



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНО УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ПРИБОРОСТРОЕНИЯ И ИНФОРМАТИКИ (МИРЭА)

Михин М.Н.

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

- Элементы математической статистики
- Статистическое оценивание
- Проверка статистических гипотез

Москва – **2016**

Оглавление

Введение	3
Глава 1. Элементы математической статистики	4
§ 1. Основные понятия и основные задачи математической статистики.....	4
§ 2. Простейшие статистические преобразования	5
§ 3. Эмпирическая функция распределения.....	9
§ 4. Полигон и гистограмма.....	11
Глава 2. Статистическое оценивание	14
§ 1. Точечные оценки. Выборочная средняя и выборочная дисперсия	14
§ 2. Метод моментов	16
§ 3. Метод максимального правдоподобия	19
§ 4. Интервальные оценки (доверительные интервалы).....	21
§ 5. Доверительный интервал для математического ожидания генеральной совокупности, имеющей нормальное распределение при известной дисперсии.....	23
§ 6. Доверительный интервал для математического ожидания генеральной совокупности, имеющей нормальное распределение при неизвестной дисперсии...	25
§ 7. Доверительный интервал для дисперсии генеральной совокупности, имеющей нормальное распределение при известном математическом ожидании.	27
§ 8. Доверительный интервал для дисперсии генеральной совокупности, имеющей нормальное распределение при неизвестном математическом ожидании	28
Глава 3. Проверка статистических гипотез.....	32
§ 1. Основные понятия	32
§ 2. Проверка гипотезы о значении математического ожидания	34
§ 3. Проверка гипотезы о равенстве математических ожиданий двух генеральных совокупностей	36
§ 4. Проверка гипотезы о значении дисперсии генеральной совокупности.....	40
§ 5. Проверка гипотезы о равенстве дисперсий двух генеральных совокупностей	42
§ 6. Проверка гипотезы о распределении. Критерий Пирсона	45
Приложения	55
Используемая литература	60

Введение

Цель данного пособия помочь студенту самостоятельно подготовиться к выполнению контрольных работ. Перед каждой рассматриваемой задачей дается тот теоретический материал, который необходим для ее решения. Если студент ранее овладел необходимым теоретическим материалом, то вводную часть каждой задачи он может опустить и перейти непосредственно к решению задачи. Необходимо отметить, что приведенный теоретический материал достаточно полно охватывает курс указанного предмета.

Глава 1. Элементы математической статистики

§ 1. Основные понятия и основные задачи математической статистики

В математической статистике исследуются способы получения выводов на основе эмпирических (опытных) данных. Основными понятиями математической статистики являются: генеральная совокупность, выборка, теоретическая функция распределения.

Определение. Генеральной совокупностью называются все возможные результаты наблюдений, которые могут быть сделаны при данном комплексе условий.

В некоторых задачах генеральную совокупность рассматривают как случайную величину X . Примером генеральной совокупности может быть все население страны. В этой совокупности нас могут интересовать, например, возраст жителей. Другим примером генеральной совокупности являются детали, изготовленные на данном станке. Эти детали могут быть качественными и бракованными.

Определение. Выборочной совокупностью (выборкой) называется множество результатов, случайно отобранных из генеральной совокупности.

Выборка должна быть репрезентативной, т.е. правильно отражать пропорции генеральной совокупности. Это достигается случайностью отбора, когда все объекты генеральной совокупности имеют одинаковую вероятность быть отобранными.

Определение. Число объектов в совокупности (генеральной или выборочной) называется ее **объемом**.

Объем генеральной совокупности обозначим символом N , а объем выборочной совокупности обозначим символом n . При этом подразумевается, что $n \ll N$.

Заметим, что сам процесс выбора можно осуществлять разными способами: выбрав объект и определив его значение, изымать объект и не допускать к последующим испытаниям (выборка без возвращения); после определения его значения объект возвращается в генеральную совокупность (выборка с возвращением). Очевидно, что при достаточно большом объеме генеральной совокупности исчезает различие между выборками с возвращением и без возвращения. Будем рассматривать случай бесконечно большого объема генеральной совокупности.

Основные задачи математической статистики:

1. Оценка значения неизвестной вероятности случайного события;
2. Определение неизвестной теоретической функции распределения;
3. Определение неизвестных параметров распределения теоретической функции распределения;
4. Проверка статистических гипотез;
5. Оценка зависимости.

§2. Простейшие статистические преобразования

Для обоснованных статистических выводов необходимо иметь выборку достаточно большого объема n . Очевидно, что использование и хранение такой выборки весьма затруднительно. Чтобы избавиться от данных проблем, используют понятие статистики.

Определение. Статистикой $S = (S_1, S_2, \dots, S_n)$ называется произвольная k -мерная функция от выборки $X_1, X_2, \dots, X_n$:

$$S_1 = S_1(X_1, X_2, \dots, X_n), \dots, S_k = S_k(X_1, X_2, \dots, X_n).$$

Как функция от случайного вектора $(X_1, X_2, \dots, X_n)$ статистика $S = (S_1, S_2, \dots, S_n)$ также будет случайным вектором.

Определение. Вариационным рядом $X_1^*, X_2^*, \dots, X_n^*$ является выборка $X_1, X_2, \dots, X_n$, элементы которой расположены в порядке возрастания элементов: $X_1^* \leq X_2^* \leq \dots \leq X_n^*$.

Очевидно, что данное преобразование не приводит к потере информации относительно теоретической функции распределения.

Для величин X_1^* и X_n^* употребляют название «крайние члены вариационного ряда».

Определение. Размахом варьирования называется разность между крайними членами вариационного ряда, т.е.

$$R = X_n^* - X_1^*. \quad (1.2.1)$$

Если среди элементов выборки $X_1, X_2, \dots, X_n$ имеются одинаковые, что происходит при наблюдении дискретной случайной величины, то целесообразно произвести группировку данных.

Определение. Значение выборки $X_1, X_2, \dots, X_n$, соответствующее отдельной группе сгруппированного ряда наблюдаемых данных, называется **вариантой**, а численность отдельной группы сгруппированного ряда наблюдаемых данных, называется **частотой** или **весом** варианты.

Если i — индекс варианты, то n_i — число значений i -ой варианты.

Определение. Отношение частоты n_i к общей сумме всех частот $\sum n_i = n$ называется **относительной частотой** варианты и обозначается $p_i^* = \frac{n_i}{n}$.

Определение. Статистическим рядом называется расположенная по возрастанию совокупность различных вариантов Z_i , представляющих выборку

$X_1, X_2, \dots, X_n$, с соответствующими им частотами или относительными частотами.

При наблюдении непрерывной случайной величины используют интервальный ряд. В этом случае весь возможный интервал, которому принадлежат значения выборки, разбивают на конечное число частичных интервалов и подсчитывают частоту попадания элементов выборки в каждый частичный интервал.

Определение. Интервальным рядом называется упорядоченная последовательность интервалов с соответствующими им частотами или относительными частотами попадания элементов выборки в каждый из этих интервалов.

Пример 1. В городе А для определения сроков гарантированного обслуживания проведено исследование величины среднего пробега автомобилей, находящихся в эксплуатации в течение двух лет. Получены следующие результаты (тыс. км.):

3,0; 25,0; 18,6; 12,1; 10,6; 18,0; 17,3; 29,1; 20,0; 18,3; 21,5; 26,7; 12,2; 14,4; 7,3; 9,1; 2,9; 5,4; 40,1; 16,8; 11,2; 9,9; 25,3; 4,2; 29,6.

Составить интервальный ряд.

Решение. Очевидно, что величина среднего пробега автомобилей, находящихся в эксплуатации в течение двух лет, является непрерывной случайной величиной. Полученные данные представляют собой выборку из $n = 25$ наблюдений. Найдем сначала минимальное и максимальное значения случайной величины (т.е. крайние члены вариационного ряда):

$$X_{\min} = 2,9 \text{ и } X_{\max} = 40,1.$$

Размах варьирования будет равен

$$R = X_n^* - X_1^* = 40,1 - 2,9 = 37,2.$$

Возьмем число частичных интервалов $l = 6$. В этом случае длина частичного интервала равна

$$h = \frac{R}{n} = \frac{X_n^* - X_1^*}{n} = \frac{37,2}{6} = 6,2.$$

Соответствующий интервальный ряд приведен в таблице 1.1.

Таблица 1.1

Номер интервала i	Средний пробег автомобилей (интервалы) $x_i < X \leq x_{i+1}$	Частота n_i	Относительная частота $p_i^* = \frac{n_i}{n}$
1	2,6 — 9,1	6	0,24
2	9,1 — 15,3	6	0,24
3	15,3 — 21,5	7	0,28
4	21,5 — 27,7	3	0,12
5	27,7 — 33,9	2	0,08
6	33,9 — 40,1	1	0,04

Пример 2. Наблюдается число выигрышей в мгновенной лотерее. В результате наблюдения получены следующие значения выигрышей (тыс. руб.):

0, 1, 0, 0, 5, 0, 10, 0, 1, 0, 0, 1, 5, 1, 0, 0, 0, 1, 0,
1, 0, 0, 0, 0, 5, 0, 5, 0, 0, 1, 1, 1, 5, 10, 0, 1, 1, 0,
5, 0, 0, 0, 0, 1, 0, 1, 0, 5, 0, 0, 0, 0, 1, 0.

Составить статистический ряд случайной величины X — выигрыша в мгновенной лотерее.

О Решение. Случайная величина X принимает 4 различных значения: 0, 1, 5 и 10. Для каждого значения подсчитаем частоту и относительную частоту. Результаты задачи представим в таблице 1.2.

Таблица 1.2

№	i	1	2	3	4
Выигрыш в мгновенной лотерее	Z_i	0	1	5	10
Частота	n_i	31	14	7	2
Относительная частота	p_i^*	31/54	7/27	7/54	1/27

§ 3. Эмпирическая функция распределения

По статистическом ряду, приведенному в таблице 7.3,

Таблица 7.3

№	1	2	...	L
Варианты Z_i	Z_1	Z_2	...	Z_L
Относительная частота p_i^*	$p_1^* = \frac{n_1}{n}$	$p_2^* = \frac{n_2}{n}$	...	$p_L^* = \frac{n_L}{n}$

можно построить эмпирическую (выборочную) функцию распределения.

Определение. Эмпирической (выборочной) функцией распределения называется функция $F^*(x)$, задающая для каждого значения x относительную частоту события $X < x$.

Следовательно, по определению

$$F^*(x) = \frac{n_x}{n}, \quad (1.3.1)$$

где n_x — число элементов выборки, значения которых меньше x .

Очевидно, что для нахождения функции распределения можно использовать формулу

$$F^*(x) = \sum_{Z_i < x} p_i^*. \quad (1.3.2)$$

Эмпирическую функцию распределения можно задать таблично или графически. Построим эмпирическую функцию распределения по данным, приведенным в таблице 7.2.

Объем выборки по условию примера $n = 54$.

Наименьшая варианта равна 0, следовательно, $n_x = 0$ при $x \leq 0$. Тогда $F^*(x) = 0/54 = 0$ при $x \leq 0$.

Если $0 < x \leq 1$, то неравенство $X < x$ выполняется для варианты $z_1 = 0$, которая встречается 31 раз, поэтому $n_x = 31$ и $F^*(x) = 31/54$.

Если $1 < x \leq 5$, то неравенство $X < x$ выполняется для вариантов $z_1 = 0$ и $z_2 = 1$, которые встречаются 31 и 14 раз соответственно, поэтому, $n_x = 31 + 14 = 45$, $F^*(x) = 45/54$ и т.д.

Результаты вычисления $F^*(x)$ приведем в таблице 7.4

Таблица 7.4

x	$F^*(x)$
$x \leq 0$	0
$0 < x \leq 1$	$F^*(x) = p_1^* = \frac{31}{54}$
$1 < x \leq 5$	$F^*(x) = p_1^* + p_2^* = \frac{31+14}{54} = \frac{45}{54}$
$5 < x \leq 10$	$F^*(x) = p_1^* + p_2^* + p_3^* = \frac{31+14+7}{54} = \frac{52}{54}$
$x > 10$	$F^*(x) = p_1^* + p_2^* + p_3^* + p_4^* = \frac{31+14+7+2}{54} = 1$

График этой функции приведен на рис. 1.

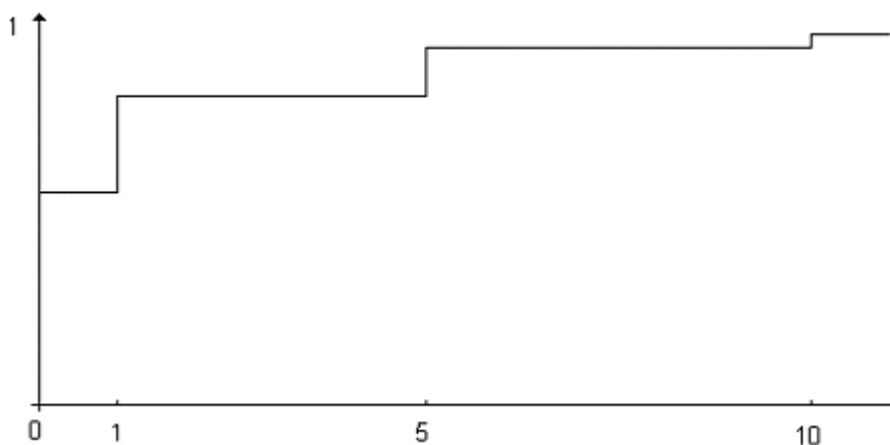


Рис. 1

В случае интервального ряда значения эмпирической функции $F^*(x)$ подсчитывают на концах частичных интервалов.

