7963

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ им. В.Ф. УТКИНА

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ЧАСТЬ 1

Методические указания к практическим занятиям

Рязань 2024

УДК 519.2

Математическая статистика. Часть 1: методические указания к практическим занятиям. / Рязан. гос. радиотехн. ун-т. им. В.Ф. Уткина; сост.: Н.В. Елкина, А.Н. Конюхов. Рязань, 2024. 40 с.

Содержат методические указания и разобранные задачи для практических занятий и самостоятельной работы по теме «Математическая статистика», изучаемой в рамках дисциплины «Математика».

Предназначены для студентов всех направлений и специальностей, изучающих высшую математику.

Табл. 25. Ил. 14. Библиогр.: 2 назв.

Случайная величина, выборка, функция распределения вероятностей, точечная оценка, доверительный интервал, квантиль, статистическая гипотеза, статистический критерий, критическая область

Печатается по решению редакционно-издательского совета Рязанского государственного радиотехнического университета.

Рецензент: кафедра высшей математики Рязанского государственного радиотехнического университета (зав. кафедрой канд. физ.-мат. наук К.В. Бухенский)

1. ЭЛЕМЕНТЫ ОПИСАТЕЛЬНОЙ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Здесь и далее под *выборкой объема n* будем понимать последовательность *n независимых одинаково распределенных* случайных величин (СВ) $X_1, X_2, ..., X_n$, распределение каждой из которых *совпадает* с распределением исследуемой СВ *X*.

На этапе *эмпирической реализации* выборки случайные величины $X_1, X_2, ..., X_n$ принимают конкретные числовые значения $x_1, x_2, ..., x_n$, называемые *вариантами*.

1.1. Описательная статистика дискретной случайной величины

Задача 1.1. Имеется выборка 45 независимых наблюдений дискретной случайной величины *X* (ДСВ *X*). Требуется: 1) составить эмпирический ряд распределения; 2) построить полигон относительных частот; 3) записать эмпирическую функцию распределения $F_X^*(x)$ и построить ее график; 4) определить выборочные характеристики ДСВ: максимальный и минимальный элементы, размах вариации, среднее значение, дисперсию, среднее квадратичное отклонение (СКО), моду, медиану. Элементы выборки: 1, 3, 4, 5, 6, 6, 4, 5, 6, 3, 4, 6, 3, 5, 6, 3, 4, 6, 4, 6, 5, 3, 3, 2, 4, 4, 2, 4, 2, 4, 3, 5, 3, 4, 1, 2, 3, 2, 3, 4, 5, 7, 8, 4, 4.

►1. Ранжируем элементы выборки в порядке возрастания:

1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 8.

Составим эмпирический ряд распределения ДСВ (первые три столбца табл. 1.1), где x_i – различные варианты $(i = \overline{1,8})$; n_i и $w_i = n_i / n$ – абсолютные и относительные частоты вариант соответственно; $n = \sum_{i=1}^{8} n_i = 45$ – объем выборки. Накопленные относительные частоты $w_i^{нак}$ вычисляем как

$$w_1^{нак} = w_1,\ w_i^{нак} = w_{i-1}^{нак} + w_i,\ (i = \overline{2,8}).$$

2. Полигон относительных частот вариант x_i в виде отрезков, последовательно соединяющих точки (x_i, w_i), изобразим по данным табл. 1.1 на рис. 1.1.

Таблица 1.1

Расчетная таблица к задаче 1.1

x_i	n_i	w_i	$w_i^{нак}$	$x_i \cdot w_i$	$x_i^2 \cdot w_i$
1	2	0,0444	0,0444	0,0444	0,0444
2	5	0,1111	0,1556	0,2222	0,4444
3	10	0,2222	0,3778	0,6666	1,9998
4	13	0,2889	0,6667	1,1556	4,6224
5	6	0,1333	0,8000	0,6665	3,3325
6	7	0,1556	0,9556	0,9336	5,6016
7	1	0,0222	0,9778	0,1554	1,0878
8	1	0,0222	1,0000	0,1776	1,4208
Σ	45	1	–	4,0219	18,5537

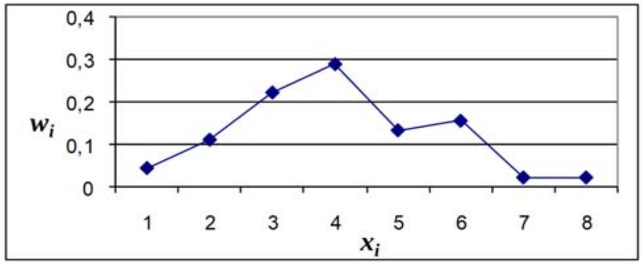

Рис. 1.1. Полигон относительных частот

3. Эмпирическую функцию распределения $F_X^*(x)$ ДСВ *X* составим по накопленным частотам $w_i^{нак}$ табл. 1.1.

$$F_X^*(x) = \begin{cases} 0 & при\ x \le 1, \\ 0{,}0444 & при\ 1 < x \le 2, \\ 0{,}1556 & при\ 2 < x \le 3, \\ 0{,}3778 & при\ 3 < x \le 4, \\ 0{,}6667 & при\ 4 < x \le 5, \\ 0{,}8000 & при\ 5 < x \le 6, \\ 0{,}9556 & при\ 6 < x \le 7, \\ 0{,}9778 & при\ 7 < x \le 8, \\ 1 & при\ x > 8. \end{cases}$$

Построение графика эмпирической функции распределения (рис. 1.2) выполнено в системе компьютерной алгебры MathCAD. Можно построить график эмпирической функции распределения и в Excel, выставив нужные параметры в настройках отображения диаграммы и применив автофигуры (стрелки).

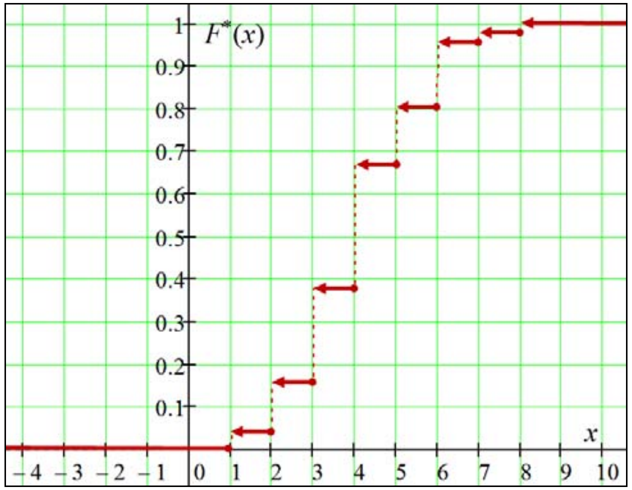

Рис. 1.2. График эмпирической функции распределения ДСВ *X*

4. Используя данные табл. 1.1, рассчитаем некоторые выборочные реализации характеристик ДСВ *X* (все их обозначаем строчными латинскими буквами).

Минимальное выборочное значение $x_{\min} = 1$; максимальное $x_{\max} = 8$.

Выборочный размах вариации: $r = x_{\max} - x_{\min} = 8 - 1 = 7$.

Выборочное среднее: $\overline{x} = \sum_{i=1}^{n} x_i w_i = 4{,}02$.

Выборочная дисперсия:

$$s^2 \overset{def}{=} \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2.$$

Можно получить более удобную формулу для расчетов:

$$s^2 = \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = \frac{n}{n-1}\frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right) =$$

$$= \frac{n}{n-1}\left(\overline{x^2} - 2\overline{x}\cdot\overline{x} + \left(\overline{x}\right)^2\right) = \frac{n}{n-1}\left(\overline{x^2} - \left(\overline{x}\right)^2\right).$$

$$s^2 = \frac{45}{44}\left(18{,}55 - 4{,}02^2\right) = 2{,}44.$$

Выборочное СКО: $s = \sqrt{s^2} = \sqrt{2{,}44} = 1{,}56$.

Выборочная мода mo в дискретном ряду – варианта с наибольшей частотой:

$$\max_{i=1,8}\{n_i\} = 13 \text{ при } x_i = 4 \;\Rightarrow\; mo = 4.$$

Медиана – значение ДСВ, которое делит вариационный ряд на две равные по численности части. В дискретном вариационном ряду с нечётным числом членов выборочной медианой будет варианта, расположенная в центре ряда; с чётным числом членов – средняя арифметическая из двух смежных вариант, стоящих в центре ряда. Так как $n = 45$, то варианта соответствующая медиане будет занимать 23-ю позицию в ранжированном ряду элементов выборки: $me = x_{23} = 4$.

