” ИСХОДНЫХ ДАННЫХ С ПРИМЕНЕНИЕМ ГИСТОГРАММ ЧАСТОТ

Для наглядности строят различные графики статического распределения [1]. Например, для дискретной случайной величины обычно строится полигон частот (Рис. 2) — ломанную, отрезки которой соединяют точки $(x_1; m_1); (x_2; m_2); \dots; (x_r; m_r)$.

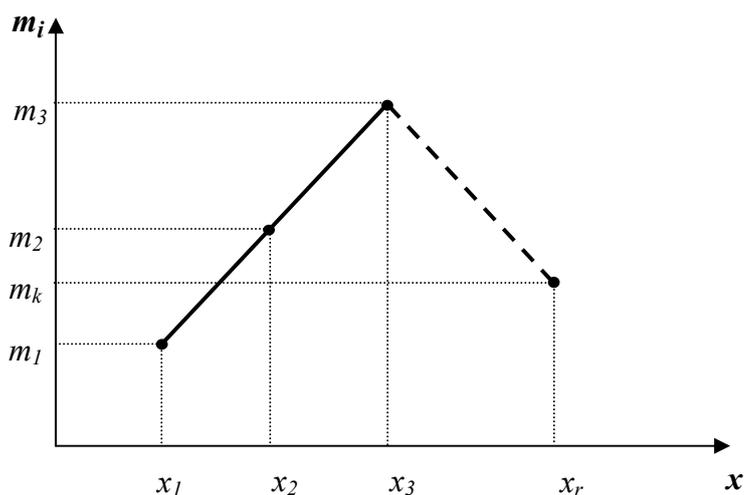


Рис. 2

В случае непрерывной случайной величины, целесообразно строить гистограмму частот (рис. 3) — ступенчатую фигуру, состоящую из прямоугольников, основанием которых служат разряды $(x_{n-1}; x_n)$ длиной h , а высоты равны значениям частот m_i соответствующих i -му разряду или значениям плотностей распределения частот в i -м разряде:

$$f^*(x) = \frac{P_i^*}{h}. \quad (15)$$

Площадь i -го прямоугольника равна $f^*(x_i) \cdot h = \frac{P_i^*}{h} \cdot h = P_i^*$, т.е. относительной частоте попадания СВ X в i -ый разряд. Площадь всей гистограммы равна сумме относительных частот всех разрядов:

$$\sum_{i=1}^r P_i^* = 1. \quad (16)$$

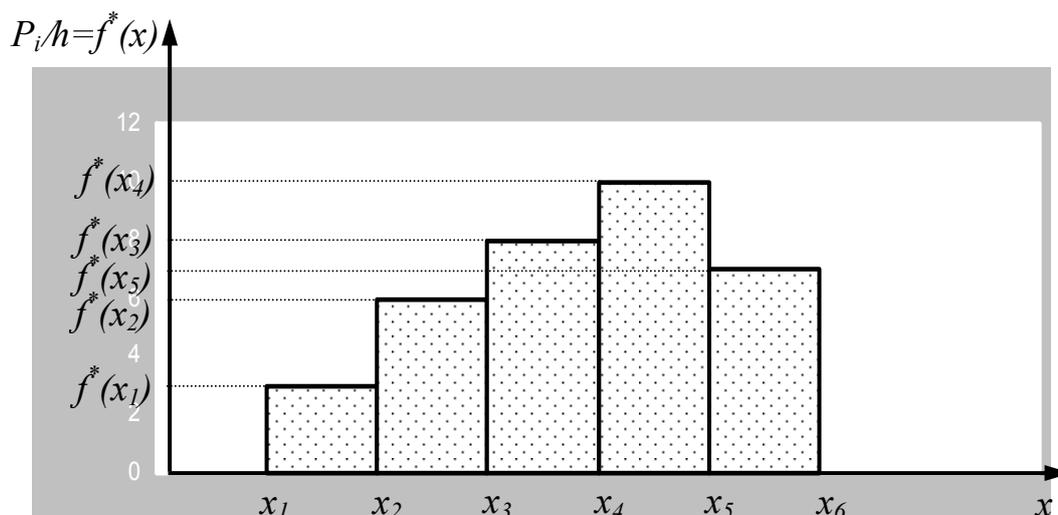


Рис. 3

Различают гистограммы мгновенные и итоговые.

Мгновенными называются гистограммы, соответствующие исходным данным за относительно короткий промежуток времени. Например, измеряют значения активной мощности, потребляемой предприятием в часы утреннего максимума энергетической системы с 8⁰⁰ до 10⁰⁰ с трёхминутным интервалом в первую декаду месяца (аналогично, во вторую и третью декаду). Соответствующая **итоговая** гистограмма строится по значениям мощности всех трёх декад.

Как правило, первоначально осуществляется построение мгновенных гистограмм, с помощью которых производится “очистка” исходных данных путём выявления на ней “островков”. В этом случае можно использовать гистограмму, у которой по оси ординат отложены значения частот m_i . На Рис. 4 представлена гистограмма распределения частот СВ активной мощности, P, МВт. На ней имеется “островок” значений (2÷2,1) МВт. Значения СВ P, попавшие в этот разряд, можно считать недостоверными и отбросить.

По сумме очищенных мгновенных гистограмм строится итоговая гистограмма. Т.к. она является объектом более глубокого анализа, то рекомендуется

при её построении использовать ординаты в значениях плотности распределения относительных частот $f^*(x)$, рассчитанных по формуле (15).

Пример 1: Результаты $n = 48$ измерений активной мощности, P , МВт, потребляемой предприятием, представлены в виде статистического ряда распределения:

Таблица 4

Разряды P , МВт	2÷2,1	2,1÷2,2	2,2÷2,3	2,3÷2,4	2,4÷2,5	2,5÷2,6	2,6÷2,7	2,7÷2,8	2,8÷2,9	2,9÷3,0
Частота m_i	1	0	0	0	5	7	10	12	7	5

Построить гистограмму данного статистического ряда. Очистить полученные исходные данные с использованием построенной гистограммы.

Решение. Результатом решения данной задачи является гистограмма, представленная на рис. 4. Значения СВ P в диапазоне (2 – 2,1)МВт можно отнести к локальному выбросу и отбросить. Для дальнейшего анализа остаются значения СВ P в диапазоне (2,4 – 3,0)МВт.

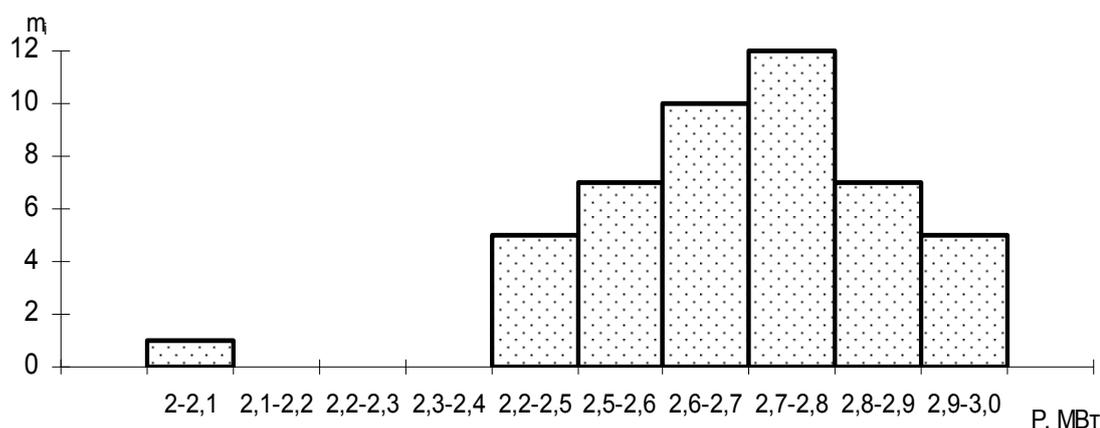


Рис. 4

Гистограмму и полигон статистического ряда удобно также использовать для предварительной оценки вида закона распределения, с которым согласуется данная выборкой. Достигается это визуальной оценкой формы гистограммы или полигона соответствующих очищенных исходных данных, насколько она соот-

ветствует известным законам распределения. Например, гистограмма, представленная на рис.4, напоминает кривую Гаусса, поэтому можно выдвинуть гипотезу о нормальном распределении СВ активной мощности, потребляемой предприятием.