Эмпирическая функция $F^*(x)$ применяется для оценивания теоретической функции распределения генеральной совокупности.

§ 4. Полигон и гистограмма

Определение. Полигоном частот (многоугольником распределения) называется ломаная линия, проходящая через точки с координатами (Z_i, n_i) , где Z_i — варианты статистического ряда, а n_i — соответствующие им частоты.

Если ломаная линия строится по точкам (Z_i, p_i^*) , где p_i^* — относительные частоты, то получаем **полигон относительных частот**.

Построим полигон относительных частот для выборки из примера 2. Используя статистический ряд, представленный в таблице 7.2, получаем полигон относительных частот, изображенный на рис. 2.

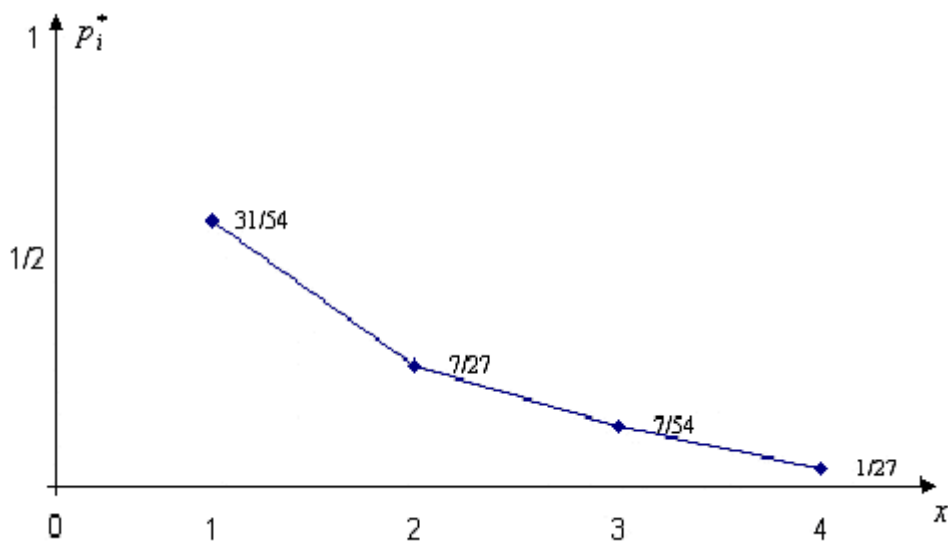


Рис. 2

В случае непрерывной случайной величины выборку преобразуют следующим образом. Всю ось абсцисс делят на интервалы (x_i, x_{i+1}) длины $\Delta_i = x_{i+1} - x_i$ и определяют функцию $f^*(x)$, которая на i -м интервале принимает значение

$$f^*(x) = \frac{p^*(x)}{\Delta_i} = \frac{n_i}{n\Delta_i}, \quad (1.4.1)$$

где n_i — число элементов выборки, попавших в интервал.

Определение. Функция $f^*(x)$, определенная соотношением (1.4.1), называется **гистограммой**.

Гистограмма является выборочной оценкой плотности вероятности.

Построим гистограмму по данным, приведенным в примере 1. Длина каждого интервала равна $\Delta_i = 6,2$. Подсчитаем значения $\frac{n_i}{n\Delta_i}$:

$x_i < X \leq x_{i+1}$	2,6 — 9,1	9,1 — 15,3	15,3—21,5	21,5—27,7	27,7—33,9	33,9—40,1
$\frac{n_i}{n\Delta_i}$	0,038710	0,038710	0,045161	0,019355	0,012903	0,006452

На рис. 3 представлена гистограмма примера 1.

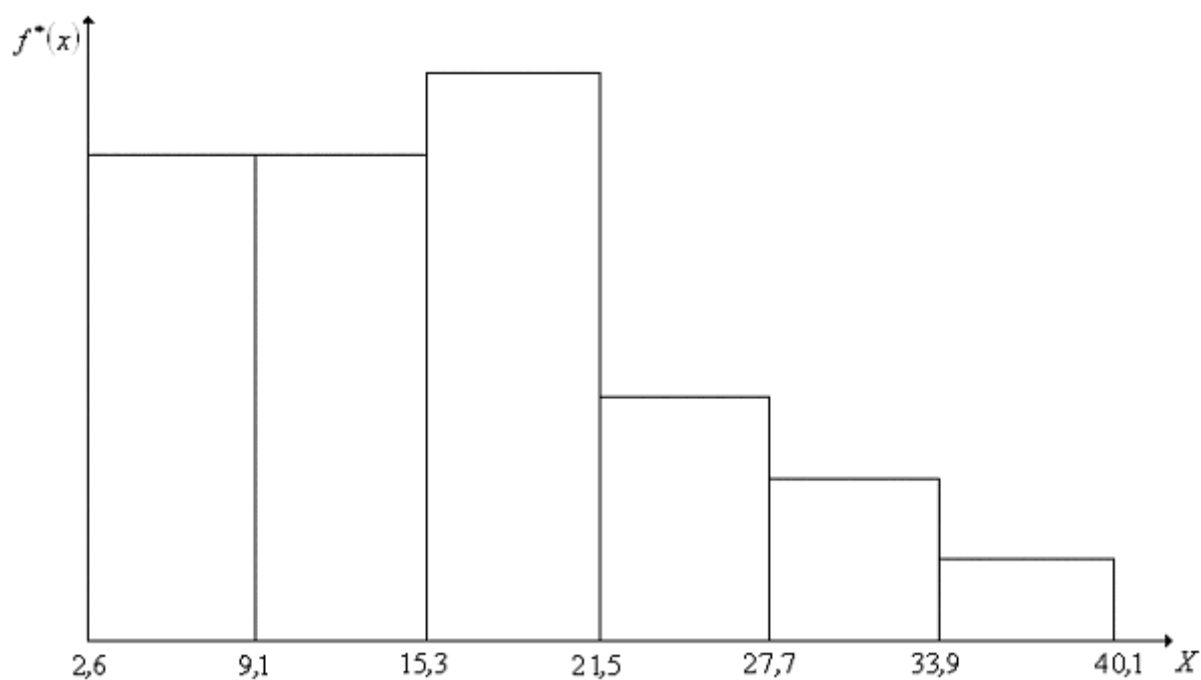


Рис. 3

Графическое изображение статистических рядов в виде полигона и гистограммы позволяет получить первоначальное представление о закономерностях, имеющих место в совокупности наблюдений.

Глава 2. Статистическое оценивание

§ 1. Точечные оценки. Выборочная средняя и выборочная дисперсия

Пусть θ — некоторый параметр закона распределения генеральной совокупности.

Определение. Точечной оценкой θ_n^* параметра θ называется произвольная функция $\theta_n^*(X_1, \dots, X_n)$ случайной выборки $X_1, \dots, X_n$.

Статическая оценка является случайной величиной и меняется в зависимости от выборки. Статическая оценка, как всякая случайная величина, имеет функцию распределения $F_{\theta^*}(x)$.

Определение. Оценка θ_n^* называется **несмещенной**, если ее математическое ожидание равно оцениваемому параметру, т.е.

$$M(\theta_n^*) = \theta. \quad (2.1.1)$$

Если оценка не является несмещенной, то она будет либо завышать значение θ , либо занижать его. В обоих случаях это приводит к систематическим ошибкам одного знака в оценке параметра θ .

Определение. Оценка θ_n^* называется **состоятельной**, если с ростом объема выборки она сходится к оцениваемому параметру. Можно рассматривать сходимость различных типов: по вероятности, с вероятностью равной единице, в среднем квадратичном и т.д. Как правило, рассматривается сходимость по вероятности, т.е. состоятельной называется оценка θ_n^* , которая для каждого $\varepsilon > 0$ при всех возможных значениях неизвестного параметра θ удовлетворяет соотношению

$$\lim_{n \rightarrow \infty} P\{|\theta_n^* - \theta| > \varepsilon\} = 0. \quad (2.1.2)$$

Состоятельность оценки обосновывает увеличение объема случайной выборки, так как при этом становится менее вероятной возможность большой ошибки в оценке параметра θ .

Определение. Несмещенная оценка θ_n^* называется **эффективной**, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра θ , вычисленных по случайным выборкам одного и того же объема.

Замечание. В дальнейшем вместо обозначения θ_n^* будем использовать θ^* .

Определение. Выборочной средней называется среднее арифметическое полученных по выборке значений

$$m^* = \frac{1}{n} \sum_{i=1}^n X_i. \quad (2.1.3)$$

Определение. Вторым выборочным моментом называется среднее арифметическое квадратов, полученных по выборке значений

$$m_2^* = \frac{1}{n} \sum_{i=1}^n X_i^2. \quad (2.1.4)$$

Определение. Выборочной дисперсией называется среднее арифметическое квадратов отклонений значений в случайной выборке от выборочной средней

$$\sigma^{2*} = \frac{1}{n} \sum_{i=1}^n (X_i - m^*)^2 = m_2^* - (m^*)^2. \quad (2.1.5)$$

Определение. Исправленной выборочной дисперсией называется произведение выборочной дисперсии на величину $\frac{n}{n-1}$, т.е.

$$S^{2*} = \frac{n}{n-1} \sigma^{2*} = \frac{1}{n-1} \sum_{i=1}^n (X_i - m^*)^2. \quad (2.1.6)$$

Пример 1. Проверить, является ли второй выборочный момент m_2^* несмещенной оценкой второго теоретического момента.

О Решение. Найдем математическое ожидание оценки m_2^* .

$$M(m_2^*) = \frac{1}{n} M(X_1^2 + \dots + X_n^2) = \frac{1}{n} (MX_1^2 + \dots + MX_n^2) = \frac{1}{n} \sum_{i=1}^n m_2 = m_2. \bullet$$

Пример 2. Проверить, является ли оценка σ^{2*} несмещенной.

О Решение.

$$\begin{aligned} M\sigma^{2*} &= M\left(\frac{1}{n} \sum_{i=1}^n (X_i - m^*)^2\right) = M\left(\frac{1}{n} \sum_{i=1}^n (X_i^2 - 2m^* X_i + (m^*)^2)\right) = \\ &= m_2 - \frac{2}{n} \sum_{i=1}^n M(m^* X_i) + M(m^*)^2 = m_2 - \frac{2}{n^2} \sum_{i=1}^n \sum_{j=1}^n M(X_i X_j) + \\ &+ \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n M(X_i X_j) = m_2 - \frac{1}{n^2} \sum_{i=1}^n M X_i^2 - \frac{1}{n^2} \sum_{\substack{i,j \\ i \neq j}} M(X_i X_j) = \\ &= m_2 - \frac{1}{n} m_2 - \frac{n-1}{n} m^2 = \frac{n-1}{n} \sigma^2 \neq \sigma^2. \end{aligned}$$

Таким образом, σ^{2*} является смещенной оценкой дисперсии σ^2 . Очевидно, что S^{2*} будет уже несмещенной оценкой дисперсии σ^2 . $\bullet$

§ 2. Метод моментов

Пусть имеется выборка $X_1, \dots, X_n$, произведенная из генеральной совокупности с теоретической функцией распределения $F(x)$, зависящей от k параметров $\theta_1, \dots, \theta_k$, которые нужно оценить. Зная функцию распределения, можно найти первые k теоретических моментов, которые будут зависеть от параметров $\theta_1, \dots, \theta_k$:

$$\begin{cases} m_1 = MX = m_1(\theta_1, \dots, \theta_k), \\ m_2 = MX^2 = m_2(\theta_1, \dots, \theta_k), \\ \dots \\ m_k = MX^k = m_k(\theta_1, \dots, \theta_k), \end{cases} \quad (2.2.1)$$

где X — случайная величина, имеющая функцию распределения $F(x)$.

Метод моментов состоит в том, что в системе (2.2.1) при большом объёме выборки n теоретические моменты $m_1, \dots, m_k$ заменяются на выборочные $m_1^*, \dots, m_k^*$, а затем, решая эту систему относительно $\theta_1, \dots, \theta_k$, находят оценки неизвестных параметров. Таким образом, в методе моментов оценки $\theta_1^*, \dots, \theta_k^*$ неизвестных параметров $\theta_1, \dots, \theta_k$ определяются из системы уравнений

$$\begin{cases} m_1^* = m_1(\theta_1^*, \dots, \theta_k^*), \\ m_2^* = m_2(\theta_1^*, \dots, \theta_k^*), \\ \dots \\ m_k^* = m_k(\theta_1^*, \dots, \theta_k^*). \end{cases} \quad (2.2.2)$$

Метод моментов был предложен в 1894 г. *К. Пирсоном*. Оценки, полученные методом моментов, как правило, являются состоятельными.