Практически одинаковые выборочные значения моды (4), медианы (4) и среднего (4,02) дают основание предположить, что закон распределения вероятностей ДСВ *X* унимодален и симметричен относительно *x*=4.◄

1.2. Описательная статистика непрерывной случайной величины

Задача 1.2. Данные о независимых наблюдениях непрерывной СВ *X* (НСВ *X*) приведены в табл. 1.2. Требуется: 1) произвести статистическую группировку данных наблюдений, построить интервальный ряд распределения; 2) построить гистограмму относительных частот; 3) построить графики эмпирической функции распределения вероятностей НСВ *X* и кумуляты; 4) определить основные выборочные характеристики НСВ *X*.

Таблица 1.2

Выборка значений НСВ *X* к задаче 1.2

10,4	18,6	10,3	26,0	45,0	18,2	17,3	19,2	25,8	18,7
28,2	25,2	18,4	17,5	41,8	14,6	10,0	37,8	10,5	16,0
18,1	16,8	38,5	37,7	17,9	29,0	10,1	28,0	12,0	14,0
14,2	20,8	13,5	42,4	15,5	17,9	19,2	10,8	12,1	12,4
12,9	12,6	16,8	19,7	18,3	36,8	15,0	37,0	13,0	19,5

►1. Определим минимальное и максимальное значение вариант в выборке и размах вариации, для чего ранжируем исходные данные (табл. 1.3). **ВНИМАНИЕ**. При выполнении расчетов в Excel данные рекомендуется расположить в виде *единст-*

венного столбца (например, А) и применить сортировку (вкладка «Главная», далее «Сортировка и фильтр»).

Таблица 1.3

Выборка, ранжированная в порядке возрастания значений

10,0	10,1	10,3	10,4	10,5	10,8	12,0	12,1	12,4	12,6
12,9	13,0	13,5	14,0	14,2	14,6	15,0	15,5	16,0	16,8
16,8	17,3	17,5	17,9	17,9	18,1	18,2	18,3	18,4	18,6
18,7	19,2	19,2	19,5	19,7	20,8	25,2	25,8	26,0	28,0
28,2	29,0	36,8	37,0	37,7	37,8	38,5	41,8	42,4	45,0

Используя функции Excel, найдем характеристики выборки:

$n = \text{СЧЕТ}(\textit{данные})$, $n = 50$; $x_{\min} = \text{МИН}(\textit{данные})$, $x_{\min} = 10{,}0$;

$$x_{\max} = \text{МАКС}(\textit{данные}),\ \ x_{\max} = 45{,}0.$$

Размах вариации: $r = x_{\max} - x_{\min} = 45{,}0 - 10{,}0 = 35{,}0.$

Так как исследуемая СВ непрерывная, то для анализа выборки составим интервальный статистический ряд. Для нахождения оптимального количества интервалов k используем формулу Стерджесса:

$k = 1 + \left[\log_2 n\right]$, $k = 1 + \text{ОКРУГЛВНИЗ}(\text{LOG}(n;2),0)$ (в Excel),

где $[x]$ – целая часть числа x.

В нашем примере $k = 1 + \left[\log_2 50\right] = 6.$

Размер (длина) интервала $h = r / k = 35{,}0 / 6 = 5{,}833.$

В Excel рекомендуется представить расчет размера интервала следующим образом (рис. 1.3).

C	D	E
Объем выборки	n=	=СЧЁТ(A:A)
Минимальное значение	xmin=	=МИН(A:A)
Максимальное значение	xmax=	=МАКС(A:A)
Размах вариации	R=	=E5-E4
Колич-во инт-ов	k=	=1+ОКРУГЛВНИЗ(LOG(E3;2);0)
Размер инт-ла	h=	=E6/E7

Рис. 1.3. Расчет размера интервала h в Excel

Полученная в нашем примере величина интервала не вполне удобна для построения интервального ряда и выполнения последующих расчетов. Поэтому, учитывая рекомендательный характер формулы Стерджесса, примем k=5. Тогда

$$h = \frac{k}{m} = \frac{35{,}0}{5} = 7{,}0.$$

Необходимо позаботиться, чтобы все элементы выборки были распределены по интервалам. В результате ручного округления величины h возможна ситуация, когда данные, лежащие на краях ранжированной выборочной совокупности, могут оказаться вне крайних интервалов. Во избежание этого примем значение левого конца первого промежутка равным

$$x_{min} - 0{,}5h = 10{,}0 - 0{,}5 \cdot 7{,}0 = 6{,}5,$$

а значение правого конца последнего промежутка равным

$$x_{\max} + 0{,}5h = 45{,}0 + 0{,}5 \cdot 7{,}0 = 48{,}5.$$

Общее число интервалов станет k=5+1=6.

Замечание. При использовании электронных таблиц вовсе не обязательно стремиться к «круглому» значению величины интервала. Скорее всего, это проблема наглядности.

В табл. 1.4 представлен эмпирический интервальный ряд распределения НСВ (первые три строки). Для подсчета количества значений, попадающих в интервал, рекомендуется использовать функцию СЧЁТЕСЛИ(). Кроме того, в табл. 1.4 рассчитаны накопленные частоты $w_i^{нак}$ для построения эмпирической функции распределения и кумуляты, значения w_i/h для построения гистограммы относительных частот, а также данные для расчета основных выборочных характеристик.

Таблица 1.4

Расчетная таблица к задаче 1.2

$(x_i;\, x_{i+1}]$	6,5-13,5	13,5-20,5	20,5-27,5	27,5-34,5	34,5-41,5	41,5-48,5	Σ
n_i	13	22	4	3	5	3	50
w_i	0,26	0,44	0,08	0,06	0,1	0,06	1
$w_i^{нак}$	0,26	0,70	0,78	0,84	0,94	1,00	-
w_i/h	0,037	0,063	0,0114	0,0086	0,014	0,0086	-
x_i^c	10,0	17,0	24,0	31,0	38,0	45,0	-
$x_i^c w_i$	2,6	7,48	1,92	1,86	3,8	2,7	20,4
$(x_i^c)^2 w_i$	26	127,16	46,08	57,66	144,4	121,5	522,8

Замечание. С целью автоматизации расчетов в Excel рекомендуется следующий макет таблицы (табл. 1.5).

Таблица 1.5

Рекомендуемый макет расчетной таблицы в Excel

x_i	x_{i+1}	$(x_i; x_{i+1}]$	n_i	w_i	$w_i^{нак}$	w_i/h	x_i^c	$x_i^c w_i$	$(x_i^c)^2 w_i$
…	…	…	..	..	…	…	…	…	…
	Σ		..	..	-	-	-	…	…

Пояснения к табл. 1.5. Первые два столбца рассчитываются на основании информации о наименьшей варианте и величине интервала h (рис. 1.3). Третий столбец, содержащий информацию о подписях категорий гистограммы, рекомендуется сформировать с использованием функций Excel СЦЕПИТЬ() и ФИКСИРОВАННЫЙ(). Четвертый – с помощью функции СЧЁТЕСЛИ() на основании массива исходных данных. В итоговых ячейках применяется функция СУММ().

2. Построим гистограмму относительных частот, отложив по оси абсцисс отрезки величиной h, а по оси ординат – столбцы высотой w/h (рис. 1.4). В Excel открываем вкладку «Вставка», затем «Гистограмма» – «Гистограмма с группировкой». Площадь, ограниченная гистограммой, равна 1. Гистограмма является выборочным аналогом функции плотности вероятности НСВ.

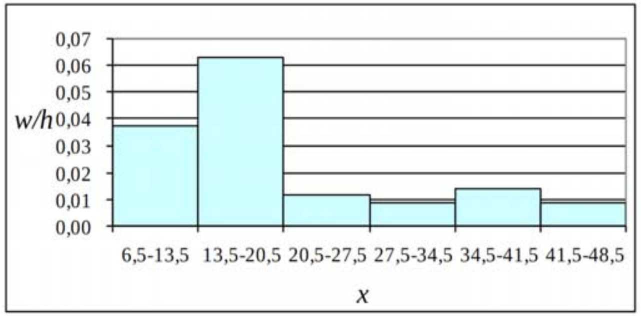

Рис. 1.4. Гистограмма относительных частот НСВ X

3. Запишем эмпирическую функцию распределения вероятностей $F_X^*(x)$ (по накопленным частотам $w_i^{нак}$ табл. 1.4).

$$F_X^*(x)=\begin{cases}0 \quad при\ x\le 13{,}5;\\ 0{,}26 \quad при\ 13{,}5<x\le 20{,}5;\\ 0{,}70 \quad при\ 20{,}5<x\le 27{,}5;\\ 0{,}78 \quad при\ 27{,}5<x\le 34{,}5;\\ 0{,}84 \quad при\ 34{,}5<x\le 41{,}5;\\ 0{,}94 \quad при\ 41{,}5<x\le 48{,}5;\\ 1 \quad при\ x>48{,}5.\end{cases}$$

Построим график эмпирической функции распределения (рис. 1.5). При использовании Excel рекомендуется выбрать гистограмму для накопленных частот без заливки и использовать автофигуры – стрелки (вкладка «Вставка» - «Фигуры»).