4. ВЫРАВНИВАНИЕ СТАТИСТИЧЕСКИХ РАСПРЕДЕЛЕНИЙ

Во всяком статистическом распределении неизбежно присутствуют элементы случайности, связанные с тем, что число опытов ограничено, что произведены именно те, а не другие опыты, давшие именно те, а не другие результаты. Только при очень большом числе опытов эти случайности сглаживаются, и явление обнаруживает в полной мере присущие ему закономерности. На практике почти никогда не имеется такого большого числа опытов (измерений). Поэтому часто приходится решать вопрос о том, как подобрать для данного статистического распределения аналитическую формулу, выражающую лишь существенные черты статистического материала и не учитывающую случайности, связанные с недостаточным объёмом опытных данных [3]. Выравнивание статистического распределения – это замена гистограммы относительных частот случайной величины плавной кривой, имеющей достаточно простое аналитическое выражение. Принципиальный вид выравнивающей плавной кривой выбирается заранее, чаще на основании визуальной оценки формы гистограммы.

В качестве плавной кривой, заменяющей гистограмму, используется плотность распределения вероятностей СВ X - $f(x)$ – производная функции распределения этой случайной величины:

$$f(x) = F'(x), \quad (18)$$

а следовательно функция распределения СВ X определяется:

$$F(x) = \int_{-\infty}^x f(t)dt \quad (19)$$

Понятие плотности распределения СВ X применимо только для непрерывной случайной величины. График кривой $f(x)$ называют кривой распределения.

Одним из наиболее распространённых распределений, с которым согласуются эмпирические распределения, является нормальное (распределение Гаусса). Объяснение этому дано русским математиком А.М. Ляпуновым [4]: если случайная величина X представляет собой сумму очень большого числа взаимно независимых случайных величин, влияние каждой из которых на всю сумму ничтожно мало, то СВ X имеет распределение, близкое к нормальному.

Нормальным называют распределение вероятностей непрерывной СВ, которое описывается плотностью:

$$f(x) = \frac{1}{\sigma(x) \cdot \sqrt{2\pi}} e^{-\frac{(x-M(x))^2}{2\sigma(x)^2}}, \quad (20)$$

где $M(x)$ – математическое ожидание СВ X , $\sigma(x)$ – среднеквадратическое отклонение СВ X - являются параметрами нормального распределения.

График плотности нормального распределения (рис.5) называют нормальной кривой (кривой Гаусса).

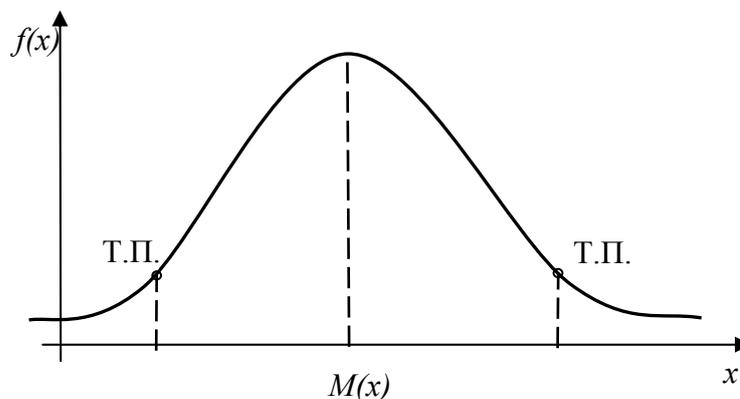


Рис.5

Свойства нормального распределения (рис.5):

1. Кривая лежит выше оси Ox , т.е. $f(x) > 0$;
2. Ветви кривой асимптотически приближаются к оси Ox , т.е. $\lim_{x \rightarrow \pm\infty} f(x) = 0$;

3. Кривая имеет максимум в точке $\left(M(x); \frac{1}{\sigma(x) \cdot \sqrt{2\pi}} \right)$; кривая симметрична относительно этой точки;

4. Кривая имеет две точки перегиба (Т.П.) с координатами

$$\left(M(x) - \sigma(x); \frac{1}{\sigma(x) \cdot \sqrt{2\pi e}} \right) \text{ и } \left(M(x) + \sigma(x); \frac{1}{\sigma(x) \cdot \sqrt{2\pi e}} \right);$$

5. Площадь под кривой, ограниченная вертикальными прямыми в точках $x = M(x) \pm \sigma(x)$, равна 0,64; $x = M(x) \pm 2 \cdot \sigma(x)$, равна 0,95; $x = M(x) \pm 3 \cdot \sigma(x)$, равна 0,9973. Площадь под кривой в пределах $]-\infty; +\infty[$ равна единице.

На форму и распределение нормальной кривой оказывает влияние значения её параметров – $M(x)$ и $\sigma(x)$:

- изменения величины математического ожидания $M(x)$ не изменяет формы нормальной кривой, а лишь перемещает её вдоль оси Ox вправо при увеличении $M(x)$, или влево, если $M(x)$ уменьшается;
- с возрастанием $\sigma(x)$ максимальная ордината нормальной кривой убывает, а сама кривая становится более полой; при убывании $\sigma(x)$ кривая становится более островершинной (рис. 6).

При $M(x) = 0$ и $\sigma(x) = 1$ нормальную кривую называют нормированной.

“Нормирование” – это переход от именованной СВ X с параметрами $M(x)$ и $\sigma(x)$ и плотностью распределения $f(x; M(x); \sigma(x))$ к нормированной СВ Z с параметрами $M(z) = 0$ и $\sigma(z) = 1$ и плотностью распределения $f(z; 0; 1)$. В этом случае осуществляется перенос начала координат в точку $x = M(x)$, а в качестве единицы масштаба выбирается $\sigma(x)$. Переход к нормированной пере-

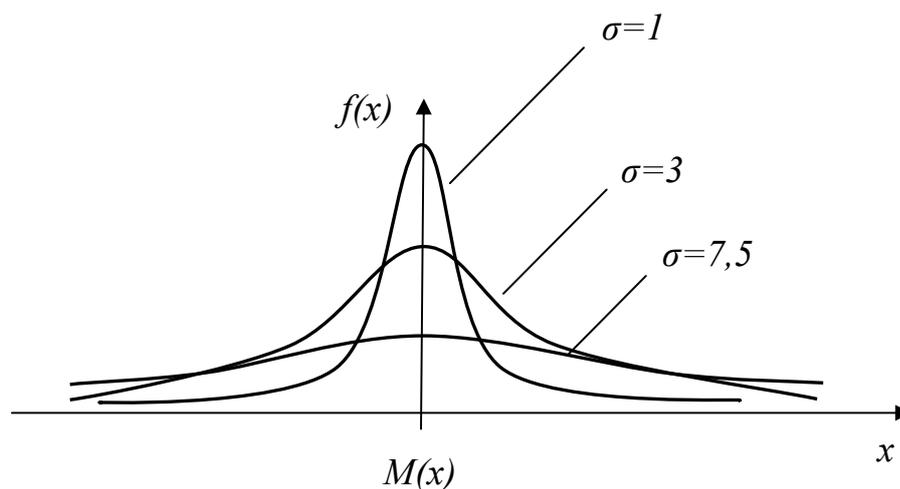


Рис.6

менной z осуществляется по формуле:

$$z = \frac{x - M(x)}{\sigma(x)} \quad (21)$$

Плотность нормированного нормального распределения $f(z)$ определяется по таблице приложения 1.