Пример 3. Выборка $X_1, \dots, X_n$ произведена из генеральной совокупности с теоретической функцией распределения, имеющей плотность показательного закона

$$p(x) = p(x, \theta) = \theta e^{-\theta x} \quad (x \geq 0).$$

Найти оценку параметра θ .

О Решение. Математическое ожидание случайной величины X , имеющей плотность показательного закона, задаётся формулой

$$m_1 = MX = \int_0^{\infty} xp(x)dx = \int_0^{\infty} x\theta e^{-\theta x} dx = \frac{1}{\theta}.$$

Используя систему (2.2.2), получаем $m_1^* = \frac{1}{\theta^*}$. Откуда окончательно получаем

$$\theta^* = \frac{1}{m^*}. \bullet$$

§ 3. Метод максимального правдоподобия

Метод максимального правдоподобия опирается на использование условий экстремума функций одной или нескольких случайных величин. В качестве такой функции используется функция правдоподобия.

Определение. Функцией правдоподобия называется функция

$$L(X_1, \dots, X_n) = L(X_1, \dots, X_n; \theta) = P(X_1; \theta) \cdots P(X_n; \theta) \quad (2.3.1)$$

в дискретном случае и

$$L(X_1, \dots, X_n) = L(X_1, \dots, X_n; \theta) = p(X_1; \theta) \cdots p(X_n; \theta) \quad (2.3.2)$$

в непрерывном случае.

В функции правдоподобия $L(X_1, \dots, X_n; \theta)$ элементы выборки $X_1, \dots, X_n$ являются фиксированными параметрами, а θ — аргументом.

Определение. Оценкой максимального правдоподобия называется такое θ^* , для которого

$$L(X_1, \dots, X_n; \theta^*) = \max_{\theta} L(X_1, \dots, X_n; \theta). \quad (2.3.3)$$

Поскольку L и $\ln(L)$ принимают максимум при одном и том же значении аргумента θ , то удобно использовать не саму функцию правдоподобия, а ее логарифм.

Определение. Уравнением правдоподобия называется уравнение

$$\frac{\partial}{\partial \theta} \ln(L(X_1, \dots, X_n; \theta)) = 0. \quad (2.3.4)$$

В случае, когда теоретическая функция распределения $F(X_1, \dots, X_n; \theta_1, \dots, \theta_k)$ зависит от нескольких параметров $\theta_1, \dots, \theta_k$, при применении метода максимального правдоподобия вместо уравнения (2.3.4) необходимо использовать систему уравнений

$$\begin{cases} \frac{\partial}{\partial \theta_1} \ln(L(X_1, \dots, X_n; \theta_1, \dots, \theta_k)) = 0, \\ \dots \\ \frac{\partial}{\partial \theta_k} \ln(L(X_1, \dots, X_n; \theta_1, \dots, \theta_k)) = 0. \end{cases} \quad (2.3.5)$$

Пример 4. Найти оценки параметров m (математическое ожидание) и σ^2 (дисперсия) нормального закона распределения.

О Решение. Нормальный закон распределения задается плотностью

$$p(x, m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Функция правдоподобия примет вид:

$$L(X_1, \dots, X_n; m, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma}} \right)^n e^{-\frac{1}{2\sigma^2}((X_1-m)^2 + \dots + (X_n-m)^2)}. \quad (2.3.6)$$

Логарифмируя выражение (2.3.6), получим

$$\begin{aligned} \ln(L(X_1, \dots, X_n; m, \sigma^2)) = & -\frac{n}{2} \ln(2\pi\sigma^2) - \\ & - \frac{1}{2\sigma^2} ((X_1 - m)^2 + \dots + (X_n - m)^2). \end{aligned} \quad (2.3.7)$$

Подставляя (2.3.7) в систему (2.3.5), получим

$$\begin{cases} \frac{\partial}{\partial m} \ln(L(X_1, \dots, X_n; m, \sigma^2)) = \frac{1}{\sigma^2} \left(\sum_{i=1}^n X_i - nm \right) = 0, \\ \frac{\partial}{\partial \sigma^2} \ln(L(X_1, \dots, X_n; m, \sigma^2)) = \frac{1}{\sigma^2} \left(\sum_{i=1}^n (X_i - m)^2 - n\sigma^2 \right) = 0. \end{cases} \quad (2.3.8)$$

Из первого уравнения системы (2.3.8) получаем

$$m^* = \frac{1}{n} \sum_{i=1}^n X_i, \quad (2.3.9)$$

а из второго уравнения системы (2.3.8), с учетом (2.3.9), получаем

$$\sigma^{2*} = \frac{1}{n} \sum_{i=1}^n (X_i - m^*)^2. \bullet$$

§ 4. Интервальные оценки (доверительные интервалы)

Статистическая оценка θ^* некоторого параметра распределения θ наблюдаемой случайной величины X , будучи функцией случайной выборки (статистикой), сама является случайной величиной, имеющей свой закон распределения и числовые характеристики (параметры) распределения. При малом числе наблюдений могут возникнуть следующие задачи:

- к каким ошибкам может привести замена параметра θ его точечной оценкой θ^* ;
- с какой степенью надежности можно ожидать, что получаемые ошибки не выйдут за известные пределы.

Определение. Интервальной оценкой (доверительным интервалом) называется числовой интервал (θ_L^*, θ_R^*) , определяемый по результатам выборки, относительно которого можно утверждать с определенной, близкой к единице, вероятности, что он включает в себе значение оцениваемого параметра генеральной совокупности, т.е.

$$P(\theta_L^* \leq \theta \leq \theta_R^*) = \gamma, \quad (2.4.1)$$

где θ_L^* и θ_R^* — нижняя и верхняя (левая и правая) границы доверительного интервала параметра θ , γ — **доверительная вероятность**.

Доверительная вероятность и уровень значимости связаны соотношением

$$\alpha + \gamma = 1. \quad (2.4.2)$$

По заданной оценке θ^* и заданной доверительной вероятности γ доверительные интервалы можно построить различными способами. На практике обычно используются два типа доверительных интервалов:

- двусторонние;

- односторонние.

Ограничимся нахождением двусторонних доверительных интервалов. Односторонние доверительные интервалы находятся аналогично.

Для каждой доверительной вероятности γ можно указать такое значение Δ , что

$$P(|\theta^* - \theta| \leq \Delta) = P(\theta^* - \Delta \leq \theta \leq \theta^* + \Delta) = \gamma. \quad (2.4.3)$$

Определение. Величину Δ , равную половине ширины доверительного интервала, называют **точностью оценки**.

Определение. Квантилем уровня γ некоторого распределения называется такое число t_γ , при котором значение соответствующей функции распределения равно γ , т.е.

$$\gamma = F(t_\gamma). \quad (2.4.4)$$

Теорема. Пусть плотность случайной величины X симметрична относительно оси OY , и пусть $P(|X| < t) = \gamma$. Тогда t — квантиль уровня $\frac{\gamma+1}{2}$ распределения случайной величины X .

Доказательство.

$$\gamma = P(|X| < t) = 1 - P(|X| \geq t) = 1 - P(X \geq t) - P(X \leq -t) = 1 - 2P(X \geq t),$$

т.к. ввиду симметричности закона распределения относительно оси OY выполнено равенство

$$P(X \geq t) = P(X \leq -t).$$

Отсюда следует, что

$$P(|X| \geq t) = \frac{1-\gamma}{2} = 1 - P(X < t) \quad \text{и} \quad P(X < t) = \frac{1+\gamma}{2}. \quad \blacksquare$$

**§ 5. Доверительный интервал для математического ожидания
генеральной совокупности, имеющей нормальное
распределение при известной дисперсии**

Пусть из генеральной совокупности X , имеющей нормальный закон распределения при известной дисперсии σ^2 и неизвестном математическом ожидании m , произведена случайная выборка $X_1, \dots, X_n$. Для оценки математического ожидания используем статистику $m^* = \frac{1}{n} \sum_{i=1}^n X_i$, которая имеет нормальное распределение с параметрами $\left(m, \frac{\sigma^2}{n}\right)$. Тогда статистика $\frac{m^* - m}{\sigma} \sqrt{n}$ имеет нормальное распределение с параметрами $(0, 1)$. Найдем вероятность отклонения $|m^* - m|$:

$$\begin{aligned} \gamma = P\left(\left|\frac{m^* - m}{\sigma} \sqrt{n}\right| < t\right) &= P\left(|m^* - m| < t \frac{\sigma}{\sqrt{n}}\right) = \\ &= P\left(m^* - t \frac{\sigma}{\sqrt{n}} < m < m^* + t \frac{\sigma}{\sqrt{n}}\right). \end{aligned} \quad (2.4.5)$$

Интервал $\left(m^* - t \frac{\sigma}{\sqrt{n}}; m^* + t \frac{\sigma}{\sqrt{n}}\right)$, определенный по (2.4.5), представляет собой доверительный интервал для математического ожидания m при известной дисперсии σ^2 . Осталось указать, как, зная доверительную вероятность γ , выбрать в (2.4.5) значение t . Ответом на этот вопрос является доказанная ранее теорема, т.е. t является квантилем уровня $\frac{\gamma + 1}{2}$ нормального распределения с параметрами $(0, 1)$.

Таким образом, доверительный интервал для математического ожидания m при известной дисперсии σ^2 строится следующим образом:

1. Вычисляем оценку m^* , которая является средним арифметическим элементов выборки $X_1, \dots, X_n$.

2. Вычисляем квантиль $t_{\frac{\gamma+1}{2}}$ нормального распределения с параметрами $(0,1)$.

3. Строим доверительный интервал, который имеет вид:

$$\left(m^* - \frac{\sigma}{\sqrt{n}} t_{\frac{\gamma+1}{2}}; m^* + \frac{\sigma}{\sqrt{n}} t_{\frac{\gamma+1}{2}} \right). \quad (2.4.6)$$

Очевидно, что точность оценки равна

$$\Delta = t_{\frac{\gamma+1}{2}} \frac{\sigma}{\sqrt{n}}. \quad (2.4.7)$$

Замечание. Квантиль $t_{\frac{\gamma+1}{2}}$ нормального распределения с параметрами $(0,1)$ можно определить разными способами, например:

- используя таблицы функции $\Phi(t)$ нормального распределения с параметрами $(0,1)$ (Приложение 3);
- используя функцию *НОРМСТОБР(вероятность)* из EXCEL, при этом аргумент «вероятность» данной функции должен быть равен $\frac{\gamma+1}{2}$;
- используя функцию *qnorm* (v, m, σ) из MATHCAD, при этом переменные данной функции должны быть равны $v = \frac{\gamma+1}{2}$, $m = 0$, $\sigma = 1$.

**§ 6. Доверительный интервал для математического ожидания
генеральной совокупности, имеющей нормальное
распределение при неизвестной дисперсии**

Пусть из генеральной совокупности X , имеющей нормальный закон распределения при неизвестной дисперсии σ^2 и неизвестном математическом ожидании m , произведена случайная выборка $X_1, \dots, X_n$. Для оценки математического ожидания используем статистику

$$T = \frac{m^* - m}{S^*} \sqrt{n-1}, \quad (2.4.8)$$

имеющую t -распределение (распределение Стьюдента) с $\nu = n-1$ числом степеней свободы.

Доверительный интервал для математического ожидания при неизвестной дисперсии имеет вид:

$$\left(m^* - \frac{S^*}{\sqrt{n-1}} t_{\frac{\gamma+1}{2}}; m^* + \frac{S^*}{\sqrt{n-1}} t_{\frac{\gamma+1}{2}} \right). \quad (2.4.9)$$

Очевидно, что точность оценки равна

$$\Delta = t_{\frac{\gamma+1}{2}} \frac{S^*}{\sqrt{n-1}}. \quad (2.4.10)$$

Таким образом, доверительный интервал для математического ожидания m при неизвестной дисперсии σ^2 строится следующим образом:

1. Вычисляем оценку m^* , которая является средним арифметическим элементов выборки $X_1, \dots, X_n$, и исправленную выборочную дисперсию

$$S^{2*} = \frac{1}{n-1} \sum_{i=1}^n (X_i - m^*)^2.$$

2. Вычисляем квантиль $t_{\frac{\gamma+1}{2}}$ t -распределения (распределения Стьюдента) с $\nu = n - 1$ числом степеней свободы.
3. Используя формулу (2.4.9), получаем доверительный интервал.

Замечание. Квантиль $t_{\frac{\gamma+1}{2}}$ t -распределения (распределения Стьюдента)

с $l = n - 1$ числом степеней свободы можно определить разными способами, например:

- используя таблицы t -распределения (распределения Стьюдента) с $l = n - 1$ числом степеней;
- используя функцию *СТЮДРАСПОБР* (вероятность; степени_свободы) из EXCEL, при этом аргумент «вероятность» данной функции должен быть равен удвоенному уровню значимости, а аргумент «степени_свободы» должен быть равен $n - 1$;
- используя функцию *qt* (p , d) из MATHCAD, при этом переменные данной функции должны быть равны $p = \gamma$, $d = n - 1$.