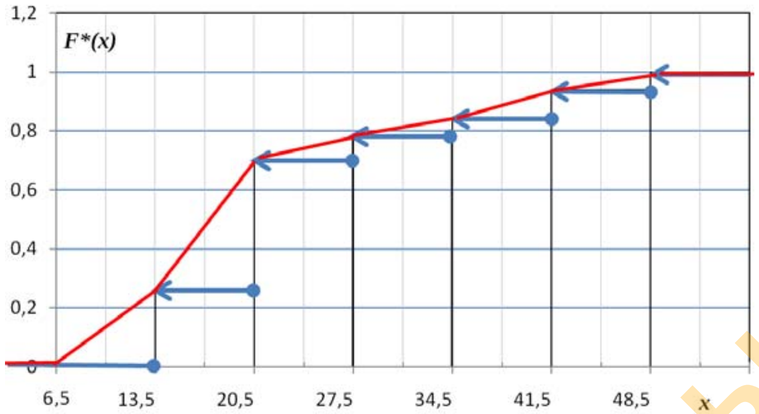

Рис. 1.5. Эмпирическая функция распределения (стрелки) и кумулята (ломаная линия)

Кумулята интервального вариационного ряда строится в той же системе координат (рис. 1.5). Левому концу первого интервала сопоставляется относительная частота, равная нулю, а правому – относительная частота этого интервала. Правому концу второго интервала соответствует накопленная относительная частота второго интервала. Правая граница последнего интервала имеет накопленную относительную частоту 1 (ломаная на рис. 1.5). Кумулята является выборочным аналогом функции распределения вероятностей НСВ.

4. Для расчета основных выборочных характеристик НСВ преобразуем интервальный ряд распределения в дискретный ряд, взяв в качестве значений случайной величины середины

интервалов (табл. 1.5, строка x_i^c): $x_i^c=0{,}5\left(x_i+x_{i+1}\right)$.

Выборочная средняя: $\overline{x}=\sum x_i^c w_i=20{,}4$.

Выборочная дисперсия:

$$s^2=\frac{n}{n-1}\left(\overline{x^2}-\left(\overline{x}\right)^2\right)=\frac{50}{49}\left(522{,}8-20{,}4^2\right)=108{,}8.$$

Выборочное СКО: $s=\sqrt{s^2}=\sqrt{108{,}8}=10{,}4$.

В интервальных рядах распределения с постоянным интервалом h выборочная мода mo определяется по формуле:

$$mo=x_{mo}+h\cdot\frac{f_{mo}-f_{mo-1}}{(f_{mo}-f_{mo-1})+(f_{mo}-f_{mo+1})},$$

где x_{mo} – начало интервала, содержащего моду; f_{mo} – частота модального интервала; f_{mo-1} – частота предмодального интервала; f_{mo+1} – частота постмодального интервала.

В Excel частоту модального интервала можно найти при помощи функции МАКС(). Начало модального интервала по табл. 1.5 – используя функции ПОИСКПОЗ(), СМЕЩ().

В нашем случае выборочная мода

$$mo=13{,}5+7{,}0\cdot\frac{22-13}{(22-13)+(22-4)}=15{,}8.$$

Выборочная медиана me определяется по формуле:

$$me=x_{me}+h\cdot\frac{0{,}5-w_{me-1}^{нак}}{w_{me}},$$

где x_{me} – начальное значение медианного интервала, т.е. интервала, накопленная относительная частота которого равна или впервые превышает $0{,}5$; h – размер медианного интервала; $w_{me-1}^{нак}$ – накопленная относительная частота предмедианного интервала; w_{me} – относительная частота медианного интервала.

$$me=13{,}5+7{,}0\cdot\frac{0{,}5-0{,}26}{0{,}44}=17{,}3.$$

Рекомендуемые функции Excel для автоматизированного расчета медианы по данным табл. 1.5: ПОИСКПОЗ(), СМЕЩ(), ПРОСМОТР(), ИНДЕКС(). Синтаксис и примеры применения функций – см. в справочной службе Excel.◄

2. ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

2.1. Точечная оценка математического ожидания

Задача 2.1. Получить точечную оценку математического ожидания ДСВ X методом моментов в случае данных: 1) несгруппированных 4, 3, 1, 0, 4, 4, 1, -5, 2, 5, 1, 4, 2, 4, 4; 2) сгруппированных (табл. 2.1).

Таблица 2.1

Исходные данные к п. 2 задачи 2.1

x_j	-8	-6	-3	-2	0	3	5	9	11	14
n_j	4	6	3	4	4	4	5	6	3	5

► *Начальным моментом* N_k (заглавная буква «ню» греческого алфавита) *порядка* k случайной величины X называется математическое ожидание СВ X^k:

$$N_k = E\left(X^k\right),\ (k=1,2,...).$$

Выборочный начальный момент N_k^* порядка k представляет собой *статистику* (то есть функцию случайной выборки) вида

$$N_k^* = N_k^*\left(X_1, X_2, ..., X_n\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^k,\ \ (k=1,2,...),$$

эмпирическая реализация которого

$$\nu_k^* = N_k^*\left(x_1, x_2, ..., x_n\right) = \frac{1}{n}\sum_{i=1}^{n} x_i^k.$$

С позиции метода моментов (ММ) N_k^* является точечной оценкой неизвестного теоретического момента N_k:

$$N_k \cong N_k^*,\ \ (k=1,2,...).$$

ММ дает *несмещенные* оценки для начальных моментов любых порядков k. Это значит, что математическое ожидание точечной оценки равно оцениваемому теоретическому параметру. Действительно, поскольку все X_i распределены так же, как и СВ X, то $E\left(X_i^k\right) = E\left(X^k\right)$ и имеет место

$$E\left(N_k^*\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(X^k\right) = \frac{1}{n} n\cdot N_k = N_k.$$

Как частный случай (k=1) точечная *несмещенная* оценка неизвестного математического ожидания $E(X)$ СВ X по методу моментов – выборочный начальный момент первого порядка:

$$E\left(X\right) \cong N_1^* = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

1. Эмпирическая реализация точечной оценки неизвестного математического ожидания $E\left(X\right)$ в случае несгруппированных данных:

$$\nu_1^* = N_1^*\left(x_1, x_2, ..., x_{15}\right) = \frac{1}{15}\sum_{i=1}^{15} x_i =$$

$$= \frac{4+3+1+0+4+4+1+(-5)+2+5+1+4+2+4+4}{15} = 2{,}3.$$

2. В случае сгруппированных данных (табл. 2.2):

$$\nu_1^* = \frac{1}{n}\sum_{j=1}^{m} x_j n_j = \frac{109}{44} = 2{,}5.$$

Таблица 2.2

Расчетная таблица к п. 2 задачи 2.1

x_j	-8	-6	-3	-2	0	3	5	9	11	14	Σ	Σ/n
n_j	4	6	3	4	4	4	5	6	3	5	44	-
$x_j n_j$	-32	-36	-9	-8	0	12	25	54	33	70	109	2,5

Ответ: 1) 2,3; 2) 2,5. ◄

Задача 2.2. Результаты наблюдений значений непрерывной случайной величины X представлены эмпирическим интервальным рядом распределения (табл. 2.3). Получить точечную оценку математического ожидания НСВ X методом моментов.

Таблица 2.3

Эмпирический ряд распределения НСВ X к задаче 2.2

Интервал X	[-3;-1)	[-1; 0,7)	[0,7; 3,4)	[3,4; 6)	[6; 8,9]
n_j	3	7	4	6	3

► Выполним дискретизацию НСВ, сопоставляя каждому промежутку с номером $j=\overline{1,m}$ его середину x_j^c (табл. 2.4). Это и будут соответствующие дискретные варианты. Задача, таким образом, сводится к предыдущей.