Пример 3. Произведено $n = 50$ измерений отклонения от заданного значения реактивной мощности (Q, вар), выдаваемой предприятием в сеть энергосистемы. Результаты обработки исходных данных сведены в таблицу 5. Выровнять статистическое распределение СВ Q нормированным нормальным законом.

Таблица 5

Разряды, Q, вар	60 ÷ 70	70 ÷ 80	80 ÷ 90	90 ÷ 100
n_i	9	18	14	9

Решение. Определим значения параметров нормального распределения по (10), (12):

$$M(Q) = \frac{1}{4} \sum_{i=1}^4 m_i \cdot Q_i = 78,8 \text{ вар};$$

$$\sigma(Q) = \sqrt{\frac{1}{4} \left(\sum_{i=1}^4 (Q_i - M(Q))^2 \cdot m_i \right)} = \sqrt{96,9} = 9,9 \text{ вар}.$$

Пронормируем это распределение. Для этого ширину разрядов h выразим в долях $\sigma(Q)$ и получим длину нормированных разрядов:

$$\Delta z = \frac{h}{\sigma(Q)} = \frac{10}{9,9} = 1,01;$$

Определяем левую границу первого разряда (по 21):

$$z_1 = \frac{60 - 78,8}{9,9} = -1,9,$$

далее рассчитываем остальные границы разрядов:

$$z_2 = z_1 + \Delta z = -1,9 + \Delta z = -1,9 + 1,01 = -0,89,$$

$$z_3 = z_2 + \Delta z = -0,89 + 1,01 = 0,12 \text{ и т.д.}$$

Аналогично нормируем середины разрядов. Результаты расчётов сводим в таблицу 6:

Таблица 6

Разряды, $Q, \text{вар}$	60 ÷ 70	70 ÷ 80	80 ÷ 90	90 ÷ 100
Середины разрядов, $Q, \text{вар}$	65	75	85	95
Частота m_i	9	18	14	9
Относительные частоты $P^* = \frac{m_i}{n}$	0,18	0,36	0,28	0,18
Разряды, z	(-1,9) ÷ (-0,89)	(-0,89) ÷ 0,12	0,12 ÷ 1,13	1,13 ÷ 2,14
Середины разрядов, z	-1,39	-0,38	0,63	1,64

$f^*(z) = \frac{P^*}{\Delta z}$	0,18	0,36	0,28	0,18
$f(z)$ – по прилож. 1	0,1518	0,3711	0,3271	0,1040

По полученным данным строим гистограмму плотности распределения относительных частот $f^*(z)$ (рис.7.). Затем по таблице приложения 1 находим значения плотности распределения вероятностей нормированного нормального распределения $f(z)$ для соответствующих значений середин разрядов.

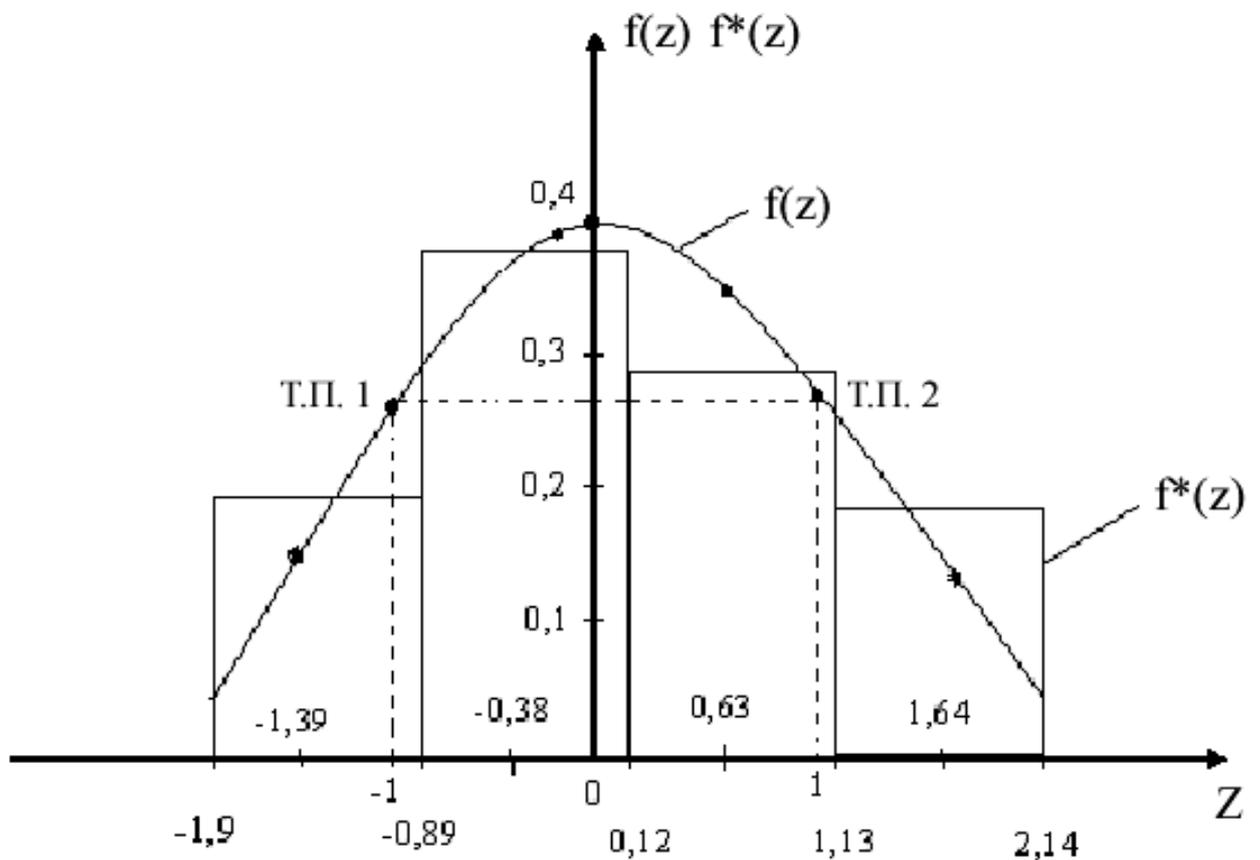


Рис.7

Для построения кривой нормированного нормального распределения следует определить ещё три дополнительных точки (рис. 7):

- 1) точка вершины имеет координату по оси абсцисс $M(x)$; для нормированной величины $M(z) = 0$; по приложению 1 для $f(0) = 0,3989$; получаем точку вершины с координатами $(0; 0,3989)$;
- 2) точки перегиба имеют координаты по оси абсцисс $M(x) \pm \sigma(x)$; для нормированной величины $M(z) = 0$, $\sigma(z) = 1$; по приложению 1 определяем $f(1) = 0,2420$; получаем точки перегиба с координатами $(-1; 0,2420)$; $(1; 0,2420)$.

5. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Во многих случаях для анализа результатов измерений исследуемой случайной величины необходимо знать закон распределения генеральной совокупности, оцениваемой с помощью полученной выборки. Если закон распределения неизвестен [1], но имеются основания предположить, что он имеет определённый вид (например, нормальный), выдвигают гипотезу: генеральная совокупность распределена по нормальному закону. Таким образом, в этой гипотезе речь идёт о виде предполагаемого распределения. Известно [5], что массовые, часто повторяющиеся события в большей или меньшей степени подчиняются нормальному закону распределения. Это касается срока службы оборудования, приборов, величин потребляемой активной и реактивной мощности, удельного электропотребления, колебания напряжения и т.д.