Замечание. При достаточно большом объеме n выборки различия между доверительными интервалами, определенными по формулам (2.4.6) и (2.4.9), мало, т.к., при $n \rightarrow \infty$, распределение Стьюдента стремится к нормальному распределению.

§ 7. Доверительный интервал для дисперсии генеральной совокупности, имеющей нормальное распределение при известном математическом ожидании.

Предположим, что выборка $X_1, \dots, X_n$ произведена из нормальной генеральной совокупности с неизвестной дисперсией при известном математическом ожидании, равном m . В качестве оценки неизвестной дисперсии σ^2 возьмем выборочную дисперсию

$$\sigma^{2*} = \frac{1}{n} \left((X_1 - m)^2 + \dots + (X_n - m)^2 \right). \quad (2.4.11)$$

В этом случае статистика $\chi^2 = \frac{n\sigma^{2*}}{\sigma^2}$ будет иметь χ^2 -распределение с n степенями свободы.

Доверительный интервал для дисперсии при известном математическом ожидании имеет вид:

$$\frac{n\sigma^{2*}}{\frac{r_{1+\gamma}}{2}} < \sigma^2 < \frac{n\sigma^{2*}}{\frac{r_{1-\gamma}}{2}}, \quad (2.4.12)$$

где r_τ — квантиль уровня τ χ^2 -распределения с n степенями свободы.

Таким образом, доверительный интервал для дисперсии при известном математическом ожидании строится следующим образом:

1. Используя выборку $X_1, \dots, X_n$, по формуле (2.4.11) вычисляем оценку дисперсии.
2. Вычисляем квантили $\frac{r_{\gamma+1}}{2}$ и $\frac{r_{1-\gamma}}{2}$ χ^2 -распределения с n степенями свободы.
3. По формуле (2.4.12) получаем искомый доверительный интервал.

Замечание. Квантили $r_{\frac{\gamma+1}{2}}$ и $r_{\frac{1-\gamma}{2}}$ χ^2 -распределения с n степенями сво-

боды можно определить разными способами, например:

- используя таблицы χ^2 -распределения с n степенями свободы;
- используя функцию *ХИ2ОБР* (вероятность; степени_свободы) из EXCEL, при этом аргумент «вероятность» данной функции должен быть равен уровню значимости, а аргумент «степени_свободы» должен быть равен n ;
- используя функцию *qchisq* (p , d) из MATHCAD, при этом переменные данной функции должны быть равны: $p = \gamma$, $d = n$.

§ 8. Доверительный интервал для дисперсии генеральной совокупности, имеющей нормальное распределение при неизвестном математическом ожидании

В качестве оценки неизвестной дисперсии σ^2 возьмем выборочную дисперсию

$$S^{2*} = \frac{1}{n-1} \left((X_1 - m^*)^2 + \dots + (X_n - m^*)^2 \right). \quad (2.4.13)$$

В этом случае статистика $\chi^2 = \frac{nS^{2*}}{\sigma^2}$ будет иметь χ^2 -распределение с $n - 1$ степенью свободы.

Доверительный интервал для дисперсии при неизвестном математическом ожидании имеет вид:

$$\frac{(n-1)S^{2*}}{r_{\frac{1+\gamma}{2}}} < \sigma^2 < \frac{(n-1)S^{2*}}{r_{\frac{1-\gamma}{2}}}, \quad (2.4.14)$$

где r_τ — квантиль уровня τ χ^2 -распределения с $n - 1$ степенью свободы.

Таким образом, доверительный интервал для дисперсии при неизвестном математическом ожидании строится следующим образом:

1. Вычисляем оценку m^* , которая является средним арифметическим элементов выборки $X_1, \dots, X_n$, и по формуле (2.4.13) вычисляем выборочную оценку дисперсии.
2. Вычисляем квантили $r_{\frac{\gamma+1}{2}}$ и $r_{\frac{1-\gamma}{2}}$ χ^2 -распределения с $n - 1$ степенью свободы.
3. По формуле (2.4.14) получаем искомый доверительный интервал.

Пример 5. Найти доверительный интервал с доверительной вероятностью 0,95 для оценки математического ожидания, нормально распределенной случайной величины X , если известны ее дисперсия $\sigma^2 = 16$, выборочная средняя $m^* = 16$ и объем выборки $n = 16$.

О Решение. По условию задачи $\gamma = 0,95$. Найдем квантиль $t_{\frac{\gamma+1}{2}} = t_{\frac{0,95+1}{2}} = t_{0,975}$ нормального распределения с параметрами $(0,1)$, используя, например EXCEL: применяя функцию *НОРМСТОБР* (0,975), получим $t_{0,975} = 1,96$.

Применив (2.4.6), получим

$$m^* - \frac{\sigma}{\sqrt{n}} t_{\frac{\gamma+1}{2}} < m < m^* + \frac{\sigma}{\sqrt{n}} t_{\frac{\gamma+1}{2}},$$

$$16 - \frac{4}{\sqrt{16}} 1,96 < m < 16 + \frac{4}{\sqrt{16}} 1,96,$$

$$14,04 < m < 17,96. \bullet$$

Пример 6. Случайная величина X распределена по нормальному закону. Статистический ряд выборки представлен таблицей:

Z_i	3	5	7	8	10	12	14
n_i	3	7	4	6	7	5	8

Требуется:

- 1) построить доверительный интервал для оценки математического ожидания, если доверительная вероятность равна 0,97;
- 2) построить доверительный интервал для оценки дисперсии, если доверительная вероятность равна 0,95.

О Решение. Для оценки математического ожидания будем использовать формулу (2.4.9), т.к. дисперсия неизвестна. Вычислим все величины, присутствующие в (2.4.9).

Объем выборки $n = 3 + 7 + 4 + 6 + 7 + 5 + 8 = 40$.

Оценка математического ожидания:

$$m^* = \frac{1}{n} \sum_{i=1}^7 n_i Z_i = \frac{3 \cdot 3 + 7 \cdot 5 + 4 \cdot 7 + 6 \cdot 8 + 7 \cdot 10 + 5 \cdot 12 + 8 \cdot 14}{40} = \frac{362}{40} = 9,05.$$

Исправленная выборочная дисперсия:

$$S^{2*} = \frac{1}{n-1} \sum_{i=1}^7 n_i (Z_i - m^*)^2 = \frac{493,9}{39} = 12,664.$$

Корень квадратный исправленной выборочной дисперсии:

$$S^* = \sqrt{S^{2*}} = \sqrt{12,664} = 3,559.$$

Для доверительной вероятности $\gamma = 0,97$ найдем квантиль

$t_{\frac{\gamma+1}{2}} = t_{\frac{0,97+1}{2}} = t_{0,985}$ t -распределения) с $\nu = n - 1 = 40 - 1 = 39$ числом степеней

свободы. Для этого используем, например, функцию $qt(p, d)$ из MATHCAD:

$$qt(0.985, 39) = 2.252.$$

Доверительный интервал имеет вид:

$$9,05 - \frac{3,559}{\sqrt{39}} 2,252 < m < 9,05 + \frac{3,559}{\sqrt{39}} 2,252$$

и окончательно

$$7,77 < m < 10,33.$$

Перейдем ко второй части задачи. Для оценки дисперсии будем использовать формулу (2.4.14), т.к. математическое ожидание неизвестно. Для данной формулы необходимо вычислить квантили $r_{\frac{\gamma+1}{2}} = r_{\frac{0,95+1}{2}} = r_{0,975}$ и

$r_{\frac{1-\gamma}{2}} = r_{\frac{1-0,95}{2}} = r_{0,025}$. Используя функцию $qchisq(p, d)$ из MATHCAD, получим:

$$r_{0,975} = qchisq(0.975, 39) = 58.12 \text{ и } r_{0,025} = qchisq(0.025, 39) = 23.654.$$

Доверительный интервал имеет вид:

$$\frac{39 \cdot 12,664}{58,12} < \sigma^2 < \frac{39 \cdot 12,664}{23,654}$$

и окончательно

$$8,50 < \sigma^2 < 20,88. \bullet$$

Глава 3. Проверка статистических гипотез

§ 1. Основные понятия

Определение. Статистической гипотезой называется любое предположение H_0 (гипотеза) о виде закона распределения генеральной совокупности или о числовых значениях параметров закона распределения.

Определение. Правило, по которому гипотеза H_0 принимается или отвергается, называется **статистическим критерием**.

Проверяемую гипотезу H_0 называют нулевой, а противоположную ей гипотезу H_1 называют альтернативной.

Схема проверки нулевой гипотезы:

1. Используя проверочные данные $X_1, \dots, X_n$ и учитывая условия задачи, принимают нулевую гипотезу H_0 и альтернативную гипотезу H_1 .
2. По случайной выборке $X_1, \dots, X_n$ определяется функция $S(X_1, \dots, X_n)$, называемая статистикой, для которой будет известен точный или приближённый закон распределения.
3. По заранее выбранной малой вероятности α определяется критическая область W_k , для которой $P(S(X_1, \dots, X_n) \in W_k) = \alpha$. И если величина $S(X_1, \dots, X_n)$, вычисляется при конкретной выборке $X_1, \dots, X_n$, окажется вне критической области W_k , то гипотеза H_0 принимается, а если она окажется в области W_k , то гипотеза H_0 отвергается (или принимается гипотеза H_1). При этом возможны 4 случая, которые представлены в таблице 9.1

Таблица 9.1

	Принимается H_0	Отвергается H_0
Верна H_0	Правильное решение	Ошибка первого рода
вероятность	$1 - \alpha$	α
Верна H_1	Ошибка второго рода	Правильное решение
вероятность	β	$1 - \beta$

Определение. Вероятность α допустить ошибку первого рода называется **уровнем значимости** критерия.

Определение. Вероятность $1 - \beta$ не допустить ошибку второго рода называется **мощностью** критерия.

Если использовать терминологию качества продукции, то α — это «риск поставщика», связанный с забраковкой по результату выборки всей партии товара, соответствующей стандарту, а β — «риск потребителя», связанный с принятием по результатам выборки партии товара, не соответствующей стандарту.

Возможны три варианта расположения критической области:

1. Правосторонняя критическая область (рис 4.1), состоящая из интервала $(t_{кр}^R, \infty)$, где $t_{кр}^R$ определяется из условия:

$$P(S(X_1, \dots, X_n) > t_{кр}^R) = \alpha. \quad (3.1.1)$$

2. Левосторонняя критическая область (рис 4.2), состоящая из интервала $(-\infty; t_{кр}^L)$, где $t_{кр}^L$ определяется из условия:

$$P(S(X_1, \dots, X_n) < t_{кр}^L) = \alpha. \quad (3.1.2)$$

3. Двусторонняя критическая область (рис 4.3), состоящая из интервалов $(-\infty; t_{кр}^L)$ и $(t_{кр}^R, \infty)$, где точки $t_{кр}^L$ и $t_{кр}^R$ определяется из условий:

$$P(S(X_1, \dots, X_n) < t_{кр}^L) = \frac{\alpha}{2} \text{ и } P(S(X_1, \dots, X_n) > t_{кр}^R) = \frac{\alpha}{2}. \quad (3.1.3)$$

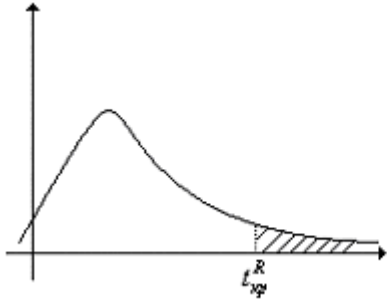


Рис. 4.1

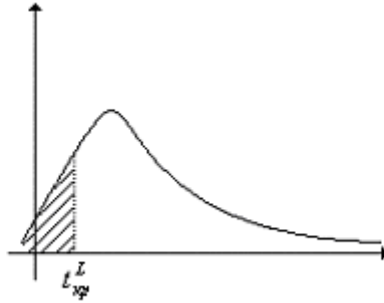


Рис. 4.2

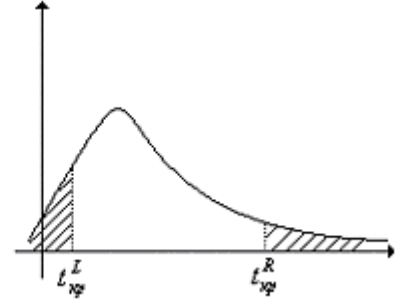


Рис. 4.3

В следующих параграфах рассмотрим несколько конкретных практических примеров.

§ 2. Проверка гипотезы о значении математического ожидания

Пусть из генеральной совокупности X , распределенной по нормальному закону с параметрами (m, σ^2) при неизвестном математическом ожидании m и неизвестной дисперсии σ^2 , взята случайная выборка $X_1, \dots, X_n$ объемом n и вычислена выборочная средняя арифметическая

$$m^* = \frac{1}{n} \sum_{i=1}^n X_i,$$

а m_0 и m_1 — определенные значения параметра m .