Таблица 2.4

Расчетная таблица к задаче 2.2

X	[-3;-1)	[-1; 0,7)	[0,7; 3,4)	[3,4; 6)	[6; 8,9]	Σ	Σ/n
n_j	3	7	4	6	3	23	-
x_j^c	-2	-0,15	2,05	4,7	7,45	-	-
$x_j^c n_j$	-6	-1,05	8,2	28,2	22,35	51,7	2,25

Эмпирическая реализация точечной оценки математического ожидания $E(X)$ НСВ *X*:

$$\nu_1^* = \frac{1}{n}\sum_{j=1}^{m} x_j n_j = \frac{51,7}{23} = 2,25.$$

Ответ: 2,25. ◄

2.2. Точечная оценка дисперсии

Задача 2.3. Методом моментов получить точечную оценку дисперсии СВ *X*, если известно, что одна из выборочных реализаций СВ *X* характеризуется следующими значениями сумм:

$$\sum_{i=1}^{6} x_i = 10,2; \qquad \sum_{i=1}^{6} x_i^2 = 121,14.$$

► *Центральным моментом* M_k (заглавная буква «мю» греческого алфавита) *порядка k* случайной величины *X* называется величина:

$$M_k = E\left[\left(X - E(X)\right)^k\right], \ (k = 2,3,...).$$

Выборочный центральный момент M_k^* порядка *k* представляет собой статистику

$$M_k^* = M_k^*\left(X_1, X_2, ..., X_n\right) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^k, \ (k = 2,3,...),$$

эмпирическая реализация которой

$$\mu_k^* = M_k^*\left(x_1, x_2, ..., x_n\right) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^k.$$

В соответствии с ММ точечная оценка неизвестной дисперсии $D(X)$ есть центральный момент второго порядка:

$$D(X) \cong M_2^* = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

Исходя из определения дисперсии с учетом свойств математического ожидания, имеем:

$$D(X) = E\left(X - E(X)\right)^2 = E\left(X^2 - 2X\cdot E(X) + \left[E(X)\right]^2\right) =$$

$$= E\left(X^2\right) - 2E(X)\cdot E(X) + \left[E(X)\right]^2 = E\left(X^2\right) - \left[E(X)\right]^2.$$

Следовательно, точечная оценка неизвестной дисперсии $D(X)$ может быть записана через начальные моменты:

$$D(X) \cong N_2^* - \left(N_1^*\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2.$$

Однако полученная таким образом оценка дисперсии является *смещенной*. Опираясь на свойства математического ожидания и дисперсии, учитывая независимость $X_i\left(i = \overline{1,n}\right)$, а также несмещенность оценок начальных моментов, получаем

$$E\left(D(X)\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - E\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] =$$

$$= E\left(N_2^*\right) - E\left[\left(\bar{X}\right)^2\right] = N_2 - \left(\left[E\left(N_1^*\right)\right]^2 + D\left(N_1^*\right)\right) =$$

$$= N_2 - N_1^2 - D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = D(X) - \frac{1}{n^2} n\cdot D(X) = \frac{n-1}{n} D(X).$$

Эмпирическая реализация точечной смещенной оценки дисперсии $D(X)$ НСВ *X*:

$$\nu_2^* - \left(\nu_1^*\right)^2 = N_2^*\left(x_1, x_2, ..., x_6\right) - \left[N_1^*\left(x_1, x_2, ..., x_6\right)\right]^2 =$$

$$= \frac{1}{6}\sum_{i=1}^{6} x_i^2 - \left(\frac{1}{6}\sum_{i=1}^{6} x_i\right)^2 = \frac{121,14}{6} - \left(\frac{10,2}{6}\right)^2 = 17,3.$$

Ответ: 17,3.◄

Задача 2.4. В условиях задачи 2.2 найти методом моментов реализацию точечной несмещенной оценки дисперсии НСВ *X*.

► Зададим величину $S^2(X)$:

$$S^2(X) \overset{def}{=} \frac{n}{n-1} D(X) = \frac{n}{n-1}\left(E\left(X^2\right) - \left[E(X)\right]^2\right).$$

Исходя из вывода задачи 2.3, можно утверждать, что *несмещенной* оценкой неизвестной дисперсии является статистика

$$S^{2*}(X_1, X_2, ..., X_n) \overset{def}{=} \frac{n}{n-1} M_2^*(X_1, X_2, ..., X_n),$$

называемая *исправленной выборочной дисперсией*.

Составим расчетную табл. 2.5.

Таблица 2.5

Расчетная таблица к задаче 2.4

X	[-3; -1)	[-1; 0,7)	[0,7;3,4)	[3,4; 6)	[6; 8,9]	Σ	Σ/n
n_j	3	7	4	6	3	23	-
x_j^c	-2	-0,15	2,05	4,70	7,45	-	-
$x_j^c n_j$	-6	-1,05	8,2	28,2	22,35	51,70	2,248
$\left(x_j^c\right)^2$	4	0,0225	4,2025	22,09	55,50	-	-
$\left(x_j^c\right)^2 n_j$	12	0,1575	16,81	132,54	166,51	328,02	14,26

В условиях нашей задачи эмпирической реализацией несмещенной оценки дисперсии $D(X)$ будет

$$s^{2*} = S^{2*}(x_1, x_2, ..., x_{23}) = \frac{n}{n-1}\left(\overline{x^2} - \left(\overline{x}\right)^2\right) =$$

$$= \frac{23}{23-1}\left(14{,}26 - 2{,}248^2\right) = 9{,}6.$$

Ответ: 9,6. ◄

2.3. Интервальные оценки математического ожидания

Задача 2.5. Пусть в условии п. 2 задачи 2.1 СВ *X* распределена по нормальному закону с известной дисперсией $\sigma^2 = 9$.

Найти доверительный интервал для неизвестного математического ожидания *a* СВ *X* с доверительной вероятностью $\gamma = 0{,}9$.

► Пусть $X_1, X_2, ..., X_n$ – случайная выборка значений СВ *X*, имеющей нормальное распределение с неизвестным математическим ожиданием *a*=*E*(*X*) и известной дисперсией $\sigma^2 = D(X)$.

Тогда СВ (среднее арифметическое случайной выборки)

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

имеет нормальное распределение (в силу устойчивости нормально распределенных независимых СВ к суммированию) с математическим ожиданием

$$E\left(\overline{X}\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} n \cdot E(X) = \frac{1}{n} n \cdot a = a$$

и дисперсией

$$\sigma^2\left(\overline{X}\right) = \sigma^2\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} n \cdot \sigma^2(X) = \frac{1}{n}\sigma^2,$$

то есть $\overline{X} \sim N\left(a, \frac{\sigma^2}{n}\right)$.

Проведя стандартизацию СВ $\overline{X}$ (т.е. её центрирование и нормирование), получим СВ *Z*, имеющую стандартное нормальное распределение (рис 2.1):

$$Z = \frac{\overline{X} - a}{\sigma / \sqrt{n}} \sim N(0{,}1).$$

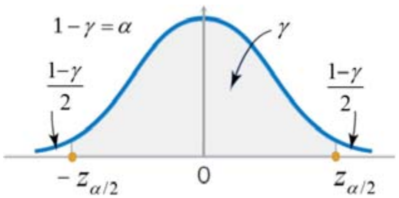

Рис. 2.1. Функция плотности вероятности стандартного нормального распределения

С учетом четности плотности нормального распределения (рис. 2.1) очевидно, что $\gamma = P\left(-z_{\alpha/2,k} < Z < z_{\alpha/2,k}\right)$. Имеем

$$P\left(-z_{\frac{1-\gamma}{2}} < \frac{\overline{X} - a}{\sigma / \sqrt{n}} < z_{\frac{1-\gamma}{2}}\right) = \gamma,$$

где γ – надежность оценки, или с учетом $\gamma = 1 - \alpha$:

$$P\left(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<a<\overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha;\;\; 0<\alpha<1,$$

где $\delta=z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ – точность оценки; $z_{\alpha/2}$ – квантиль нормального распределения, определяемая из условия (рис. 2.1)

$$\Phi(+\infty)-\Phi(z_{\alpha/2})=\frac{\alpha}{2}\;\Leftrightarrow\;\Phi(z_{\alpha/2})=1-\frac{\alpha}{2}\;\Leftrightarrow z_{\alpha/2}=\Phi^{-1}\left(1-\frac{\alpha}{2}\right),$$

где $\Phi(z)$ – функция распределения СВ $Z\sim N(0,1)$:

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{\frac{-t^2}{2}}dt,\; -\infty<z<+\infty .$$

Таким образом, доверительный интервал

$$\overline{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<a<\overline{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

покрывает неизвестный параметр – математическое ожидание a нормально распределенной СВ $\overline{X}$ с заданной доверительной вероятностью $\gamma=1-\alpha$.

Значение $z_{\alpha/2}$ можно найти с помощью встроенной функции Excel НОРМСТОБР$\left(1-\frac{\alpha}{2}\right)$ или НОРМСТОБР$\left(\frac{1+\gamma}{2}\right)$.

В нашем случае имеем

$$z_{\alpha/2}=\Phi^{-1}\left(\frac{1+\gamma}{2}\right)=\Phi^{-1}\left(\frac{1+0,9}{2}\right)=\Phi^{-1}(0,95)=1,645.$$

В п. 2 задачи 2.1 была найдена выборочная средняя $\overline{x}=2,5$ при объеме выборки $n=\Sigma n_i=44$. Дисперсия по условию $\sigma^2=9$.

Точность оценки

$$\delta=z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=\frac{1,645\cdot 3}{\sqrt{44}}=0,74.$$

Доверительный интервал для математического ожидания a

$$2,5-0,74<a<2,5+0,74 \quad\text{или}\quad 1,76<a<3,24 .$$

Ответ: $1,76<a<3,24\;\;(\gamma=0,9)$. ◄

Задача 2.6. В университете, где обучаются N=2000 студентов дневного отделения, была сформирована случайная бесповторная выборка с целью анализа количества пропуска занятий студентом в течение месяца. Полученные при этом результаты представлены в табл. 2.6. Найти границы, в которых с вероятностью 0,997 заключено математическое ожидание a случайной величины X – «количество пропусков занятий студентом в течение месяца».