Возможен случай, когда закон распределения известен, а его параметры неизвестны. В этом случае выдвигаются гипотезы о предполагаемой величине параметра известного распределения.

Статистической называют гипотезу о виде неизвестного распределения или о параметрах известного распределения [1].

Нулевой (основной) называют выдвинутую гипотезу H_0 . *Конкурирующей* (альтернативной) называют гипотезу H_1 , которая противоречит нулевой.

Выдвинутая гипотеза может быть правильной или неправильной, поэтому возникает необходимость её проверки. Поскольку для проверки используют статистические методы, её называют статистической проверкой гипотезы.

При проверке статистических гипотез по выбранным данным могут быть допущены два вида ошибок:

- ошибка первого рода состоит в том, что будет отвергнута правильная гипотеза;
- ошибка второго рода состоит в том, что будет принята неправильная гипотеза.

Последствия этих ошибок могут оказаться различными. Например, если отвергнуто правильное решение “продолжить строительство жилого дома”, то эта ошибка первого рода повлечет материальный ущерб; если же принято неправильное решение “продолжить строительство”, несмотря на опасность обвала стройки, то эта ошибка второго рода может повлечь гибель людей.

Вероятность совершить ошибку первого рода называется *уровнем значимости* и обозначается α .

В большинстве случаев на практике применяются *статистические критерии значимости* [2] — это критерии, на основании применения которых с заранее заданным уровнем значимости принимается только одно решение: “Отклонить проверяемую нулевую гипотезу”. Если же для проверяемой нулевой гипотезы нет основания для её отклонения данным критерием, то утверждается, что результаты выборки не дают основания для отклонения выдвинутой нулевой гипотезы. При этом выбор уровня значимости α осуществляется до проверки гипотезы. При проведении технических исследований обычно принимается уровень значимости $\alpha = 0,05$ исходя из того, что в теории вероятностей и математической статистике события, вероятность наступления которых не превышает 5%(0,05), считают практически невозможными.

Проверка статистических гипотез осуществляется на основании данных выборки. Путем проверки соответствующих гипотез решаются такие часто возни-

кающие задачи, как сравнение однопоточных производств по удельному расходу, сравнение новой и старой технологий производства какого-либо изделия, проверки точности работы различных приборов, регуляторов и т.п. Для проверки нулевой гипотезы используют *статистический критерий* K — специально подобранную случайную величину, точное или приближённое распределение которой известно. *Наблюдаемым значением* $K_{набл.}$ называют значение критерия, вычисленное по выборкам. В зависимости от вида рассматриваемого распределения критерии имеют разные обозначения: Z — критерий нормального распределения; F — критерий распределения Фишера; t — критерий распределения Стьюдента и т.д.

ПРОВЕРКА ГИПОТЕЗЫ О РАВЕНСТВЕ МАТЕМАТИЧЕСКИХ ОЖИДАНИЙ ДВУХ СЛУЧАЙНЫХ ВЕЛИЧИН, ИМЕЮЩИХ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

Требуется при заданном уровне значимости α проверить нулевую гипотезу $H_0: M(X) = M(Y)$ о равенстве математических ожиданий двух нормальных совокупностей с неизвестными, но одинаковыми дисперсиями (в случае независимых малых выборок) при альтернативной гипотезе $H_0: M(X) \neq M(Y)$.

Такая задача ставится потому, что, как правило, математическое ожидание выборок оказываются численно различными. Возникает вопрос: значимо или незначимо это различие. Если подтвердится, что нулевая гипотеза справедлива, то различие математических ожиданий незначимо (т.е. они равны) и объясняется случайными причинами. Если справедливой окажется альтернативная гипотеза, то различие математических ожиданий считается значимым.

Для проверки этой гипотезы надо вычислить наблюдаемое значение критерия [1]:

$$t_{\text{набл.}} = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1) \cdot S(X)^2 + (n_2 - 1) \cdot S(Y)^2}} \cdot \sqrt{\frac{n_1 \cdot n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}, \quad (22)$$

где \bar{x}, \bar{y} – средние значения выборок СВ X и Y , определяемые по формуле (10); n_1, n_2 – объёмы выборок СВ X и Y ; $S^2(X), S^2(Y)$ – дисперсии соответствующих выборок, определяемые по формуле (13).

По таблице критических точек распределения Стьюдента (таблица приложения 3) по заданному уровню значимости α и числу степеней свободы

$$k = n + m - 2 \quad (23)$$

определяется критическое значение $t_{\text{крит.}}(\alpha; k)$. Если $|t_{\text{набл.}}| < t_{\text{крит.}}(\alpha, k)$, отвергнуть нулевую гипотезу нет оснований; если $|t_{\text{набл.}}| > t_{\text{крит.}}(\alpha, k)$, нулевую гипотезу отвергают.

В случае, если нет оснований считать дисперсии сравниваемых выборок одинаковыми то, прежде чем сравнивать математические ожидания, следует [4] проверить гипотезу о равенстве дисперсий (п. 4.2.).

Пример 4. Произведены замеры отклонений напряжений на шинах ТП в дневное и ночное время суток (малые выборки): $n_D = 5$, $n_H = 6$. Оценки математических ожиданий этих выборок $\bar{U}_D = 3,3 \text{ В}$; $\bar{U}_H = 2,48 \text{ В}$ и несмещенные оценки дисперсий $S_{U_D}^2 = 0,25 \text{ В}^2$ и $S_{U_H}^2 = 0,108 \text{ В}^2$. При уровне значимости $\alpha = 0,05$ проверить нулевую гипотезу о равенстве средних значений напряжения в ночное и дневное время суток $H_0: M(U_D) = M(U_H)$.

Решение. Рассчитаем наблюдаемое значение критерия Стьюдента:

$$t_{\text{набл.}} = \frac{3,3 - 2,48}{\sqrt{(5 - 1) \cdot 0,25 + (6 - 1) \cdot 0,108}} \cdot \sqrt{\frac{5 \cdot 6 (5 + 6 - 2)}{5 + 6}} = 3,27.$$

По таблице приложения 3 определим критическое значение критерия при числе степеней свободы $k = 5 + 6 - 2 = 9$: $t_{\text{крит.}}(0,05; 9) = 2,26$.

Так как $|t_{набл.}| > t_{крит.}$, гипотезу о равенстве двух центров распределений отвергаем, т.е. средние значения напряжения в дневное и в ночное время суток различны.

5.2 ПРОВЕРКА ГИПОТЕЗЫ О РАВЕНСТВЕ ДИСПЕРСИЙ ДВУХ СЛУЧАЙНЫХ ВЕЛИЧИН, ИМЕЮЩИХ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ

На практике задача сравнения дисперсий возникает, если требуется сравнить точность приборов, инструментов, самих методов измерений и т.д. Очевидно, предпочтительнее тот прибор, инструмент и метод, который обеспечивает наименьшее рассеяние результатов измерений, т.е. наименьшую дисперсию.

Пусть генеральные совокупности X и Y распределены нормально. Из них извлечены выборки объемами соответственно n и m , по которым рассчитаны несмещенные оценки дисперсий $S^2(X)$ и $S^2(Y)$. Требуется при заданном уровне значимости проверить нулевую гипотезу, состоящую в том, что дисперсии генеральных совокупностей равны между собой: $H_0 : D(X) = D(Y)$ при $H_1 : D(X) \neq D(Y)$.