Для проверки нулевой гипотезы $H_0 : m = m_0$ при конкурирующей гипотезе $H_1 : m = m_1$ используют статистику

$$T = \frac{m^* - m}{S^*} \sqrt{n-1}, \quad (3.2.1)$$

которая при выполнении нулевой гипотезы H_0 имеет t -распределение (распределение Стьюдента) с $\nu = n - 1$ степени свободы.

Учитывая соотношения (3.1.1) — (3.1.3), определяем границы $t_{кр}^L$ и $t_{кр}^R$ критической области по таблицам t -распределения (распределения Стьюдента) или используя функцию *СТЬЮДРАСПОБР(вероятность; степени_свободы)* из EXCEL, или используя функцию *qt(p, d)* из MATHCAD.

При проверке нулевой гипотезы $H_0 : m = m_0$ при известной дисперсии $H_1 : m_X \neq m_Y$ используют статистику

$$T = \frac{m^* - m}{\sigma} \sqrt{n}, \quad (3.2.2)$$

которая при выполнении нулевой гипотезы H_0 имеет нормальное распределение с параметрами $(0,1)$.

Учитывая соотношения (3.1.1) — (3.1.3), определяем границы $t_{кр}^L$ и $t_{кр}^R$ критической области по таблицам нормального распределения с параметрами $(0,1)$ или используя функцию *НОРМСТОБР(вероятность)* из EXCEL, или используя функцию *qnorm(v, m, σ)* из MATHCAD.

Пример 1. Поставщик двигателей утверждает, что средний срок их службы равен 800 ч. Для выборки из 17 двигателей средний срок службы оказался равным 865 ч. При «исправленном» среднем квадратичном отклонении 120 ч. Проверить гипотезу о том, что значение 800 является математическим ожиданием при 1% уровне значимости.

О Решение. Предположим, что случайная величина среднего времени службы подчинена нормальному закону о числовом значении математического ожидания нормально распределённой величины (генеральной средней) при неизвестной генеральной дисперсии. В этом случае в качестве критерия выбирают функцию

$$T = \frac{m^* - m}{S^*} \sqrt{n-1},$$

где m^* — выборочная средняя, m — математическое ожидание, S^* — «исправленное» выборочное среднее квадратичное отклонение.

Случайная величина T имеет t -распределение (распределение Стьюдента) с $\nu = n - 1$ степени свободы.

Требуется найти критическую область для нулевой гипотезы $H_0 : m = 800$ при альтернативной гипотезе $H_1 : m > 800$.

Критическая область правосторонняя, $\nu = n - 1$ находим из условия $P(T > t_{кр}^R) = \alpha$.

При $\alpha = 0,01$ и $\nu = 17 - 1 = 16$ в таблице t -распределения находим $t_{кр}^R = 1,746$. Таким образом, критическая область $W_{кр} = (1,746 ; +\infty)$. Найдем расчетное значение $t_{расч}$, полагая $m = 800$:

$$t_{расч} = \frac{m^* - m}{S^*} \sqrt{n-1} = \frac{865 - 800}{120} \sqrt{16} = \frac{65}{30} \approx 2,0167.$$

Очевидно, что найденное значение $t_{расч} \approx 2,0167$ попадает в критическую область. Таким образом, нулевая гипотеза не подтверждается. ●

§ 3. Проверка гипотезы о равенстве математических ожиданий двух генеральных совокупностей

Пусть X и Y — две генеральные совокупности с известными дисперсиями σ_X^2 и σ_Y^2 и неизвестными математическими ожиданиями m_X и m_Y . Из генеральных совокупностей взяты две независимые выборки $X_1, \dots, X_n$ и $Y_1, \dots, Y_k$ и вычислены выборочные математические ожидания (средние) m_X^* и

m_Y^* . Для проверки гипотезы о равенстве математических ожиданий $H_0: m_X = m_Y$ используют статистику

$$T = \frac{m_X^* - m_Y^*}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{k}}}, \quad (3.3.1)$$

которая при выполнении нулевой гипотезы H_0 имеет нормальное распределение с параметрами $(0,1)$.

При неизвестных дисперсиях генеральных совокупностей либо требуется достаточно большой объем выборки для надежной и точной оценки, либо требуется, чтобы эти дисперсии были одинаковы, в противном случае известные критерии малоэффективны. Если дисперсии генеральных совокупностей равны $\sigma_X^2 = \sigma_Y^2$, то для проверки нулевой гипотезы $H_0: m_X = m_Y$, используют статистику

$$T = \frac{m_X^* - m_Y^*}{\sqrt{\left(\frac{1}{n_X} + \frac{1}{n_Y}\right) \cdot \frac{S_X^{2*}(n_X - 1) + S_Y^{2*}(n_Y - 1)}{n_X + n_Y - 2}}}, \quad (3.3.2)$$

имеющую t -распределение (распределение Стьюдента) с $\nu = n_X + n_Y - 2$ степенями свободы.

Выбор критической области зависит от вида альтернативной гипотезы:

- при альтернативной гипотезе $H_1: m_X > m_Y$ необходимо выбрать правостороннюю критическую область;
- при альтернативной гипотезе $H_1: m_X < m_Y$ необходимо выбрать левостороннюю критическую область;
- при альтернативной гипотезе $H_1: m_X \neq m_Y$ необходимо выбрать двустороннюю критическую область.

Пример 2. Средний ежедневный объём продаж за I квартал 2004 года для 17 торговцев района А составляет 15 тысяч рублей при “исправленном” среднем квадратичном отклонении 2,5 тысяч рублей, а для 10 торговцев района В — 13 тысяч рублей при «исправленном» среднем квадратичном отклонении 3 тысячи рублей. Существенно ли различие в объёме продаж в районах А и В при 5%-м уровне значимости.

О Решение. Решим задачу, предположив, что ежедневный объём продаж подчинён нормальному закону при неизвестных параметрах распределения. Предположим, что дисперсии объёмов продаж одинаковы. Найдём критическую область для нулевой гипотезы

$$H_0 : m_X = m_Y$$

при альтернативной гипотезе

$$H_1 : m_X \neq m_Y ,$$

где m_X — математическое ожидание объёма продаж для района А, а m_Y — для района В.

В качестве критерия необходимо использовать функцию

$$T = \frac{m_X^* - m_Y^*}{\sqrt{\left(\frac{1}{n_X} + \frac{1}{n_Y}\right) \cdot \frac{S_X^{2*}(n_X - 1) + S_Y^{2*}(n_Y - 1)}{n_X + n_Y - 2}}},$$

где n_X и n_Y объёмы выборок, а m_X^* и m_Y^* оценки параметров m_X и m_Y .

Функция T имеет t -распределение (распределение Стьюдента) с $v = n_X + n_Y - 2 = 17 + 10 - 2 = 25$ степенями свободы.

По таблице t -распределения для $l = 25$ и 5%-го уровня значимости (для двусторонней критической области) находим:

$$P(T < t_{кр}^L) = \frac{0,05}{2} \Rightarrow t_{кр}^L = -2,06,$$

$$P(T > t_{кр}^R) = \frac{0,05}{2} \Rightarrow t_{кр}^R = 2,06.$$

Это означает, что критической областью является

$$W_k = (-\infty; -2,06) \cup (2,06; \infty).$$

Вычислим $t_{расч}$:

$$t_{расч} = \frac{15-13}{\sqrt{\left(\frac{1}{17} + \frac{1}{10}\right) \frac{2,5^2 16 + 3^2 9}{25}}} = 1,86.$$

Полученное значение критерия $t_{расч}$ не принадлежит критической области, следовательно, гипотеза $H_0 : m_X = m_Y$ принимается. ●

§ 4. Проверка гипотезы о значении дисперсии генеральной совокупности

Пусть X — генеральная совокупность, распределенная по нормальному закону с неизвестной дисперсией σ^2 . Из генеральной совокупности взята случайная выборка $X_1, \dots, X_n$ и вычислена выборочная дисперсия σ^{2*} .

Требуется проверить нулевую гипотезу

$$H_0 : \sigma^2 = \sigma_0^2,$$

где σ_0^2 — заданное значение дисперсии генеральной совокупности. Для проверки нулевой гипотезы используют статистику

$$T = \frac{n\sigma^{2*}}{\sigma_0^2}, \quad (3.4.1)$$

которая при выполнении нулевой гипотезы H_0 имеет χ^2 -распределение с $\nu = n - 1$ степенями свободы.

Границы $t_{кр}^L$ и $t_{кр}^R$ критической области определяем по таблицам χ^2 -распределения с $\nu = n - 1$ степенями свободы или используя функцию *ХИ2ОБР (вероятность; степени_свободы)* из EXCEL, или используя функцию *qchisq(p, d)* из MATHCAD.

Пример 3. Точность работы автоматической линии проверяют по дисперсии контролируемого размера, которая не должна превышать $0,1 \text{ мм}^2$. По результатам выборочного контроля получены следующие данные:

Контролируемый размер, Z_i , мм ²	43,0	43,5	43,8	44,4	44,6
Частота n_i	3	7	10	8	2

Требуется проверить при уровне значимости 0,01, обеспечивает ли линия требуемую точность.

О Решение. Задача состоит в проверке гипотезы о значении дисперсии

$$H_0 : \sigma^2 = 0,1$$

при альтернативной гипотезе

$$H_1 : \sigma^2 > 0,1.$$

Таким образом, необходимо построить правостороннюю критическую область. Расчетное значение критерия вычисляем по формуле

$$t_{расч} = \frac{n\sigma^{2*}}{\sigma_0^2}.$$

Следовательно, по данным статистического ряда необходимо вычислить выборочную дисперсию. Дальнейшие вычисления удобно оформить в виде таблицы:

№	Z_i	n_i	$n_i Z_i$	$Z_i - m^*$	$(Z_i - m^*)^2$	$n_i (Z_i - m^*)^2$
1	43,0	3	129,0	-0,863	0,745	2,236
2	43,5	7	304,5	-0,363	0,132	0,924
3	43,8	10	438,0	-0,063	0,004	0,040
4	44,4	8	355,2	0,537	0,288	2,304
5	44,6	2	89,2	0,737	0,543	1,085
Итого		30	1315,9			6,590
Среднее			43,863			0,220

$$m^* = \frac{\sum n_i Z_i}{\sum n_i} = \frac{1315,9}{30} = 43,863;$$

$$\sigma^{2*} = \frac{\sum n_i (Z_i - m^*)^2}{\sum n_i} = \frac{6,59}{30} = 0,22;$$

$$t_{расч} = \frac{n\sigma^{2*}}{\sigma_0^2} = \frac{30 \cdot 0,22}{0,1} = 66.$$

Из условия $P(T > t_{кр}^R) = 0,01$, при числе степеней свободы равным $\nu = n - 1 = 30 - 1 = 29$ находим границу правосторонней критической области. По таблице χ^2 - распределения получаем $t_{кр}^R = 49,5878$.

Заметим, что аналогичный результат можно получить с помощью EXCEL или MATHCAD:

$$t_{кр}^R = \text{ХИ2ОБР}(0,01; 29) = 49,5878$$

или

$$t_{кр}^R = \text{qchisq}(0,99, 29) = 49,5878.$$

Это означает, что критической областью является $W_k = (49,5878; \infty)$. Очевидно, что расчетное значение $t_{расч} = 66$ находится внутри критической области, следовательно, нулевая гипотеза отвергается, т.е. автоматическая линия не обеспечивает заданную точность и требуется ее регулировка. ●

§ 5. Проверка гипотезы о равенстве дисперсий двух генеральных совокупностей

Пусть X и Y — две генеральные совокупности, распределенные по нормальному закону с неизвестными дисперсиями σ_X^2 и σ_Y^2 . Из генеральных со-

вокупностей взяты две независимые выборки $X_1, \dots, X_n$ и $Y_1, \dots, Y_k$ и вычислены исправленные выборочные дисперсии S_X^{2*} и S_Y^{2*} .

Требуется проверить нулевую гипотезу

$$H_0 : \sigma_X^2 = \sigma_Y^2.$$

В данном случае используется статистика

$$T = \frac{S_X^{2*}}{S_Y^{2*}}, \quad (3.5.1)$$

которая имеет F -распределение (распределение Фишера) с $v_1 = n - 1$ и $v_2 = k - 1$ степенями свободы, если $S_X^{2*} > S_Y^{2*}$, и

$$T = \frac{S_X^{2*}}{S_Y^{2*}}, \quad (3.5.2)$$

с числом степеней свободы $v_1 = k - 1$ и $v_2 = n - 1$, если $S_X^{2*} < S_Y^{2*}$.