Таблица 2.6

Фактические данные за месяц март 2024 г.

Пропуски занятий, ед.	1-5	5-9	9-13	13-17	17-21	Σ
Количество студентов	15	20	45	12	8	100

► Используем формулу из предыдущей задачи. Несмотря на то, что дисперсия СВ неизвестна, выборка большая (условно большой выборкой можно считать совокупность объемом n>30) и наблюдения независимы. В таком случае условия центральной предельной теоремы предполагаются выполненными и, следовательно, СВ X распределена «почти» нормально. В качестве оценки дисперсии возьмем статистику S^{2*}. Учтем также, что выборка *бесповторная*. В таком случае реализация доверительного интервала для математического ожидания a:

$$\overline{x}-z_{\alpha/2}\frac{s^*}{\sqrt{n}}\sqrt{1-\frac{n}{N}}<a<\overline{x}+z_{\alpha/2}\frac{s^*}{\sqrt{n}}\sqrt{1-\frac{n}{N}},$$

где N – объем генеральной совокупности.

По условию доверительная вероятность $\gamma=1-\alpha=0,997$; значение $z_{\alpha/2}$ найдем или по таблице значений интегральной функции распределения N(0,1), или с помощью функции Excel НОРМСТОБР$\left(0,5(1+\gamma)\right)$. Получим $z_{\alpha/2}=2,97$ при $\gamma=0,997$ (правило «трех сигм»).

Для нахождения реализаций выборочной средней $\overline{x}$ и выборочной исправленной дисперсии s^{2*} дополним табл. 2.6 серединами частичных интервалов x_j^c и другими данными, необходимыми для расчетов (табл. 2.7). Имеем

$$\bar{x}=10{,}12\,;\quad s^{2*}=\frac{100}{100-1}\left(121{,}5-\left(10{,}12\right)^2\right)=19{,}26.$$

Точность оценки с учетом бесповторного отбора:

$$\delta=\frac{z_{\alpha/2}s^*}{\sqrt{n}}\sqrt{1-\frac{n}{N}}=\frac{2{,}97\cdot\sqrt{19{,}26}}{\sqrt{100}}\sqrt{1-\frac{100}{2000}}=1{,}27\,.$$

Таблица 2.7

Расчетная таблица к задаче 2.6

Интервал X	1-5	5-9	9-13	13-17	17-21	Σ	Σ/n
x_j^c	3	7	11	15	19	-	-
n_j	15	20	45	12	8	100	-
$x_j^c n_j$	45	140	495	180	152	1012	10,12
$\left(x_j^c\right)^2 n_j$	135	980	5445	2700	2888	12148	121,5

Доверительный интервал для математического ожидания a СВ X с надежностью $\gamma=0{,}997$:

$$10{,}12-1{,}27\le a\le 10{,}12+1{,}27;\quad 8{,}85\le a\le 11{,}39.$$

Ответ: с вероятностью 0,997 среднее количество пропусков занятий студентом в марте месяце 2024 года заключено в пределах от 8,85 до 11,39. ◄

Задача 2.7. Найти реализацию доверительного интервала математического ожидания СВ X, распределенной по нормальному закону с неизвестным математическим ожиданием и дисперсией, если известно, что одна из её эмпирических реализаций объемом n=6 характеризуются следующими числовыми показателями:

$$\sum_{i=1}^{6}x_i=-39{,}6\,;\qquad \sum_{i=1}^{6}x_i^2=323{,}16\,.$$

Доверительная вероятность $\gamma=0{,}995$.

► Пусть $X_1, X_2, ..., X_n$ – выборка значений СВ X, имеющей нормальное распределение с неизвестным математическим ожиданием a и неизвестной дисперсией σ^2. Тогда СВ

$$T=\frac{\overline{X}-a}{S/\sqrt{n}}\sim t\left(n-1\right),$$

то есть имеет t-распределение (Стьюдента) с n-1 степенями свободы (рис. 2.2).

Функция плотности t-распределения четная, унимодальная, но имеет более «тяжелые хвосты», нежели стандартное нормальное распределение. При $n\to\infty$ распределение Стьюдента вырождается в стандартное нормальное распределение.

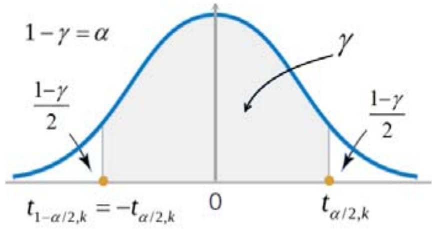

Рис. 2.2. Функция плотности вероятности t-распределения с k степенями свободы

С учетом четности распределения Стьюдента (рис. 2.2) квантиль $t_{\alpha/2,n-1}$ при $\alpha=1-\gamma$ определяется из условия

$$P\left(T>t_{\alpha/2,n-1}\right)=\frac{\alpha}{2},\quad 0<\alpha<1.$$

Следуя логике задачи 2.6, можно записать доверительный интервал математического ожидания a с надежностью $\gamma=1-\alpha$ нормально распределенной СВ при неизвестной дисперсии:

$$\bar{x}-t_{\alpha/2,n-1}\frac{s^*}{\sqrt{n}}<a<\bar{x}+t_{\alpha/2,n-1}\frac{s^*}{\sqrt{n}},$$

где s^* – выборочное исправленное среднее квадратичное отклонение; $t_{\alpha/2,n-1}$ – квантиль t-распределения с n-1 степенями свободы.

Определяется значение $t_{\alpha/2,n-1}$ по заданным α и n-1 по таблице критических точек распределения Стьюдента либо с помощью встроенной функции Excel СТЬЮДРАСПОБР$\left(\alpha;n-1\right)$ или СТЬЮДРАСПОБР$\left(1-\gamma;n-1\right)$.

Замечание. При использовании таблиц для **двусторонней** критической области, ровно как и функции СТЬЮДРАСПОБР, **делить α пополам не нужно**!

Вернемся к нашей задаче. Найдем выборочную среднюю

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{-39{,}6}{6}=-6{,}6$$

и выборочное исправленное СКО:

$$s^*=\sqrt{\frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}x_i^2-\overline{x}^2\right)}=\sqrt{\frac{6}{6-1}\left(\frac{323{,}16}{6}-(-6{,}6)^2\right)}=3{,}516.$$

Для коэффициента доверия $\gamma=0{,}995$, то есть при $\alpha=1-0{,}995=0{,}005$ и $df=n-1=6-1=5$ степеней свободы, найдем квантиль распределения Стьюдента:

$$t_{0{,}005/2;\,5}=\text{СТЬЮДРАСПОБР}(0{,}005;5)=4{,}77.$$

Искомый доверительный интервал для математического ожидания a:

$$-6{,}6-4{,}77\cdot\frac{3{,}516}{\sqrt{6}}<a<-6{,}6+4{,}77\cdot\frac{3{,}516}{\sqrt{6}}.$$

Ответ: $-13{,}45<a<0{,}25\quad(\gamma=0{,}995)$. ◄

Задача 2.8. Из генеральной совокупности извлечена выборка (табл. 2.8).

Таблица 2.8

Исходные данные к задаче 2.8

x_j	-2	1	2	3	4	5
n_j	2	1	2	2	2	1

Оценить с надежностью γ=0,95 математическое ожидание *a* нормально распределенной случайной величины *X* по выборочной средней при помощи доверительного интервала.

► Составим расчетную таблицу 2.9.

Таблица 2.9

Расчетная таблица к задаче 2.9

x_j	-2	1	2	3	4	5	Σ	Σ/n
n_j	2	1	2	2	2	1	10	-
$x_j\cdot n_j$	-4	1	4	6	8	5	20	2,0
$x_j^2\cdot n_j$	8	1	8	18	32	25	92	9,2

Из табл. 2.9 имеем: $\overline{x}=2{,}0$; $s^{2*}=\frac{10}{10-1}\cdot\left(9{,}2-2^2\right)=5{,}78.$

Найдем квантиль *t*-распределения Стьюдента для $df=n-1=9$ степеней свободы при $\alpha=1-\gamma=1-0{,}95=0{,}05$:

$$t_{0{,}05/2;9}=\text{СТЬЮДРАСПОБР}(0{,}05;9)=2{,}26.$$

Точность оценки:

$$\delta=\frac{t_{0{,}05/2;9}\cdot\sqrt{s^{*2}}}{\sqrt{n}}=\frac{2{,}26\cdot\sqrt{5{,}78}}{\sqrt{10}}=1{,}71.$$

Доверительный интервал:

$$2-1{,}71<a<2+1{,}71;\quad 0{,}29<a<3{,}71.$$

Ответ: $0{,}29<a<3{,}71\quad(\gamma=0{,}95)$. ◄

2.4. Интервальные оценки дисперсии и среднего квадратичного отклонения

Задача 2.9. Для оценки неизвестного параметра $\sigma^2(X)$ нормально распределенной СВ X сделана выборка *n*=25 единиц и вычислено $s^{2*}=0{,}64$. Найти доверительный интервал, покрывающий дисперсию $\sigma^2(X)$ с вероятностью $\gamma=0{,}95$.