Такая задача встает потому, что чаще всего несмещенные оценки дисперсии различны, и возникает вопрос, значимо или нет это различие. Если окажется, что нулевая гипотеза справедлива, т.е. дисперсии генеральных совокупностей равны, то различие несмещенных оценок дисперсий незначимо и объясняется случайными факторами. В противном случае численное различие между дисперсиями признается значимым, а дисперсии неравными.

В качестве критерия проверки нулевой гипотезы о равенстве дисперсий двух генеральных совокупностей используется критерий Фишера, наблюдаемое значение которого определяется по формуле:

$$F_{набл.} = \frac{S_{\delta}^2 - \text{большая из двух дисперсий}}{S_m^2 - \text{меньшая из двух дисперсий}} \quad (24)$$

Далее по таблице приложения 5 находят критическое значение этого критерия $F_{крит.}(k_1; k_2)$ при числе степеней свободы, определяемом как

$$k = n - 1. \quad (25)$$

Если $F_{набл} < F_{крит}$, нет оснований отвергать нулевую гипотезу; если $F_{набл} > F_{крит}$ – нулевая гипотеза о равенстве двух дисперсий отвергается.

Пример 5. По результатам измерений двумя амперметрами значений силы тока во вторичной обмотке трансформатора тока сформировано две независимые выборки объемами $n_1=10$, $n_2=18$, по которым определены соответствующие несмещенные оценки дисперсий $S_1^2 = 1,23 A^2$, $S_2^2 = 0,41 A^2$. При уровне значимости $\alpha = 0,05$ проверить нулевую гипотезу о равенстве дисперсий генеральных совокупностей, т.е. что приборы имеют одинаковую точность.

Решение. Определяем наблюдаемое значение критерия Фишера:

$$F_{набл.} = \frac{1,23}{0,41} = 3.$$

По таблице приложения 5 находим критическое значение этого критерия при $k_1 = 10 - 1 = 9$, $k_2 = 18 - 1 = 17$:

$$F_{крит.}(k_1; k_2) = F(9; 17) = 2,5.$$

Поскольку $F_{набл.} > F_{кр.}$, следовательно, нулевая гипотеза о равенстве дисперсий двух генеральных совокупностей отвергается, т.е. приборы имеют разную точность.

При этом для осуществлений измерений следует предпочесть тот прибор, который обеспечивает большую точность, т.е. обладает меньшей дисперсией - $S_2^2 = 0,41 A^2$.

5.3 ПРОВЕРКА ГИПОТЕЗЫ О НОРМАЛЬНОМ РАСПРЕДЕЛЕНИИ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ

При обработке результатов эксперимента по виду гистограммы, полигона или из каких-либо других соображений выдвигается гипотеза о законе распределения СВ, производится оценка параметров этого распределения (п.2,3,4). Дальнейшая задача состоит в проверке выдвинутой гипотезы о законе распределения, насколько хорошо подобрана вероятностная модель ряда наблюдений.

Проверка гипотезы о виде предполагаемого распределения производится с помощью непараметрических критериев согласия. Наиболее распространенным среди них является критерий χ^2 (Пирсона).

При использовании критерия χ^2 (Пирсона) для проверки гипотезы о нормальном распределении генеральной совокупности следует иметь в виду, что объём выборки из этой совокупности должен составлять не менее 50 значений случайной величины.

Порядок проверки нулевой гипотезы о виде закона распределения с использованием критерия χ^2 (Пирсона):

1. Определяются вероятности попадания СВ X в соответствующие разряды $x_{i-1}; x_i$:

$$P(x_{i-1} < x < x_i) = \Phi_0\left(\frac{x_i - \bar{x}}{S(x)}\right) - \Phi_0\left(\frac{x_{i-1} - \bar{x}}{S(x)}\right), \quad (26)$$

где Φ_0 – функция Лапласа, значение которой определяется в соответствии с таблицей приложения 2.

2. Вычисляют наблюдаемое значение критерия $\chi_{набл}^2$:

$$\chi_{набл}^2 = n \sum_{i=1}^r \frac{(P_i - P_i^*)^2}{P_i} = \sum_{i=1}^r \frac{(m_i - nP_i)^2}{nP_i}, \quad (27)$$

где r – число разрядов; P_i^* – относительная частота попадания СВ X в i -ый разряд; P_i – вероятность попадания СВ X в i -ый разряд; m_i – частота попадания СВ X в i -ый разряд; n – объём выборки.

3. Находят критическое значение критерия $\chi_{крит.}^2(\alpha; k)$ (по таблице приложения 4) в зависимости от α – уровня значимости и k – числа степеней свободы:

$$k = r - 1 - S, \quad (28)$$

где r – число разрядов; S - число параметров предполагаемого распределения, которые оценены по данным выборки.

3. Если $\chi_{набл.}^2 < \chi_{крит.}^2$, гипотеза о предполагаемом виде закона распределения верна; если $\chi_{набл.}^2 > \chi_{крит.}^2$, гипотеза о предполагаемом виде закона распределения отвергается.

Для определения вероятности попадания СВ X в соответствующий разряд применяется функция Лапласа:

$$\Phi_0(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt, \quad (29)$$

где z – нормированная СВ X , рассчитываемая по формуле (21).

Функция Лапласа равна вероятности попадания СВ z в интервал от $z = 0$ до $z = z_1$ и геометрически она представляет собой площадь под нормированной кривой, ограниченную вертикальными прямыми $z = 0$ и $z = z_1$ (рис. 5).

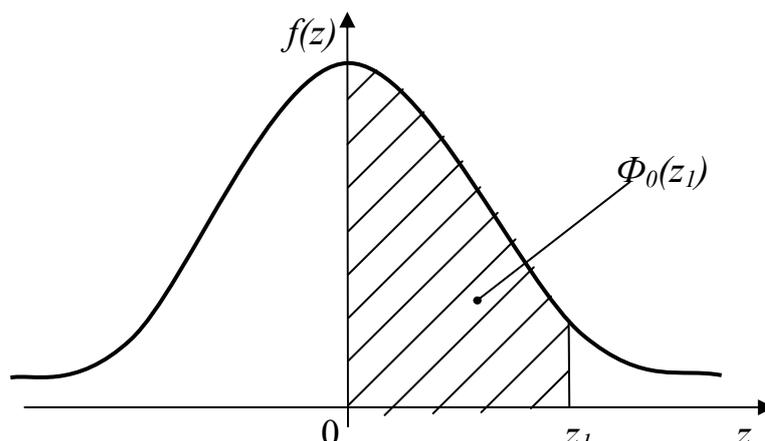


Рис. 5

С использованием формулы Лапласа можно определить вероятность попадания СВ в некоторый интервал по (26), а также вероятность неперевышения некоторой заданной величины:

$$P(X < x_1) = 0,5 + \Phi_0\left(\frac{x_1 - \bar{x}}{S(x)}\right). \quad (30)$$

Свойства функции Лапласа:

1. $\Phi_0(0) = 0$;
2. $\Phi_0(-\infty) = -0,5$;
3. $\Phi_0(\infty) = 0,5$;
4. $\Phi_0(-z) = -\Phi_0(z)$.

Значения $\Phi_0(z)$ табулированы и определяются по таблице приложения 2.

Пример 6. На основании исходных данных и произведенных расчётов примера 3 (п.4), результаты расчетов которого приведены в таблице 7, проверить с помощью критерия χ^2 (Пирсона) гипотезу о выравнивании статистического распределения СВ реактивной мощности Q нормированным нормальным законом.