Если задаться уровнем значимости α , то можно построить критические области для проверки нулевой гипотезы при двух альтернативных гипотезах:

а) $H_1 : \sigma_X^2 > \sigma_Y^2$, если $S_X^{2*} > S_Y^{2*}$, или $H_1 : \sigma_X^2 < \sigma_Y^2$, если $S_X^{2*} < S_Y^{2*}$. В этом случае критическая область правосторонняя $(t_{кр}^R; +\infty)$. Граница $t_{кр}^R$ критической области определяется из условия $P(T(v_1, v_2) > t_{кр}^R) = \alpha$;

б) $H_1 : \sigma_X^2 \neq \sigma_Y^2$. В этом случае критическая область двусторонняя. Однако можно использовать только правостороннюю область $(t_{кр}^R; +\infty)$, где граница $t_{кр}^R$ определяется из условия $P(T(v_1 = n - 1, v_2 = k - 1) > t_{кр}^R) = \frac{\alpha}{2}$, если

$S_X^{2*} > S_Y^{2*}$, и из условия $P(T(v_1 = k - 1, v_2 = n - 1) > t_{кр}^R) = \frac{\alpha}{2}$, если $S_X^{2*} < S_Y^{2*}$.

Пример 4. При измерении производительности двух агрегатов получены следующие результаты в кг вещества за час работы:

№ замера	1	2	3	4	5
Агрегат А	14,1	10,1	14,7	13,7	14,0
Агрегат В	14	14,5	13,7	12,7	14,1

При уровне значимости $\alpha = 0,7$ проверить гипотезу о равенстве дисперсий.

О Решение. Проверим нулевую гипотезу $H_0 : \sigma_X^2 = \sigma_Y^2$ при альтернативной гипотезе $H_1 : \sigma_X^2 \neq \sigma_Y^2$. Вычислим «исправленные» выборочные дисперсии S_X^{2*} и S_Y^{2*} . Для этого сначала найдем выборочные средние m_X^* и m_Y^* :

$$m_X^* = \frac{14,1 + 10,1 + 14,7 + 13,7 + 14}{5} = 13,32;$$

$$m_Y^* = \frac{14 + 14,5 + 13,7 + 12,7 + 14,1}{5} = 13,8.$$

Тогда

$$S_X^{2*} = \frac{(14,1 - 13,32)^2 + (10,1 - 13,32)^2 + (14,7 - 13,32)^2 + (13,7 - 13,32)^2 + (14 - 13,32)^2}{4} = 3,372;$$

$$S_Y^{2*} = \frac{(14 - 13,8)^2 + (14,5 - 13,8)^2 + (13,7 - 13,8)^2 + (12,7 - 13,8)^2 + (14,1 - 13,8)^2}{4} = 0,46.$$

Учитывая, что $S_X^{2*} > S_Y^{2*}$, определяет $t_{расч}$:

$$t_{расч} = \frac{S_X^{2*}}{S_Y^{2*}} = \frac{3,372}{0,46} = 7,3304.$$

Критическое значение $t_{кр}^R$ находим из условия

$$P(T(v_1 = 5 - 1, v_2 = 5 - 1) > t_{кр}^R) = \frac{\alpha}{2} = 0,05$$

По таблице F -распределения (распределения Фишера) с $v_1 = 4$ и $v_2 = 4$ степенями свободы определяем $t_{кр}^R = 6,39$.

Так как число $t_{расч} = 7,3304$ попадает в критическую область $(6,39; +\infty)$, то гипотезу о равенстве дисперсий отвергаем. ●

Замечание. Границу критической области было можно определить и не используя таблицы, например:

- используя функцию *FPАСПОБР* (вероятность; степени_свободы1; степени_свободы2) из EXCEL. При этом задаваемый уровень значимости используется как аргумент «вероятность». В рассматриваемом примере получаем $t_{кр}^R = FPPACPOB(0,04; 4; 4) = 6,39$;
- используя функцию *qF(P,d1,d2)* из MATHCAD, где P — доверительная вероятность $P = 1 - \alpha$, $d1$ и $d2$ степени свободы. В рассматриваемом примере получаем $t_{кр}^R = qF(0.95, 4, 4) = 6,39$.

§ 6. Проверка гипотезы о распределении.

Критерий Пирсона

Пусть $X_1, \dots, X_n$ — выборка, произведенная из генеральной совокупности X , с неизвестной функцией распределения $F(x)$. Проверяется нулевая гипотеза H_0 , утверждающая, что генеральная совокупность распределена по закону, имеющему функцию распределения $F(x)$, равную функции $F_0(x)$,

т.е. проверяется нулевая гипотеза $H_0 : F(x) = F_0(x)$. При этом альтернативной гипотезой является $H_0 : F(x) \neq F_0(x)$.

Наибольшее применение при проверке согласования закона распределения, т.е. проверке нулевой гипотезы $H_0 : F(x) = F_0(x)$, является критерий Пирсона или критерий χ^2 (хи-квадрат).

Наблюдаемое значение критерия (статистика) вычисляется по следующей формуле:

$$\chi^2 = \sum_{i=1}^L \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^L \frac{n_i^2}{np_i} - n. \quad (3.6.1)$$

Согласно теореме К. Пирсона, при $n \rightarrow \infty$ статистика (3.6.1) имеет χ^2 -распределение с $\nu = L - r - 1$ степенями свободы, где L — число групп (интервалов) выборки, r — число параметров предполагаемого распределения.

Схема проверки нулевой гипотезы $H_0 : F(x) = F_0(x)$.

1. По выборке $X_1, \dots, X_n$ строят статистический ряд; он может быть как дискретным, так и интервальным. Рассмотрим для определенности интервальный ряд:

$[x_i, x_{i+1})$	$[x_0, x_1)$	$[x_1, x_2)$	...	$[x_{L-2}, x_{L-1})$	$[x_{L-1}, x_L)$
n_i	n_1	n_2	...	n_{L-1}	n_L

2. По данным предыдущих исследований или по предварительным данным делают предположение (выдвигают гипотезу) о модели закона распределения генеральной совокупности X .

Для предположения модели закона распределения генеральной совокупности X можно использовать графическое представление выборки.

3. Используя выборочные данные, строят оценки параметров выбранной модели закона распределения.

4. Определяют теоретические вероятности попадания случайной величины в интервалы $[x_i, x_{i+1})$ по формуле

$$p_i = F(x_{i+1}) - F(x_i), \quad (3.6.1)$$

где $F(x)$ — предполагаемая функция распределения генеральной совокупности X .

4. Определяют расчетное значение критерия К. Пирсона по формуле

$$\chi^2_{расч} = \sum_{i=1}^L \frac{(n_i - np_i)^2}{np_i}$$

или

$$\chi^2_{расч} = \sum_{i=1}^L \frac{n_i^2}{np_i} - n.$$

5. Выбрав уровень значимости α , находят критическую область (она всегда правосторонняя) $(t_{кр}^R; \infty)$. Границу критической области можно найти одним из следующих:

- используя таблицы χ^2 -распределения с $\nu = L - r - 1$ степенями свободы;
- используя функцию *ХИ2ОБР (вероятность; степени_свободы)* из EXCEL, при этом аргумент «вероятность» данной функции должен быть равен уровню значимости, а аргумент «степени_свободы» должен быть равен $\nu = L - r - 1$;
- используя функцию *qchisq(p, d)* из MATHCAD, при этом переменные данной функции должны быть равны: $p = 1 - \alpha$, $d = L - r - 1$.

6. Если расчетное значение $\chi^2_{расч}$ попадает в критическую область $(t_{кр}^R; \infty)$, то нулевая гипотеза $H_0 : F(x) = F_0(x)$ отвергается и принимается альтернативная гипотеза.

Замечание. При применении критерия К. Пирсона в каждом интервале должно быть не менее 5 элементов выборки (т.е. $n_i \geq 5$). Если это условие не выполняется, то число интервалов надо уменьшить путем объединения соседних интервалов.

Пример 5. Получены значения случайной величины X .

0,54	0,7	0,66	0,55	0,57	0,56	0,62	0,65	0,5	0,67
0,66	0,58	0,57	0,53	0,53	0,53	0,46	0,72	0,41	0,55
0,6	0,61	0,45	0,55	0,63	0,6	0,53	0,51	0,59	0,57
0,59	0,64	0,66	0,56	0,59	0,58	0,55	0,55	0,53	0,53
0,41	0,54	0,45	0,61	0,46	0,52	0,59	0,63	0,4	0,69
0,53	0,48	0,41	0,57	0,6	0,57	0,56	0,56	0,42	0,6
0,54	0,59	0,56	0,5	0,46	0,62	0,57	0,42	0,72	0,47
0,47	0,44	0,52	0,62	0,64	0,41	0,51	0,4	0,49	0,67
0,52	0,58	0,65	0,42	0,68	0,53	0,65	0,59	0,43	0,61
0,56	0,62	0,7	0,49	0,7	0,55	0,5	0,52	0,71	0,48

Необходимо:

1. Найти выборочные характеристики: выборочное среднее и исправленную выборочную дисперсию.
2. Построить доверительной вероятностью 0,95 доверительные интервалы для математического ожидания и дисперсии.
3. Построить гистограмму.
4. Проверить гипотезу о нормальном распределении случайной величины X при уровне значимости $\alpha = 0,05$.

○ Решение. Учитывая, что количество значений равно 100, определяем выборочное среднее и исправленную выборочную дисперсию:

$$m^* = \frac{1}{100} (0,54 + 0,7 + 0,66 + 0,55 + \dots + 0,71 + 0,48) = \frac{55,69}{100} = 0,5569.$$

$$\begin{aligned} S^{2*} &= \frac{1}{100-1} \left((0,54 - 0,5569)^2 + (0,7 - 0,5569)^2 + \dots + (0,48 - 0,5569)^2 \right) = \\ &= \frac{0,64839}{99} = 0,00655. \end{aligned}$$

Замечание. Выборочное среднее можно определить используя функцию *СРЗНАЧ(число1; число2; ...)* из EXCEL, а исправленную выборочную дисперсию можно определить используя функцию *ДИСП(число1;число2; ...)* из EXCEL.

Перейдем к построению доверительных интервалов. При построении доверительного интервала для математического ожидания считаем, что при этом дисперсия не известна. Как известно из предыдущей главы, необходимо использовать формулу

$$m^* - \frac{S^*}{\sqrt{n-1}} t_{\frac{\gamma+1}{2}} < m < m^* + \frac{S^*}{\sqrt{n-1}} t_{\frac{\gamma+1}{2}},$$

в которой неизвестна только величина $t_{\frac{\gamma+1}{2}}$, являющаяся квантилем

t -распределения с $\nu = n - 1$ числом степеней свободы. Для доверительной вероятности $\gamma = 0,95$ найдем квантиль

$$t_{\frac{\gamma+1}{2}} = t_{\frac{0,95+1}{2}} = t_{0,975}$$

t -распределения с $\nu = n - 1 = 100 - 1 = 99$ числом степеней свободы. Для этого используем, например, функцию $qt(p, d)$ из MATHCAD: $qt(0.975, 99) = 1.984$ или функцию *СТЮДРАСПОБ(вероятность; степени_свободы)* из EXCEL: *СТЮДРАСПОБ(2 · (1 - 0,975); 99) = 1,984* или Приложение 5.

Определяем точность оценки

$$\frac{S^*}{\sqrt{n-1}} t_{\frac{\gamma+1}{2}} = \frac{\sqrt{S^{2*}}}{\sqrt{100-1}} t_{0,975} = \frac{\sqrt{0,00655}}{\sqrt{99}} 1,984 = 0,0161.$$

Таким образом, получаем доверительный интервал

$$0,5569 - 0,0161 < m < 0,5569 + 0,0161 \\ 0,5408 < m < 0,5730.$$

Построим доверительный интервал для дисперсии при неизвестном математическом ожидании. В этом случае необходимо использовать формулу

$$\frac{(n-1)S^{2*}}{\frac{r_{1+\gamma}}{2}} \leq \sigma^2 \leq \frac{(n-1)S^{2*}}{\frac{r_{1-\gamma}}{2}},$$

где r_ν — квантиль уровня ν χ^2 -распределения с $n-1$ степенью свободы.

Для данной формулы необходимо вычислить квантили

$$\frac{r_{\gamma+1}}{2} = \frac{r_{0,95+1}}{2} = r_{0,975} \text{ и } \frac{r_{1-\gamma}}{2} = \frac{r_{1-0,95}}{2} = r_{0,025}.$$

Используя функцию $qchisq(p, d)$ из MATHCAD, получим:

$$r_{0,975} = qchisq(0.975, 99) = 128.422 \text{ и } r_{0,025} = qchisq(0.025, 99) = 73.361.$$

Аналогичный результат можно получить используя функцию *ХИ2ОБР* (вероятность; степени_свободы) из EXCEL:

$$r_{0,975} = \text{ХИ2ОБР}(1 - 0.975, 99) = 128,422,$$

$$r_{0,025} = \text{ХИ2ОБР}(1 - 0.025, 99) = 73,361.$$

Таким образом, получаем доверительный интервал

$$\frac{(n-1)S^{2*}}{r_{0,975}} < \sigma^2 < \frac{(n-1)S^{2*}}{r_{0,025}}, \\ \frac{99 \cdot 0,00655}{128,422} < \sigma^2 < \frac{99 \cdot 0,00655}{73,361} . \\ 0,0051 < \sigma^2 < 0,0088$$

Перейдем к построению гистограммы. Построим интервальный ряд. Найдем сначала минимальное и максимальное значения случайной величины (т.е. крайние члены вариационного ряда):

$$X_{\min} = 0,4 \text{ и } X_{\max} = 0,72.$$

Для нахождения минимального и максимального значений случайной величины можно использовать функции *МИН* (*число1;число2; ...*) и *МАКС* (*число1;число2; ...*) из EXCEL. Размах варьирования будет равен

$$R = X_n^* - X_1^* = 0,72 - 0,4 = 0,32.$$

Возьмем число частичных интервалов $l = 8$. В этом случае длина частичного интервала равна

$$h = \frac{R}{n} = \frac{X_n^* - X_1^*}{n} = \frac{0,32}{8} = 0,04.$$

Соответствующий интервальный ряд приведен в таблице 9.2.