► Пусть $X_1,X_2,\ldots,X_n$ – выборка значений СВ *X*, имеющей нормальное распределение с неизвестным математическим ожиданием *a* и неизвестной дисперсией σ^2. Тогда СВ

$$X_0^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi_{n-1}^2,$$

т.е. имеет распределение «хи-квадрат» с *n*-1 степенями свободы (рис. 2.3).

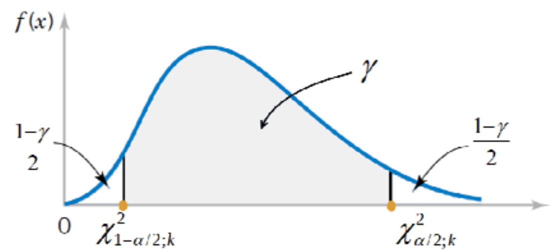

Рис. 2.3. Функция плотности вероятности распределения «хи-квадрат» с *k*≥3 степенями свободы

Эмпирическая реализация интервальной оценки дисперсии σ^2 нормально распределенной СВ X:

$$\frac{(n-1)s^{2*}}{\chi^2_{\alpha/2;n-1}} < \sigma^2 < \frac{(n-1)s^{2*}}{\chi^2_{1-\alpha/2;n-1}},$$

где $\chi^2_{\alpha/2;n-1}$ и $\chi^2_{1-\alpha/2;n-1}$ – верхний и нижний квантили распределения χ^2 с $(n-1)$ степенями свободы (рис. 2.3), которые можно найти либо по таблице критических точек распределения χ^2 (Пирсона), либо с помощью встроенной функции Excel ХИ2ОБР$\left(\frac{1\pm\gamma}{2};n-1\right)$.

В данной задаче $n=25$, $\gamma=0{,}95$. Тогда квантили распределения «хи-квадрат»:

$$\chi^2_{0{,}05/2;24-1} = \text{ХИ2ОБР}\left(\frac{1-0{,}95}{2};25-1\right) = 39{,}36;$$

$$\chi^2_{1-0{,}05/2;24-1} = \text{ХИ2ОБР}\left(\frac{1+0{,}95}{2};25-1\right) = 12{,}40.$$

Доверительный интервал для неизвестной дисперсии:

$$\frac{0{,}64\cdot(25-1)}{39{,}36} < \sigma^2 < \frac{0{,}64\cdot(25-1)}{12{,}40}; \quad 0{,}39 < \sigma^2 < 1{,}24.$$

Интервальная оценка неизвестного среднего квадратичного отклонения σ получается извлечением квадратного корня из пределов доверительного интервала для σ^2.

Ответ: $0{,}39 < \sigma^2 < 1{,}24$; $0{,}62 < \sigma < 1{,}11$ $(\gamma = 0{,}95)$.

3. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

3.1. Проверка гипотезы о виде распределения

Задача 3.1. Для заданной выборки на уровне значимости $\alpha=0{,}05$ проверить гипотезы о законе распределения случайной величины X: 1) H_0: СВ X имеет показательное распределение вероятностей; 2) H_0: СВ X распределена равномерно. Использовать критерий согласия Пирсона. Выборка задана в виде интервального статистического ряда (табл. 3.1).

Таблица 3.1

Исходные данные к задаче 3.1

x_i, x_{i+1}	[0;5)	[5-10)	[10-15)	[15-20)	[20-25)	[25;30]
n_i	17	10	12	9	10	7

Формулируя гипотезу о возможном виде распределения, полезно визуализировать выборочные данные в виде гистограммы абсолютных или относительных частот (рис. 3.1). Можно предположить, что СВ X распределена экспоненциально. Не исключается и равномерное распределение.

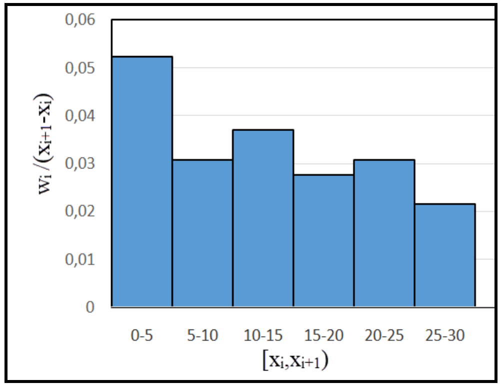

Рис. 3.1. Гистограмма относительных частот к задаче 3.1

В качестве критерия для оценки «схожести» выборочного распределения с гипотетическим распределением применяется статистика X_0^2 (заглавная буква «хи» греческого алфавита), основанная на сравнении наблюдаемых N_i и теоретических N_i' частот по всем m частичным интервалам случайной выборки

$$\mathrm{X}_0^2 = \sum_{i=1}^{m}\frac{(N_i - N_i')^2}{N_i'}.$$

В случае, если выборочное распределение соответствует гипотетическому распределению, данная статистика имеет распределение «хи квадрат» (Пирсона) с числом степеней свободы $df = m-1-r$, где m – число интервалов выборки; r – число параметров предполагаемого распределения.

1. Основная гипотеза H_0: $X \sim E_\lambda$ (СВ X распределена по показательному закону с неизвестным параметром λ). Альтер-

нативная гипотеза $H_1: X \not\sim E_\lambda$.

Сделаем точечную оценку неизвестного параметра распределения λ по выборке на основе ММ. Запишем функцию плотности вероятностей и функцию распределения СВ X

$$X \sim E_\lambda \Leftrightarrow f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0, \\ 0, & \text{иначе}, \end{cases} \quad F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0, \\ 0, & \text{иначе} \end{cases}$$

и найдем начальный теоретический момент первого порядка (математическое ожидание СВ X):

$$E(X) = \int_0^{+\infty} \lambda e^{-\lambda x} dx = \frac{1}{\lambda}.$$

Найдем числовые характеристики эмпирической реализации СВ X, выбрав в качестве вариант центры заданных интервалов x_i^c. Вычисления оформим в табл. 3.2.

Таблица 3.2

Расчетная таблица к задаче 3.1

x_i, x_{i+1}	x_i^c	n_i	$x_i^c n_i$
[0;5)	2,5	17	42,5
[5-10)	7,5	10	75
[10-15)	12,5	12	150
[15-20)	17,5	9	157,5
[20-25)	22,5	10	225
[25;30]	27,5	7	192,5
Σ	-	65	842,5
Σ / n	-	-	12,96

Характеристики выборки: $n = 65$; $\overline{x} = 12{,}96$.

Следовательно, в качестве реализации точечной оценки неизвестного теоретического параметра λ принимаем

$$\lambda^* = \frac{1}{\overline{x}} = \frac{1}{12{,}96} = 0{,}07715.$$

Оценки теоретических вероятностей P_i' и теоретических абсолютных частот N_i':

$$P_i' = F_X(x_{i+1}) - F_X(x_i) = e^{-\lambda^* x_i} - e^{-\lambda^* x_{i+1}},$$

$$N_i' = nP_i' = n\left(e^{-\lambda^* x_i} - e^{-\lambda^* x_{i+1}}\right),$$

где x_i и x_{i+1} – левая и правая границы i-го интервала.

Эмпирическая реализация (наблюдаемое значение) статистики X_0^2 по результатам табл. 3.3:

$$\chi_0^2 = \sum_{i=1}^{m} \frac{\left(n_i - N_i'\right)^2}{N_i'} = 10{,}995.$$

Таблица 3.3

К расчету χ_0^2 в случае гипотезы $H_0: X \sim E_\lambda$

x_i, x_{i+1}	n_i	P_i'	N_i'	$\left(n_i - N_i'\right)^2 / N_i'$
[0;5)	17	0,320	20,804	0,696
[5-10)	10	0,218	14,145	1,215
[10-15)	12	0,148	9,618	0,590
[15-20)	9	0,101	6,540	0,926
[20-25)	10	0,068	4,447	6,936
≥25	7	0,145	9,446	0,633
Σ	65	1	65	10,995

Замечание. При подсчете теоретических вероятностей и частот правый конец заключительного интервала принимается бесконечно удаленным (+∞), так как этого требует поведение функции плотности показательного распределения вероятностей. Поэтому при подсчете теоретической вероятности в выделенной клетке табл. 3.3 используется формула $P_i' = e^{-\lambda^* x_i}$.