Таблица 7

Разряды, z	$(-1,9) \div (-0,89)$	$(-0,89) \div 0,12$	$0,12 \div 1,13$	$1,13 \div 2,14$
Частоты, P^*	0,18	0,36	0,28	0,18
Вероятности, P	0,158	0,3611	0,323	0,121

Вероятность попадания СВ z в соответствующий разряд определяется с использованием функции Лапласа по формуле (4). Вероятность попадания в первый разряд:

$$P(-1,9 < z < -0,89) = \Phi_0(-0,89) - \Phi_0(-1,9) = \Phi_0(0,89) + \Phi_0(1,9) = -0,3133 + 0,4713 = 0,158.$$

Аналогично находим вероятности попадания СВ z в остальные разряды.

Наблюдаемое значение критерия (по 27):

$$\chi_{набл.}^2 = n \cdot \sum_{i=1}^r \frac{(P_i^* - P_i)^2}{P_i} = 50 \cdot \sum_{i=1}^4 \frac{(P_i^* - P_i)^2}{P_i} = 1,88$$

По таблице приложения 4 находим для $\alpha = 0,05$ и $k = 4 - 1 - 2 = 1$

$\chi_{крит.}^2(0,05; 1) = 3,8$. Т.к. $\chi_{набл.}^2 < \chi_{крит.}^2$, нет оснований отвергать нулевую гипотезу о нормальном распределении случайной величины реактивной мощности Q .

6 ОПРЕДЕЛЕНИЕ ИНТЕРВАЛЬНЫХ ОЦЕНОК МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Точечная оценка неизвестного параметра Θ , найденного по выборке объёмом n из генеральной совокупности с функцией распределения $F(x)$, не позволяет непосредственно ответить на вопрос, какую ошибку мы совершаем, принимая вместо точного значения неизвестного параметра некоторое его приближенное значение (оценку) $\bar{\Theta} = u(x_1, x_2, \dots, x_n)$. При выборке малого объёма точечная оценка может значительно отличаться от оцениваемого параметра, т. е. приводить к грубым ошибкам.

В связи с этим, при выборках малого объёма нужно пользоваться интервальной оценкой, основанной на определении некоторого интервала, внутри которого с определённой вероятностью находятся неизвестные значения параметра Θ .

Пусть найденная по результатам выборки объёмом n статическая характеристика $\bar{\Theta} = u(x_1, x_2, \dots, x_n)$ является точечной оценкой неизвестного параметра Θ . Чем меньше разность $|\Theta - \bar{\Theta}|$, тем лучше качество оценки, тем точнее оценка. Таким образом, положительное число ε , называемое *предельной погрешностью*, характеризует точность оценки

$$|\Theta - \bar{\Theta}| < \varepsilon \quad (31)$$

Однако статистические методы не позволяют утверждать, что оценка $\bar{\Theta}$ обязательно удовлетворяет неравенству (31), можно лишь говорить о вероятности γ , с которой это неравенство осуществляется.

Доверительной вероятностью (надёжностью) называют вероятность $P = (1 - \alpha) = \gamma$ выполнения неравенства (31), где α – уровень значимости. Обычно доверительной вероятностью оценки задаются заранее. В технических расчётах обычно полагают $\gamma = 1 - \alpha = 1 - 0,05 = 0,95$. Доверительная вероятность точечной оценки показывает, что при извлечении выборок объёмом n из одной и той же генеральной совокупности с функцией распределения $F(x)$ в $(1 - \alpha) \cdot 100\%$ случаях параметр Θ будет покрываться интервалом:

$$P(|\bar{\Theta} - \Theta| < \varepsilon) = \gamma, \quad (32)$$

$$P(\bar{\Theta} - \varepsilon < \Theta < \bar{\Theta} + \varepsilon) = \gamma. \quad (33)$$

Т.е. неизвестный параметр Θ заключён внутри интервала $(\bar{\Theta} - \varepsilon; \bar{\Theta} + \varepsilon)$, который называется **доверительным интервалом (точностью)**, покрывающим неизвестный параметр Θ с заданной доверительной вероятностью. Длина доверительного интервала равна 2ε .

В практических задачах важную роль играет длина доверительного интервала, причем, чем меньше его длина, тем точнее оценка. Если же длина доверительного интервала велика, то оценка малоприспособна для практики.

Рассмотренные величины ε и γ связаны между собой и, кроме того, зависят от объёма выборки n . Поэтому при интервальном оценивании неизвестного параметра может решаться одна из трех возможных задач:

- 1) найти доверительный интервал, определяемый предельной погрешностью ε , покрывающий с заданной доверительной вероятностью γ неизвестный параметр, оценка которого определяется на основании выборки объёмом n ;

- 2) определить, с какой доверительной вероятностью γ покрывает доверительный интервал длиной 2ε искомый параметр, оцениваемый по выборке объёмом n ;
- 3) найти минимальный объём выборки n , на основании которого можно было бы оценить искомый параметр с предельной погрешностью, не превышающей ε и доверительной вероятностью γ .

В случае, когда СВ X распределена по нормальному закону, а среднеквадратическое отклонение неизвестно, точечные оценки параметров нормального распределения – математического ожидания и среднеквадратического отклонения – определяются по выборке объёмом n по формулам (10 - 14).

Значение предельной погрешности ε , определяющей длину доверительного интервала, рассчитываются по формуле:

$$\varepsilon = t_{\alpha,k} \cdot \frac{S(x)}{\sqrt{n}}, \quad (34)$$

где $t_{\alpha,k}$ – квантиль распределения Стьюдента, определяемый по таблице приложения 3 в зависимости от α -уровня значимости и числа степеней свободы $k = n - 1$.

Доверительная вероятность того, что доверительный интервал покрывает неизвестный параметр, определяется решением задачи, обратной выше. Сначала находится значение квантиля распределения Стьюдента:

$$t_{\alpha,k} = \frac{\varepsilon \cdot \sqrt{n}}{S(x)}, \quad (35)$$

а затем по таблице приложения 3 определяется уровень значимости α , которому соответствует данное значение квантиля. Искомая вероятность $\gamma = 1 - \alpha$.

Минимальный объём выборки, обеспечивающий непревышение заданной предельной погрешности при заданной доверительной вероятности можно рассчитать по формуле:

$$n_{\min} = t_{\alpha, k}^2 \cdot \frac{S^2(x)}{\varepsilon^2} \quad (36)$$

Анализируя формулы (34,35,36), можно сделать следующие выводы:

- при увеличении объёма выборки n уменьшается предельная погрешность ε , т.е. возрастает точность оценки;
- увеличение доверительной вероятности γ приводит к увеличению значения квантиля $t_{\alpha, k}$ (см. приложение 3), а, следовательно, к увеличению длины доверительного интервала (32), т.е. к снижению точности оценки.

Пример 7. По результатам 10 измерений потребляемой активной мощности P , кВт, рассчитаны $\bar{M}(P) = 21 \text{ кВт}$, $S(P) = 3,8 \text{ кВт}$. Построить доверительный интервал для математического ожидания СВ P . Доверительную вероятность принять $\gamma = 1 - \alpha = 0,95$.

Решение: Найдём предельную погрешность оценки по (34):

$$\varepsilon = t_{\alpha, k} \frac{S(P)}{\sqrt{n}}.$$

Значение квантиля $t_{\alpha, k}$ для $\alpha = 0,05$ и $k = n - 1 = 10 - 1 = 9$ определим по таблице приложения 3 распределения Стьюдента: $t_{0,05; 9} = 2,26$.

$$\varepsilon = 2,26 \cdot \frac{3,8}{\sqrt{10}} = 2,72 \text{ кВт}.$$

Рассчитаем границы доверительного интервала:

$$\bar{a} - \varepsilon = 21 - 2,72 = 18,28 \text{ кВт};$$

$$\bar{a} + \varepsilon = 21 + 2,72 = 23,72 \text{ кВт}.$$

Таким образом с доверительной вероятностью $\gamma = 1 - \alpha = 0,95$ истинное значение математического ожидания СВ P лежит в диапазоне $(18,28 \text{ кВт}; 23,72 \text{ кВт})$.