Таблица 9.2

Номер интервала i	Средний пробег автомобилей (интервалы) $x_i < X \leq x_{i+1}$	Частота n_i	Относительная частота $f_i^* = \frac{n_i}{nh}$
1	0,4 — 0,44	10	2,5
2	0,44 — 0,48	8	2
3	0,48 — 0,52	9	2,25
4	0,52 — 0,56	21	5,25
5	0,56 — 0,6	21	5,25
6	0,6 — 0,64	13	3,25
7	0,64 — 0,68	10	2,5
8	0,68 — 0,72	8	2

Гистограмма приведена на рис 5.

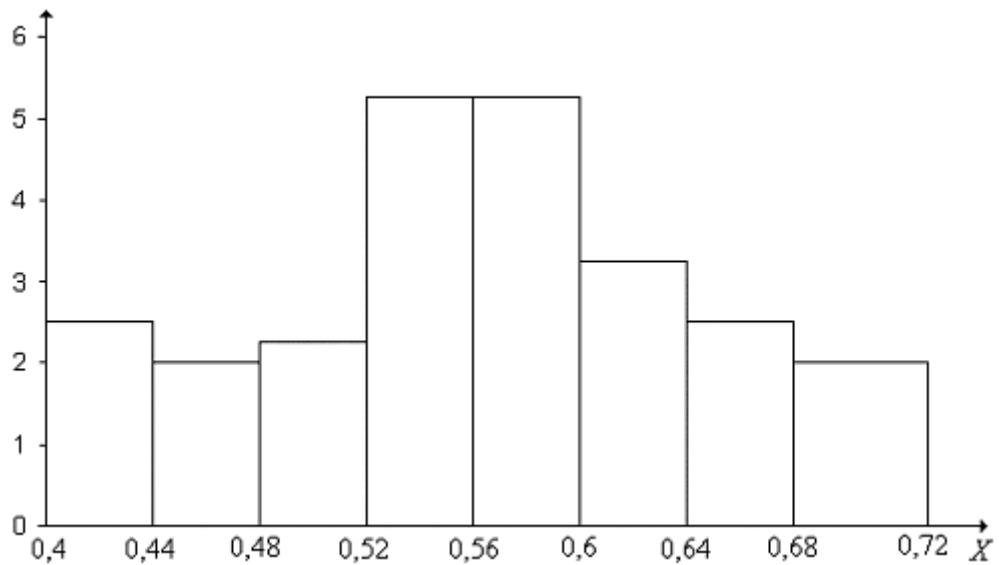


Рис. 5

Проверим гипотезу о распределении случайной величины X . Найдем теоретические вероятности попадания случайной величины в интервалы $[x_i, x_{i+1})$ по формуле $p_i = F(x_{i+1}) - F(x_i)$, где $F(x)$ — функция распределения нормального закона. Так как случайная величина X , подчиненная нормальному закону распределения, определена на $(-\infty; \infty)$, то крайние интервалы в ряде распределения следует заменить на $(-\infty; 0,4)$ и $(0,72; \infty)$. Тогда

$$\begin{aligned}
 p_1 &= P(-\infty < X < 0,4) = \Phi\left(\frac{0,44 - m^*}{\sqrt{S^{2*}}}\right) - \Phi(-\infty) = \\
 &= \Phi\left(\frac{0,44 - 0,5569}{\sqrt{0,00655}}\right) - 0 = \Phi(-1,44) = 1 - \Phi(1,44) = 1 - 0,9251 = 0,0749.
 \end{aligned}$$

$$\begin{aligned}
 p_2 &= P(0,4 < X < 0,44) = \Phi\left(\frac{0,48 - m^*}{\sqrt{S^{2*}}}\right) - \Phi\left(\frac{0,44 - m^*}{\sqrt{S^{2*}}}\right) = \\
 &= \Phi\left(\frac{0,48 - 0,5569}{\sqrt{0,00655}}\right) - \Phi\left(\frac{0,44 - 0,5569}{\sqrt{0,00655}}\right) = \Phi(-0,95) - \Phi(-1,44) = \\
 &= 1 - \Phi(0,95) - (1 - \Phi(1,44)) = 1 - 0,8289 - (1 - 0,9251) = 0,0962.
 \end{aligned}$$

$$\begin{aligned}
p_3 &= P(0,48 < X < 0,52) = \Phi\left(\frac{0,52 - m^*}{\sqrt{S^{2*}}}\right) - \Phi\left(\frac{0,48 - m^*}{\sqrt{S^{2*}}}\right) = \\
&= \Phi\left(\frac{0,52 - 0,5569}{\sqrt{0,00655}}\right) - \Phi\left(\frac{0,48 - 0,5569}{\sqrt{0,00655}}\right) = \Phi(-0,46) - \Phi(-0,95) = \\
&= 1 - \Phi(0,46) - (1 - \Phi(1,95)) = 1 - 0,6772 - (1 - 0,8289) = 0,1517.
\end{aligned}$$

$$\begin{aligned}
p_4 &= P(0,52 < X < 0,56) = \Phi\left(\frac{0,56 - m^*}{\sqrt{S^{2*}}}\right) - \Phi\left(\frac{0,52 - m^*}{\sqrt{S^{2*}}}\right) = \\
&= \Phi\left(\frac{0,56 - 0,5569}{\sqrt{0,00655}}\right) - \Phi\left(\frac{0,52 - 0,5569}{\sqrt{0,00655}}\right) = \Phi(0,04) - \Phi(-0,46) = \\
&= \Phi(0,04) - (1 - \Phi(0,46)) = 0,516 - (1 - 0,6772) = 0,1932.
\end{aligned}$$

$$\begin{aligned}
p_5 &= P(0,56 < X < 0,6) = \Phi\left(\frac{0,6 - m^*}{\sqrt{S^{2*}}}\right) - \Phi\left(\frac{0,56 - m^*}{\sqrt{S^{2*}}}\right) = \\
&= \Phi\left(\frac{0,6 - 0,5569}{\sqrt{0,00655}}\right) - \Phi\left(\frac{0,56 - 0,5569}{\sqrt{0,00655}}\right) = \\
&= \Phi(0,53) - \Phi(0,04) = 0,7019 - 0,516 = 0,1859.
\end{aligned}$$

$$\begin{aligned}
p_6 &= P(0,6 < X < 0,64) = \Phi\left(\frac{0,64 - m^*}{\sqrt{S^{2*}}}\right) - \Phi\left(\frac{0,6 - m^*}{\sqrt{S^{2*}}}\right) = \\
&= \Phi\left(\frac{0,64 - 0,5569}{\sqrt{0,00655}}\right) - \Phi\left(\frac{0,6 - 0,5569}{\sqrt{0,00655}}\right) = \\
&= \Phi(1,03) - \Phi(0,53) = 0,8485 - 0,7019 = 0,1466.
\end{aligned}$$

$$\begin{aligned}
p_7 &= P(0,64 < X < 0,68) = \Phi\left(\frac{0,68 - m^*}{\sqrt{S^{2*}}}\right) - \Phi\left(\frac{0,64 - m^*}{\sqrt{S^{2*}}}\right) = \\
&= \Phi\left(\frac{0,68 - 0,5569}{\sqrt{0,00655}}\right) - \Phi\left(\frac{0,64 - 0,5569}{\sqrt{0,00655}}\right) = \\
&= \Phi(1,52) - \Phi(1,03) = 0,9357 - 0,8485 = 0,0872.
\end{aligned}$$

$$\begin{aligned}
p_8 &= P(0,68 < X < \infty) = \Phi(\infty) - \Phi\left(\frac{0,68 - m^*}{\sqrt{S^{2*}}}\right) = 1 - \\
&- \Phi\left(\frac{0,68 - 0,5569}{\sqrt{0,00655}}\right) = 1 - \Phi(1,52) = 1 - 0,9357 = 0,0643.
\end{aligned}$$

Дальнейшие вычисления удобно оформить в виде таблицы (табл. 9.3).

Таблица 9.3

Номер интервала i	Средний пробег автомобилей (интервалы) $x_i < X \leq x_{i+1}$	Частота n_i	Теоретические частоты p_i	np_i	$\frac{n_i^2}{np_i}$
1	0,4 — 0,44	10	0,0749	7,49	13,3511
2	0,44 — 0,48	8	0,0962	9,62	6,6528
3	0,48 — 0,52	9	0,1517	15,17	5,3395
4	0,52 — 0,56	21	0,1932	19,32	22,8261
5	0,56 — 0,6	21	0,1859	18,59	23,7224
6	0,6 — 0,64	13	0,1466	14,66	11,5280
7	0,64 — 0,68	10	0,0872	8,72	11,4679
8	0,68 — 0,72	8	0,0643	6,43	9,9533
Итого		100	1	100	104,8411

Определяем расчетное значение критерия К. Пирсона:

$$\chi^2_{расч} = \sum_{i=1}^8 \frac{n_i^2}{np_i} - n = 13,3511 + 6,6528 + 5,3395 + 22,8261 + 23,7224 + \\ + 11,528 + 11,4679 + 9,9533 - 100 = 104,8411 - 100 = 4,8411.$$

Находим число степеней свободы. По выборке были рассчитаны два параметра, значит, $r = 2$. Количество интервалов 8, т.е. $L = 8$. Следовательно, $\nu = 8 - 2 - 1 = 5$. Зная, что $\alpha = 0,05$ и $\nu = 5$, находим границу правосторонней критической области $t_{кр}^R = 11,07$ (см. Приложение 4). Таким образом, критической областью является интервал $(11,07; \infty)$.

Так как расчетное значение критерия К. Пирсона $\chi^2_{расч} = 4,8411$ не попадает в критическую область, то нет оснований отвергнуть проверяемую гипотезу о нормальном законе распределения. ●

Приложения

Приложение 1

Таблица значений функции $p_n(m) = \frac{\lambda^m}{m!} e^{-\lambda}$

m	Значения λ						
	0,1	0,2	0,3	0,4	0,5	0,6	0,7
0	0,904837	0,818731	0,740818	0,670320	0,606531	0,548812	0,496585
1	0,090484	0,163746	0,222245	0,268128	0,303265	0,329287	0,347610
2	0,004524	0,016375	0,033337	0,053626	0,075816	0,098786	0,121663
3	0,000151	0,001092	0,003334	0,007150	0,012636	0,019757	0,028388
4	0,000004	0,000055	0,000250	0,000715	0,001580	0,002964	0,004968
5		0,000002	0,000015	0,000057	0,000158	0,000356	0,000696
6			0,000001	0,000004	0,000013	0,000036	0,000081
7					0,000001	0,000003	0,000008
8							0,000001

m	Значения λ						
	0,8	0,9	1	2	3	4	5
0	0,449329	0,406570	0,367879	0,135335	0,049787	0,018316	0,006738
1	0,359463	0,365913	0,367879	0,270671	0,149361	0,073263	0,033690
2	0,143785	0,164661	0,183940	0,270671	0,224042	0,146525	0,084224
3	0,038343	0,049398	0,061313	0,180447	0,224042	0,195367	0,140374
4	0,007669	0,011115	0,015328	0,090224	0,168031	0,195367	0,175467
5	0,001227	0,002001	0,003066	0,036089	0,100819	0,156293	0,175467
6	0,000164	0,000300	0,000511	0,012030	0,050409	0,104196	0,146223
7	0,000019	0,000039	0,000073	0,003437	0,021604	0,059540	0,104445
8	0,000002	0,000004	0,000009	0,000859	0,008102	0,029770	0,065278
9			0,000001	0,000191	0,002701	0,013231	0,036266
10				0,000038	0,000810	0,005292	0,018133
11				0,000007	0,000221	0,001925	0,008242
12				0,000001	0,000055	0,000642	0,003434
13					0,000013	0,000197	0,001321
14					0,000003	0,000056	0,000472
15					0,000001	0,000015	0,000157
16						0,000004	0,000049
17						0,000001	0,000014
18							0,000004
19							0,000001