При заданном уровне значимости $\alpha = 0{,}05$ и числе степеней свободы $k = 6 - 1 - 1 = 4$ (у показательного распределения единственный параметр λ) найдем критическую точку правосторонней критической области распределения «хи квадрат» с помощью соответствующей функции Excel:

$$\chi^2_{0,05;4} = \text{ХИ2ОБР}(0{,}05; 4) = 9{,}488.$$

Так как $\chi_0^2 > \chi^2_{0,05;4}$, то основная гипотеза H_0 о распределении СВ X по показательному закону отвергается.

2. Выдвинем основную гипотезу $H_0 : X \sim U[a,b]$ (СВ X распределена по равномерному закону с параметрами a и b):

$$X \sim U[a,b] \Leftrightarrow f_X(x) = \begin{cases} (b-a)^{-1}, & a \le x \le b, \\ 0, & \text{иначе.} \end{cases}$$

Оценка параметров равномерного распределения методом максимального правдоподобия дает

$$a \cong X_{\min},\ b \cong X_{\max}.$$

Выборочные реализации оценок: $a^* = x_{\min} = 0$; $b^* = x_{\max} = 30$.

Плотность вероятностей при $a^* \le x \le b^*$ постоянна, причем эмпирическая ее оценка:

$$f^*(x) = \frac{1}{b^* - a^*} = \frac{1}{30-0} = 0{,}03333.$$

Теоретические вероятности и частоты:

$$P_i' = f^*(x)\cdot(x_i - x_{i-1});\quad N_i' = nP_i'.$$

Таблица 3.4

К расчету χ_0^2 в случае гипотезы $H_0 : X \sim U[a,b]$

$x_i,\ x_{i+1}$	n_i	P_i'	N_i'	$(n_i - N_i')^2 / N_i'$
[0;5)	17	0,167	10,833	3,510
[5-10)	10	0,167	10,833	0,064
[10-15)	12	0,167	10,833	0,126
[15-20)	9	0,167	10,833	0,310
[20-25)	10	0,167	10,833	0,064
[25-30]	7	0,167	10,833	1,356
Σ	65	1	65	5,431

Эмпирическая реализация (наблюдаемое значение) статистики «хи квадрат» X_0^2 по результатам табл. 3.4:

$$\chi_0^2 = \sum_{i=1}^{m} \frac{(n_i - N_i')^2}{N_i'} = 5{,}431.$$

Критическую точку правосторонней критической области найдем при уровне значимости $\alpha = 0{,}05$ и числе степеней свободы $df = m - 1 - r = 6 - 1 - 2 = 3$ (равномерное распределение имеет $r = 2$ параметра):

$$\chi^2_{0,05;3} = \text{ХИ2ОБР}(0{,}05;3) = 7{,}81.$$

Так как $\chi_0^2 < \chi^2_{0,05;3}$, то нет оснований отвергнуть гипотезу H_0 о равномерном распределении СВ X.

Ответ: гипотеза о равномерном распределении СВ X с параметрами $a = 0$, $b = 30$ согласуется с опытными данными на уровне значимости α=0,05. ◄

Задача 3.2. Для заданной выборки (табл. 3.5) на уровне значимости $\alpha = 0{,}05$ проверить гипотезу о том, что СВ X имеет нормальное распределение вероятностей.

Таблица 3.5

Исходные данные к задаче 3.2

Интервал	2-5	5-8	8-11	11-14	14-17	17-20
Частота n_i	6	8	15	22	14	5

► Составим вспомогательную таблицу для расчета выборочных характеристик (табл. 3.6).

Таблица 3.6

К расчету основных выборочных характеристик задачи 3.2

Интервалы	2-5	5-8	8-11	11-14	14-17	17-20	Σ
x_i^c	3,5	6,5	9,5	12,5	15,5	18,5	-
n_i	6	8	15	22	14	5	70
$x_i^c n_i$	21	52	142,5	275	217	92,5	800
$(x_i^c)^2 n_i$	73,5	338	1353,8	3437,5	3363,5	1711,3	10277,5

Выборочные характеристики: $n = 70$; $\bar{x} = \frac{800}{70} = 11{,}43$;

$$s = \sqrt{\frac{70}{70-1}\left(\frac{10277{,}5}{70} - 11{,}43^2\right)} = 4{,}06.$$

Основная гипотеза H_0: $X \sim N(a,\sigma^2)$, т.е. СВ X распределена по нормальному закону. Реализации точечных оценок неизвестных параметров:

$$a = \bar{x} = 11{,}43,\ \sigma = s = 4{,}06.$$

Выполним стандартизацию СВ X и вариант соответственно:

$$Z = \frac{X-a}{\sigma};\quad z_i = \frac{x_i - \bar{x}}{s}.$$

Теоретические частоты находим по формуле $N_i' = nP_i'$, где $n = 70$ – объем выборки, $P_i' = \Phi(z_{i+1}) - \Phi(z_i)$; $\Phi(x)$ – функция распределения СВ $Z \sim N(0,1)$, значения которой можно найти с помощью функции Excel $\Phi(x) = \text{НОРМСТРАСП}(x;1)$, где параметр «1» означает что будет возвращено значение интегральной функции распределения.

Таблица 3.7

К расчету χ_0^2 в случае гипотезы H_0: $X \sim N(a,\sigma^2)$

x_i	2	5	8	11	14	17	Σ
x_{i+1}	5	8	11	14	17	20	-
z_i	-∞	-1,585	-0,845	-0,106	0,634	1,374	-
z_{i+1}	-1,585	-0,845	-0,106	0,634	1,374	+∞	-
$\Phi(z_i)$	0	0,056	0,199	0,458	0,737	0,915	-
$\Phi(z_{i+1})$	0,056	0,199	0,458	0,737	0,915	1	-
P_i'	0,056	0,142	0,259	0,279	0,178	0,085	1
N_i'	3,95	9,97	18,13	19,54	12,48	5,93	70
n_i	6	8	15	22	14	5	70
$\frac{(n_i - N_i')^2}{N_i'}$	1,062	0,390	0,540	0,311	0,185	0,146	2,64

Наблюдаемое значение статистики «хи квадрат» X_0^2:

$$\chi_0^2 = \sum_{i=1}^{s} \frac{(n_i - N_i')^2}{N_i'} = 2{,}64.$$

Критическое значение при уровне значимости $\alpha = 0{,}05$ и числе степеней свободы $df = s-1-r = 6-1-2 = 3$ (нормальное распределение имеет $r = 2$ параметра):

$$\chi^2_{0,05;3} = \text{ХИ2ОБР}(0{,}05;3) = 7{,}81.$$

Так как $\chi_0^2 < \chi^2_{0,05;3}$, то нет оснований отвергнуть H_0.

Ответ: гипотеза $H_0: X \sim N(11.43,\ 4.06^2)$ согласуется с опытными данными (α=0,05).◄

3.2. Проверка гипотез о числовых значениях параметров

Рассмотрим проверку гипотезы о среднем значении *a* нормально распределенной СВ *X* (будем говорить «нормальной генеральной совокупности»). Напомним, что несмещенной состоятельной точечной оценкой *a* является средняя $\bar{X}$ случайной выборки $X_1, X_2, ..., X_n$.

Задача 3.3. Из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma = 5{,}2$ извлечена выборка объема $n = 100$ и по ней найдена выборочная средняя $\bar{x} = 27{,}56$. Требуется на уровне значимости $\alpha = 0{,}05$ проверить нулевую гипотезу $H_0: a = a_0 = 26{,}00$ при альтернативной гипотезе $H_1: a \neq a_0 = 26{,}00$. В случае, если нулевая гипотеза отвергнута, выяснить при каком минимальном уровне значимости (*P*-уровень значимости) все еще можно отвергнуть H_0.

►Рассматривается случай нормального распределения с известной дисперсией σ^2. Стандартизируем (центрируем и нормируем) СВ *X*, то есть перейдем от распределения $X \sim N(a,\sigma^2)$ к стандартному нормальному распределению $Z \sim N(0,1)$. При этом в качестве точечной оценки неизвестного генерального среднего *a* берем $\bar{X}$. Также учитываем, что среднее квадратичное отклонение (СКО) среднего n независимых одинаково распределенных СВ $X \sim N\left(a, \sigma^2\right)$ равно $\sigma / \sqrt{n}$. Следовательно,

$$Z=\frac{\left(\bar{X}-a\right)}{\sigma / \sqrt{n}}=\frac{\left(\bar{X}-a\right)\sqrt{n}}{\sigma},$$

и если нулевая гипотеза верна $H_0: a=a_0$, то статистика

$$Z_0=\frac{\left(\bar{X}-a_0\right)\sqrt{n}}{\sigma}$$

имеет известное распределение $N(0,1)$.