Пример 8. По выборке объемом $n = 16$ результатов измерения погрешностей измерения амперметра I, A определены средние значения погрешности измерения $\bar{M}(I) = a = 20,2 A$ и оценка среднеквадратического отклонения $S(I) = 0,8 A$. Определить вероятность непревышения предельной погрешности $\varepsilon = 0,43 A$.

Решение: Определим по (35) значение квантиля распределения Стьюдента:

$$t_{\alpha,k} = \varepsilon \frac{\sqrt{n}}{S(I)} = 0,43 \cdot \frac{\sqrt{16}}{0,8} = 2,13$$

По таблице приложения 3 найдём значение уровня значимости α , которому соответствует данное значение квантиля, при $k = n - 1 = 15$:

$$\alpha = 0,05.$$

Тогда искомая вероятность: $\gamma = 1 - \alpha = 1 - 0,05 = 0,95$.

Пример 9. Произведено $n = 18$ измерений значения СВ коэффициента мощности, $\cos \varphi$. По результатам замеров рассчитаны значения $\bar{M}(\cos \varphi) = a = 0,984$ и $S(\cos \varphi) = 0,092$. Найти минимальное число измерений, которое необходимо выполнить, что бы с доверительной вероятностью $\gamma = 1 - \alpha = 0,95$ можно было утверждать, что предельная погрешность точечной оценки значения коэффициента мощности не превысила $0,003$.

Решение: Искомую величину определим по (36):

$$n_{\min} = t_{\alpha,k}^2 \frac{S^2(\cos \varphi)}{\varepsilon^2}.$$

Значение $t_{\alpha,k}^2$ находим по таблице приложения 3 для $\alpha = 0,05$ и $k = n - 1 = 17$:

$$t_{0,05;17} = 2,11.$$

$$n_{\min} = 2,11^2 \cdot \frac{0,0092^2}{0,003^2} = 41,87 \approx 42 \text{ измерения.}$$

Таким образом, для того, чтобы предельная погрешность истинного значения $\cos \varphi$ не превышала $\varepsilon = 0,003$ с доверительной вероятностью 0,95, необходимо помимо 18 проделанных измерений произвести ещё $42 - 18 = 24$ измерения.

ПРИЛОЖЕНИЯ

Приложение 1

Таблица значений функции $f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

	0	1	2	3	4	5	6	7	8	9
0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013

	0	1	2	3	4	5	6	7	8	9
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
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	0001	0001

Приложение 2

Таблица значений функции Лапласа $\Phi_0(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-t^2/2} dt$

z	$\Phi_0(z)$	z	$\Phi_0(z)$	z	$\Phi_0(z)$	z	$\Phi_0(z)$
0,00	0,0000	0,24	0,0948	0,48	0,1844	0,72	0,2642
0,01	0,0040	0,25	0,0987	0,49	0,1879	0,73	0,2673
0,02	0,0080	0,26	0,1026	0,50	0,1915	0,74	0,2703
0,03	0,0120	0,27	0,1064	0,51	0,1950	0,75	0,2734
0,04	0,0160	0,28	0,1103	0,52	0,1985	0,76	0,2764
0,05	0,0199	0,29	0,1141	0,53	0,2019	0,77	0,2794
0,06	0,0239	0,30	0,1179	0,54	0,2054	0,78	0,2823
0,07	0,0279	0,31	0,1217	0,55	0,2088	0,79	0,2852
0,08	0,0319	0,32	0,1255	0,56	0,2123	0,80	0,2881
0,09	0,0359	0,33	0,1293	0,57	0,2157	0,81	0,2910
0,10	0,0398	0,34	0,1331	0,58	0,2190	0,82	0,2939
0,11	0,0438	0,35	0,1368	0,59	0,2224	0,83	0,2967
0,12	0,0478	0,36	0,1406	0,60	0,2257	0,84	0,2995
0,13	0,0517	0,37	0,1443	0,61	0,2291	0,85	0,3023
0,14	0,0557	0,38	0,1480	0,62	0,2324	0,86	0,3051
0,15	0,0596	0,39	0,1517	0,63	0,2357	0,87	0,3078
0,16	0,0636	0,40	0,1554	0,64	0,2389	0,88	0,3106
0,17	0,0675	0,41	0,1591	0,65	0,2422	0,89	0,3133
0,18	0,0714	0,42	0,1628	0,66	0,2454	0,90	0,3159
0,19	0,0753	0,43	0,1664	0,67	0,2486	0,91	0,3186
0,20	0,0793	0,44	0,1700	0,68	0,2517	0,92	0,3212
0,21	0,0832	0,45	0,1736	0,69	0,2549	0,93	0,3238
0,22	0,0871	0,46	0,1772	0,70	0,2580	0,94	0,3264
0,23	0,0910	0,47	0,1808	0,71	0,2611	0,95	0,3289

z	$\Phi_0(z)$	z	$\Phi_0(z)$	z	$\Phi_0(z)$	z	$\Phi_0(z)$
0,96	0,3315	1,37	0,4147	1,78	0,4625	2,30	0,4909
0,97	0,3340	1,38	0,4162	1,79	0,4633	2,38	0,4913
0,98	0,3365	1,39	0,4177	1,80	0,4641	2,40	0,4918
0,99	0,3389	1,40	0,4192	1,81	0,4649	2,42	0,4922
1,00	0,3413	1,41	0,4207	1,82	0,4056	2,44	0,4927
1,01	0,3438	1,42	0,4222	1,83	0,4664	2,46	0,4931
1,02	0,3461	1,43	0,4235	1,84	0,4671	2,48	0,4934
1,03	0,3485	1,44	0,4251	1,85	0,4678	2,50	0,4938
1,04	0,3508	1,45	0,4265	1,86	0,4686	2,52	0,4941
1,05	0,3531	1,46	0,4279	1,87	0,4693	2,54	0,4945
1,06	0,3554	1,47	0,4292	1,83	0,4699	2,50	0,4948
1,07	0,3577	1,48	0,4306	1,89	0,4706	2,58	0,4951
1,08	0,3599	1,49	0,4319	1,90	0,4713	2,60	0,4953
1,09	0,3621	1,50	0,4332	1,91	0,4719	2,62	0,4956
1,10	0,3643	1,51	0,4345	1,92	0,4726	2,64	0,4959
1,11	0,3665	1,52	0,4357	1,93	0,4732	2,66	0,4961
1,12	0,3686	1,53	0,4370	1,94	0,4738	2,68	0,4963
1,13	0,3708	1,54	0,4382	1,95	0,4744	2,70	0,4965
1,14	0,3729	1,55	0,4394	1,96	0,4750	2,72	0,4967
1,15	0,3749	1,56	0,4406	1,97	0,4756	2,74	0,4969
1,16	0,3770	1,57	0,4418	1,98	0,4761	2,76	0,4971
1,17	0,3790	1,58	0,4429	1,99	0,4767	2,78	0,4973
1,18	0,3810	1,59	0,4441	2,00	0,4772	2,80	0,4974
1,19	0,3830	1,60	0,4452	2,02	0,4783	2,82	0,4976
1,20	0,3849	1,61	0,4463	2,04	0,4793	2,84	0,4977
1,21	0,3869	1,62	0,4474	2,06	0,4803	2,86	0,4979
1,22	0,3883	1,63	0,4484	2,08	0,4812	2,88	0,4980
1,23	0,3907	1,64	0,4495	2,10	0,4821	2,90	0,4981
1,24	0,3925	1,65	0,4505	2,12	0,4830	2,92	0,4982
1,25	0,3944	1,60	0,4515	2,14	0,4838	2,94	0,4984
1,26	0,3962	1,67	0,4525	2,16	0,4846	2,96	0,4985
1,27	0,3980	1,68	0,4535	2,18	0,4854	2,98	0,4986
1,28	0,3997	1,69	0,4545	2,20	0,4861	3,00	0,49865
1,29	0,4015	1,70	0,4554	2,22	0,4868	3,20	0,49931
1,30	0,4032	1,71	0,4564	2,24	0,4875	3,40	0,49966
1,31	0,4049	1,72	0,4573	2,26	0,4881	3,60	0,499841
1,32	0,4066	1,73	0,4582	2,28	0,4887	3,80	0,499928
1,33	0,4082	1,74	0,4591	2,30	0,4893	4,00	0,499968
1,34	0,4099	1,75	0,4599	2,32	0,4898	4,50	0,499997
1,35	0,4115	1,76	0,4608	2,34	0,4904	5,00	0,499997
1,36	0,4131	1,77	0,4616				