Плотность стандартного нормального распределения

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

t	0	1	2	3	4	5	6	7	8	9
0,0	0,39894	0,39892	0,39886	0,39876	0,39862	0,39844	0,39822	0,39797	0,39767	0,39733
0,1	0,39695	0,39654	0,39608	0,39559	0,39505	0,39448	0,39387	0,39322	0,39253	0,39181
0,2	0,39104	0,39024	0,38940	0,38853	0,38762	0,38667	0,38568	0,38466	0,38361	0,38251
0,3	0,38139	0,38023	0,37903	0,37780	0,37654	0,37524	0,37391	0,37255	0,37115	0,36973
0,4	0,36827	0,36678	0,36526	0,36371	0,36213	0,36053	0,35889	0,35723	0,35553	0,35381
0,5	0,35207	0,35029	0,34849	0,34667	0,34482	0,34294	0,34105	0,33912	0,33718	0,33521
0,6	0,33322	0,33121	0,32918	0,32713	0,32506	0,32297	0,32086	0,31874	0,31659	0,31443
0,7	0,31225	0,31006	0,30785	0,30563	0,30339	0,30114	0,29887	0,29659	0,29431	0,29200
0,8	0,28969	0,28737	0,28504	0,28269	0,28034	0,27798	0,27562	0,27324	0,27086	0,26848
0,9	0,26609	0,26369	0,26129	0,25888	0,25647	0,25406	0,25164	0,24923	0,24681	0,24439
1,0	0,24197	0,23955	0,23713	0,23471	0,23230	0,22988	0,22747	0,22506	0,22265	0,22025
1,1	0,21785	0,21546	0,21307	0,21069	0,20831	0,20594	0,20357	0,20121	0,19886	0,19652
1,2	0,19419	0,19186	0,18954	0,18724	0,18494	0,18265	0,18037	0,17810	0,17585	0,17360
1,3	0,17137	0,16915	0,16694	0,16474	0,16256	0,16038	0,15822	0,15608	0,15395	0,15183
1,4	0,14973	0,14764	0,14556	0,14350	0,14146	0,13943	0,13742	0,13542	0,13344	0,13147
1,5	0,12952	0,12758	0,12566	0,12376	0,12188	0,12001	0,11816	0,11632	0,11450	0,11270
1,6	0,11092	0,10915	0,10741	0,10567	0,10396	0,10226	0,10059	0,09893	0,09728	0,09566
1,7	0,09405	0,09246	0,09089	0,08933	0,08780	0,08628	0,08478	0,08329	0,08183	0,08038
1,8	0,07895	0,07754	0,07614	0,07477	0,07341	0,07206	0,07074	0,06943	0,06814	0,06687
1,9	0,06562	0,06438	0,06316	0,06195	0,06077	0,05959	0,05844	0,05730	0,05618	0,05508
2,0	0,05399	0,05292	0,05186	0,05082	0,04980	0,04879	0,04780	0,04682	0,04586	0,04491
2,1	0,04398	0,04307	0,04217	0,04128	0,04041	0,03955	0,03871	0,03788	0,03706	0,03626
2,2	0,03547	0,03470	0,03394	0,03319	0,03246	0,03174	0,03103	0,03034	0,02965	0,02898
2,3	0,02833	0,02768	0,02705	0,02643	0,02582	0,02522	0,02463	0,02406	0,02349	0,02294
2,4	0,02239	0,02186	0,02134	0,02083	0,02033	0,01984	0,01936	0,01888	0,01842	0,01797
2,5	0,01753	0,01709	0,01667	0,01625	0,01585	0,01545	0,01506	0,01468	0,01431	0,01394
2,6	0,01358	0,01323	0,01289	0,01256	0,01223	0,01191	0,01160	0,01130	0,01100	0,01071
2,7	0,01042	0,01014	0,00987	0,00961	0,00935	0,00909	0,00885	0,00861	0,00837	0,00814
2,8	0,00792	0,00770	0,00748	0,00727	0,00707	0,00687	0,00668	0,00649	0,00631	0,00613
2,9	0,00595	0,00578	0,00562	0,00545	0,00530	0,00514	0,00499	0,00485	0,00470	0,00457

Функция распределения стандартного нормального закона

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (\Phi(-x) = 1 - \Phi(x))$$

t	0	1	2	3	4	5	6	7	8	9
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861

Критические точки распределения χ^2

Число степеней свободы	Уровень значимости α									
	0,01	0,025	0,05	0,075	0,1	0,9	0,925	0,95	0,975	0,99
1	6,63	5,02	3,84	3,17	2,71	0,02	0,01	0,00	0,00	0,00
2	9,21	7,38	5,99	5,18	4,61	0,21	0,16	0,10	0,05	0,02
3	11,34	9,35	7,81	6,90	6,25	0,58	0,47	0,35	0,22	0,11
4	13,28	11,14	9,49	8,50	7,78	1,06	0,90	0,71	0,48	0,30
5	15,09	12,83	11,07	10,01	9,24	1,61	1,39	1,15	0,83	0,55
6	16,81	14,45	12,59	11,47	10,64	2,20	1,94	1,64	1,24	0,87
7	18,48	16,01	14,07	12,88	12,02	2,83	2,53	2,17	1,69	1,24
8	20,09	17,53	15,51	14,27	13,36	3,49	3,14	2,73	2,18	1,65
9	21,67	19,02	16,92	15,63	14,68	4,17	3,78	3,33	2,70	2,09
10	23,21	20,48	18,31	16,97	15,99	4,87	4,45	3,94	3,25	2,56
11	24,73	21,92	19,68	18,29	17,28	5,58	5,12	4,57	3,82	3,05
12	26,22	23,34	21,03	19,60	18,55	6,30	5,82	5,23	4,40	3,57
13	27,69	24,74	22,36	20,90	19,81	7,04	6,52	5,89	5,01	4,11
14	29,14	26,12	23,68	22,18	21,06	7,79	7,24	6,57	5,63	4,66
15	30,58	27,49	25,00	23,45	22,31	8,55	7,97	7,26	6,26	5,23
16	32,00	28,85	26,30	24,72	23,54	9,31	8,71	7,96	6,91	5,81
17	33,41	30,19	27,59	25,97	24,77	10,09	9,45	8,67	7,56	6,41
18	34,81	31,53	28,87	27,22	25,99	10,86	10,21	9,39	8,23	7,01
19	36,19	32,85	30,14	28,46	27,20	11,65	10,97	10,12	8,91	7,63
20	37,57	34,17	31,41	29,69	28,41	12,44	11,73	10,85	9,59	8,26
21	38,93	35,48	32,67	30,92	29,62	13,24	12,50	11,59	10,28	8,90
22	40,29	36,78	33,92	32,14	30,81	14,04	13,28	12,34	10,98	9,54
23	41,64	38,08	35,17	33,36	32,01	14,85	14,06	13,09	11,69	10,20
24	42,98	39,36	36,42	34,57	33,20	15,66	14,85	13,85	12,40	10,86
25	44,31	40,65	37,65	35,78	34,38	16,47	15,64	14,61	13,12	11,52
26	45,64	41,92	38,89	36,98	35,56	17,29	16,44	15,38	13,84	12,20
27	46,96	43,19	40,11	38,18	36,74	18,11	17,24	16,15	14,57	12,88
28	48,28	44,46	41,34	39,38	37,92	18,94	18,05	16,93	15,31	13,56
29	49,59	45,72	42,56	40,57	39,09	19,77	18,85	17,71	16,05	14,26
30	50,89	46,98	43,77	41,76	40,26	20,60	19,66	18,49	16,79	14,95
31	52,19	48,23	44,99	42,95	41,42	21,43	20,48	19,28	17,54	15,66
32	53,49	49,48	46,19	44,13	42,58	22,27	21,30	20,07	18,29	16,36
33	54,78	50,73	47,40	45,31	43,75	23,11	22,12	20,87	19,05	17,07
34	56,06	51,97	48,60	46,49	44,90	23,95	22,94	21,66	19,81	17,79
35	57,34	53,20	49,80	47,66	46,06	24,80	23,76	22,47	20,57	18,51
36	58,62	54,44	51,00	48,84	47,21	25,64	24,59	23,27	21,34	19,23
37	59,89	55,67	52,19	50,01	48,36	26,49	25,42	24,07	22,11	19,96
38	61,16	56,90	53,38	51,17	49,51	27,34	26,25	24,88	22,88	20,69
39	62,43	58,12	54,57	52,34	50,66	28,20	27,09	25,70	23,65	21,43
40	63,69	59,34	55,76	53,50	51,81	29,05	27,93	26,51	24,43	22,16
99	134,64	128,42	123,23	119,91	117,41	81,45	79,51	77,05	73,36	69,23
100	135,81	129,56	124,34	121,02	118,50	82,36	80,41	77,93	74,22	70,06

Критические точки распределения Стьюдента

Число степеней свободы	Уровень значимости α (двусторонняя критическая область)								
	0,01	0,025	0,05	0,075	0,1	0,2	0,25	0,3	0,5
1	63,66	25,45	12,71	8,45	6,31	3,08	2,41	1,96	1,00
2	9,92	6,21	4,30	3,44	2,92	1,89	1,60	1,39	0,82
3	5,84	4,18	3,18	2,68	2,35	1,64	1,42	1,25	0,76
4	4,60	3,50	2,78	2,39	2,13	1,53	1,34	1,19	0,74
5	4,03	3,16	2,57	2,24	2,02	1,48	1,30	1,16	0,73
6	3,71	2,97	2,45	2,15	1,94	1,44	1,27	1,13	0,72
7	3,50	2,84	2,36	2,09	1,89	1,41	1,25	1,12	0,71
8	3,36	2,75	2,31	2,05	1,86	1,40	1,24	1,11	0,71
9	3,25	2,69	2,26	2,01	1,83	1,38	1,23	1,10	0,70
10	3,17	2,63	2,23	1,99	1,81	1,37	1,22	1,09	0,70
11	3,11	2,59	2,20	1,97	1,80	1,36	1,21	1,09	0,70
12	3,05	2,56	2,18	1,95	1,78	1,36	1,21	1,08	0,70
13	3,01	2,53	2,16	1,94	1,77	1,35	1,20	1,08	0,69
14	2,98	2,51	2,14	1,92	1,76	1,35	1,20	1,08	0,69
15	2,95	2,49	2,13	1,91	1,75	1,34	1,20	1,07	0,69
16	2,92	2,47	2,12	1,90	1,75	1,34	1,19	1,07	0,69
17	2,90	2,46	2,11	1,90	1,74	1,33	1,19	1,07	0,69
18	2,88	2,45	2,10	1,89	1,73	1,33	1,19	1,07	0,69
19	2,86	2,43	2,09	1,88	1,73	1,33	1,19	1,07	0,69
20	2,85	2,42	2,09	1,88	1,72	1,33	1,18	1,06	0,69
21	2,83	2,41	2,08	1,87	1,72	1,32	1,18	1,06	0,69
22	2,82	2,41	2,07	1,87	1,72	1,32	1,18	1,06	0,69
23	2,81	2,40	2,07	1,86	1,71	1,32	1,18	1,06	0,69
24	2,80	2,39	2,06	1,86	1,71	1,32	1,18	1,06	0,68
25	2,79	2,38	2,06	1,86	1,71	1,32	1,18	1,06	0,68
26	2,78	2,38	2,06	1,85	1,71	1,31	1,18	1,06	0,68
27	2,77	2,37	2,05	1,85	1,70	1,31	1,18	1,06	0,68
28	2,76	2,37	2,05	1,85	1,70	1,31	1,17	1,06	0,68
29	2,76	2,36	2,05	1,85	1,70	1,31	1,17	1,06	0,68
30	2,75	2,36	2,04	1,84	1,70	1,31	1,17	1,05	0,68
40	2,70	2,33	2,02	1,83	1,68	1,30	1,17	1,05	0,68
50	2,68	2,31	2,01	1,82	1,68	1,30	1,16	1,05	0,68
60	2,66	2,30	2,00	1,81	1,67	1,30	1,16	1,05	0,68
70	2,65	2,29	1,99	1,81	1,67	1,29	1,16	1,04	0,68
80	2,64	2,28	1,99	1,80	1,66	1,29	1,16	1,04	0,68
90	2,63	2,28	1,99	1,80	1,66	1,29	1,16	1,04	0,68
100	2,63	2,28	1,98	1,80	1,66	1,29	1,16	1,04	0,68

Используемая литература

- *Гмурман В.Е.* Теория вероятностей и математическая статистика. — 9-е изд., стер.— М.: Высшая школа, 2003.— 479 с.
- *Гнеденко Б.В.* Курс теории вероятностей. — 8-е изд., испр. и доп.— М.: Едиториал УРСС, 2005.— 448 с.
- *Красс М.С., Чупрынов Б.П.* Основы математики и ее приложения в экономическом образовании. — М.: Дело, 2002.
- *Кузнецов Б.Т.* Математика. — М.: ЮНИТИ-ДАНА, 2004.
- *Черняк А.А., Новиков В.А., Мельников О.И., Кузнецов А.В.* Математика для экономистов на базе Mathcad. — СПб.: БХВ-Петербург, 2003