Критические точки распределений Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29

Критические точки распределения χ^2

Число степеней свободы k	Уровень значимости α				
	0,01	0,025	0,05	0,10	0,20
1	6,6	5,0	3,8	2,706	1,642
2	9,2	7,4	6,0	4,605	3,219
3	11,3	9,4	7,8	6,251	4,642
4	13,3	11,1	9,5	7,779	5,989
5	15,1	12,8	11,1	9,236	7,289
6	16,8	14,4	12,6	10,645	8,558
7	18,5	16,0	14,1	12,017	9,803
8	20,1	17,5	15,5	13,362	11,030
9	21,7	19,0	16,9	14,684	12,242
10	23,2	20,5	18,3	15,987	13,442
11	24,7	21,9	19,7	17,275	14,631
12	26,2	23,3	21,0	18,549	15,812
13	27,7	24,7	22,4	19,812	16,985
14	29,1	26,1	23,7	21,064	18,151
15	30,6	27,5	25,0	22,307	19,311
16	32,0	28,8	26,3	23,542	20,465
17	33,4	30,2	27,6	24,769	21,615
18	34,8	31,5	28,9	25,989	22,760
19	36,2	32,9	30,1	27,204	23,900
20	37,6	34,2	31,4	28,412	25,038
21	38,9	35,5	32,7	29,615	26,171
22	40,3	36,8	33,9	30,813	27,301
23	41,6	38,1	35,2	32,007	28,429
24	43,0	39,4	36,4	33,196	29,553
25	44,3	40,6	37,7	34,382	30,675
26	45,6	41,9	38,9	35,563	31,795
27	47,0	43,2	40,1	36,741	32,912
28	48,3	44,5	41,3	37,916	34,027
29	49,6	45,7	42,6	39,087	35,139
30	50,9	47,0	43,8	40,256	36,250

Распределения Фишера

Критические точки:

 k_1 – число степеней свободы большей дисперсии; k_2 – число степеней свободы меньшей дисперсии.

$k_2 \backslash k_1$	1	2	3	4	5	6	8	12	24	∞
1	161,45	99,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,512	18,999	19,163	19,248	19,298	19,329	19,371	19,414	19,453	19,496
3	10,129	9,552	9,276	9,118	9,014	8,941	8,844	8,744	8,638	8,527
4	7,710	6,945	6,591	6,388	6,257	6,164	6,041	5,912	5,774	5,628
5	6,607	5,786	5,410	5,192	5,050	4,950	4,818	4,678	4,527	4,365
6	5,9871	5,143	4,756	4,534,	4,388	4,284	4,147	4,000	3,841	3,669
7	5,591	4,737	4,347	4,121	3,972	3,866	3,725	3,574	3,410	3,230
8	5,317	4,459	4,067	3,838	3,688	3,580	3,438	3,284	3,116	2,928
9	5,117	4,256	3,863	3,633	3,482	3,374	3,230	3,073	2,900	2,707
10	4,965	4,103	3,708	3,478	3,326	3,217	3,072	2,913	2,737	2,538
11	4,844	3,982	3,587	3,3571	3,204	3,094	2,948	2,788	2,609	2,405
12	4,747	3,885	3,490	3,259	3,106	2,999	2,848	2,686	2,505	2,296
13	4,66	3,805	3,410	3,179	3,025	2,915	2,767	2,604	2,420	2,207
14	4,600	3,739	3,344	3,112	2,958	2,848	2,699	2,534	2,349	2,131
15	4,543	3,683	3,287	3,056	2,901	2,790	2,641	2,475	2,288	2,066
16	4,494	3,634	3,239	3,0071	2,853	2,7411	2,591	2,424	2,2351	2,010
17	4,45	3,592	3,197	2,965	2,810	2,699	2,548	2,381	2,190	1,961
18	4,414	3,555	3,160	2,928	2,773	2,661	2,510	2,342	2,150	1,917
19	4,381	3,522	3,127	2,895	2,740	2,629	2,477	2,308	2,114	1,878
20	4,351	3,493	3,098	2,866	2,711	2,599	2,447	2,278	2,083	1,843
21	4,325	3,467	3,072	2,840	2,6851	2,573	2,421	2,2501	2,054	1,812
22	4,301	3,443	3,049	2,817	2,661	2,549	2,397	2,226	2,028	1,783
23	4,279	3,422	3,028	2,795	2,640	2,528	2,375	2,203	2,005	1,757
24	4,260	3,403	3,009	2,777	2,621	2,508	2,355	2,183	1,984	1,733
25	4,242	3,385	2,991	2,759	2,603	2,490	2,337	2,165	1,965	1,711
26	4,225	3,369	2,975	2,743	2,5871	2,474	2,321	2,148	1,947	1,691
27	4,210	3,354	2,961	2,728	2,572	2,459	2,305	2,132	1,930	1,672
28	4,196	3,340	2,947	2,714	2,558	2,445	2,292	2,118	1,915	1,654
29	4,183	3,328	2,934	2,702	2,545	2,432	2,278	2,104	1,901	1,638
30	4,171	3,316	2,922	2,690	2,534	2,421	2,266	2,092	1,887	1,622
40	4,085	3,232	2,839	2,608	2,449	2,336	2,180	2,004	1,793	1,509
60	4,001	3,151	2,758	2,523	2,368	2,254	2,097	1,918	1,700	1,389
120	3,920	3,072	2,680	2,447	2,290	2,175	2,106	1,834	1,608	1,254
∞	3,841	2,996	2,605	2,372	2,214	2,098	1,938	1,752	1,517	1,000

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие для вузов. М.: Высш. шк., 2001. — 479 с.
2. Герасимович А.И. Математическая статистика: [Учебное пособие для инженерно-технических и экономических специальностей ВТУЗов]. — Мн.: Выш. школа, 1983. — 279 с.
3. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и её инженерные приложения. — М.: Наука, 1988. — 480 с.
4. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. — М.: Наука, 1965. — 512 с.
5. Математические задачи сельской электрификации. Синьков В.М., Пересыпкина С.И., Филиппов Н.М. — Киев: Вища школа, 1978. — 288с.

Учебное издание

КОЗЛОВСКАЯ Влада Борисовна

МАТЕМАТИЧЕСКИЕ ЗАДАЧИ ЭНЕРГЕТИКИ

Учебно- методическое пособие
к практическим занятиям и
курсовому проектированию

Редактор

Корректор

Компьютерная верстка

Подписано в печать

Формат 60x84 1/16. Бумага типографическая №2.

Печать офсетная. Гарнитура Таймс.

Усл.печ.

Уч.-изд.л.

Тираж

Заказ

Издатель и полиграфическое исполнение:

Белорусский национальный технический университет

Лицензия ЛВ №155 от 30.01.2003. 220013, Минск, проспект Ф.Скорины, 65.