Выборочная реализация статистики Z_0 называется наблюдаемым значением z_0. В нашем случае

$$z_0=\frac{\left(\bar{x}-a_0\right)\sqrt{n}}{\sigma}=\frac{(27{,}56-26{,}00)\sqrt{100}}{5{,}2}=3{,}0.$$

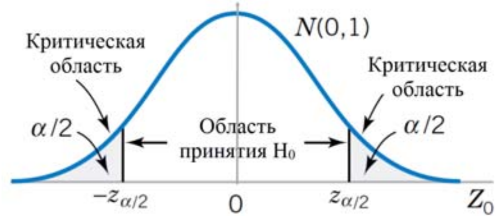

Рис. 3.2. Двусторонняя критическая область для гипотезы H_0 (альтернативная гипотеза $H_1: a \neq a_0$)

Альтернативная гипотеза имеет вид $H_1: a \neq a_0$, поэтому критическая область – двусторонняя (рис. 3.2). В силу симметрии распределения левая критическая точка определяется как $-z_{\alpha/2}$.

Найдем правую критическую точку $z_{\alpha/2}$ из равенства:

$$\frac{\alpha}{2}=\Phi(+\infty)-\Phi\left(z_{\alpha/2}\right) \Leftrightarrow \Phi\left(z_{\alpha/2}\right)=1-\frac{\alpha}{2};$$

$$z_{\alpha/2}=\Phi^{-1}\left(1-\frac{\alpha}{2}\right).$$

Используя Excel, получаем правую критическую точку

$$z_{0{,}05/2}=\text{НОРМСТОБР}\left(1-\frac{0{,}05}{2}\right)=1{,}96.$$

Так как $\left|z_0\right|>z_{0{,}05/2}$ (3,0>1,96), то наблюдаемое значение критерия попадает в критическую область, следовательно, нулевую гипотезу отвергаем на заданном уровне значимости α=0,05.

Выясним, при каком минимальном уровне значимости (P-уровень значимости) все еще можно отвергнуть H_0. В случае двусторонней критической области

$$\left(1-\frac{P}{2}\right)=\left|\Phi\left(z_{набл}\right)\right| \Leftrightarrow P=2\left(1-\left|\Phi\left(z_{набл}\right)\right|\right)$$

имеем $P=2\left(1-\left|\Phi(3)\right|\right)=0{,}0027$.

Примечание: $\Phi(3)=\text{НОРМСТРАСП}(3;1)=0{,}99865$.

Ответ: выборочная средняя (27,56) и гипотетическая генеральная средняя (26,00) отличаются значимо (α=0,05). Значение P-уровня значимости (P=0.0027<<0.05) свидетельствует о высокой надёжности вывода об отклонении нулевой гипотезы.◄

Задача 3.4. В соответствии с технологическим регламентом масса детали является нормально распределенной СВ с математическим ожиданием $a_0=0{,}50$ г, и СКО $\sigma=0{,}11$ г. Выборочная проверка 121 образца показала, что средняя масса детали из этой партии равна $\bar{x}=0{,}53$ г. Требуется при уровне значимости $\alpha=0{,}01$ проверить нулевую гипотезу $H_0: a=a_0$ при альтернативной гипотезе $H_1: a>a_0$. Найти P-уровень значимости.

► Наблюдаемое значение Z_0-статистики:

$$z_0=\frac{\left(\bar{x}-a_0\right)\sqrt{n}}{\sigma}=\frac{(0{,}53-0{,}5)\sqrt{121}}{0{,}11}=3{,}0.$$

По условию альтернативная гипотеза имеет вид $a>a_0$, поэтому критическая область – правосторонняя (рис. 3.3).

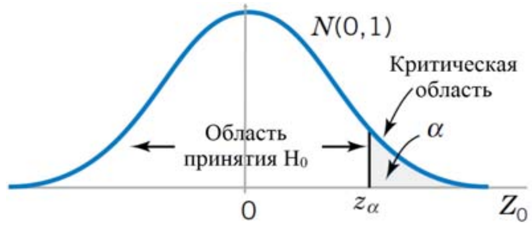

Рис. 3.3. Правосторонняя критическая область для H_0 ($H_1: a>a_0$)

Найдем критическую точку из условия

$$\alpha = \Phi(+\infty) - \Phi(z_\alpha) \Leftrightarrow \Phi(z_\alpha) = 1-\alpha;\ z_\alpha = \Phi^{-1}(1-\alpha).$$

$$z_{0,01} = \text{НОРМСТОБР}(1-0,01) = 2,33.$$

Так как $z_0 > z_{0,01}$, то наблюдаемое значение критерия попадает в критическую область. Следовательно, нулевую гипотезу отвергаем на уровне значимости α=0,01.

Р-уровень для правосторонней критической области:

$$1-P = \Phi(z_0);$$

$$P = 1-\Phi(z_0) = 1-\Phi(3,0) = 1-0,99865 = 0,00135.$$

Ответ: масса детали значимо отличается от номинала (α=0,01). Рассчитанное значение *P*-уровня значимости (*P*=0,00135<<0,01) свидетельствует о достаточно высокой надёжности данного статистического вывода. ◄

Задача 3.5. Отдел технического контроля произвел выборочный анализ изделий (табл. 3.8), поперечный размер которых предположительно является нормально распределенной СВ с математическим ожиданием $a_0 = 35,0$ мм.

Таблица 3.8

Исходные данные к задаче 3.5

Размер	x_i	34,8	34,9	35,0	35,1	35,3
Число изделий	n_i	5	6	4	3	2

Требуется при уровне значимости $\alpha = 0,05$ проверить нулевую гипотезу $H_0:\ a = a_0 = 35,0$ при альтернативной гипотезе 1) $H_1: a > a_0 = 35,0$; 2) $H_1: a < a_0 = 35,0$; 3) $H_1: a \neq a_0 = 35,0$.

►В данном случае генеральная дисперсия σ^2 неизвестна. Логичным представляется в этом случае в качестве оценки неизвестной дисперсии использовать ее выборочную несмещенную оценку – исправленную дисперсию. В курсе математической статистики доказывается, что если $X_1, X_2, ..., X_n$ – случайная выборка из нормального распределения со средним значением *a*, то случайная величина

$$T = \frac{(\bar{X}-a)}{S/\sqrt{n}} = \frac{(\bar{X}-a)\sqrt{n}}{S}$$

имеет *t*-распределение (Стьюдента) с количеством степеней свободы $df = n-1$. Здесь СВ $S = \sqrt{\frac{n}{n-1}\left(\overline{X^2} - \bar{X}^2\right)}$ – исправленное среднее квадратичное отклонение в случайной выборке.

Таким образом, если нулевая гипотеза верна $H_0: a = a_0$, то статистика

$$T_0 = \frac{(\bar{X}-a_0)\sqrt{n}}{S}$$

имеет известное распределение и может выступать критерием при тестировании гипотез.

Выборочная реализация статистики (наблюдаемое значение) рассчитывается по формуле

$$t_0 = \frac{(\bar{x}-a_0)\sqrt{n}}{s},$$

где $s = \sqrt{\frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right)}$ – реализация выборочного СКО.

Расчет числовых характеристик заданной выборки выполним в табл. 3.9.

Таблица 3.9

Расчетная таблица к задаче 3.5

x_i	34,8	34,9	35,0	35,1	35,3	Σ	$\frac{1}{n}\Sigma$
n_i	2	3	4	6	5	20	-
$x_i n_i$	69,6	104,7	140	210,6	176,5	701,4	35,07
$x_i^2 n_i$	2422	3654	4900	7392	6230	24598	1230

Выборочные характеристики:

$$\bar{x} = 35,07;\quad s = \sqrt{\frac{20}{20-1}\left(1230 - 35,07^2\right)} = 0,166.$$

Наблюдаемое значение критерия:

$$t_0 = \frac{(\overline{x} - a_0)\sqrt{n}}{s} = \frac{(35{,}07 - 35)\sqrt{20}}{0{,}166} = 1{,}89.$$

1. По условию альтернативная гипотеза имеет вид $H_1 : a > a_0$, следовательно, критическая область – правосторонняя. В Excel найдем критическое значение, учитывая, что функция СТЬЮДРАСПОБР$(\alpha; n-1)$ дает значения критической точки для двусторонней критической области, т.е. для α/2. Это значит, что в случае односторонней критической области следует *удвоить* значение α. Таким образом, в правостороннем случае критическая точка

$$t_{\alpha,n-1} = \text{СТЬЮДРАСПОБР}(2\alpha; n-1).$$

Критическая точка для правосторонней критической области:

$$t_{0,05;19} = \text{СТЬЮДРАСПОБР}(2 \cdot 0{,}05; 19) = 1{,}73.$$

Так как $t_0 > t_{0,05;19}$ (1,89>1,73), то нулевую гипотезу отвергаем при уровне значимости α=0,05.

2. Если альтернативная гипотеза имеет вид $H_1 : a < a_0$, то критическая область – левосторонняя. В силу симметрии критическая точка левосторонней критической области:

$$-t_{0,05;19} = -1{,}73.$$

Так как $t_0 > -t_{0,05;19}$ (1,89>-1,73), то нет оснований отвергать нулевую гипотезу.

3) В случае альтернативной гипотезы $H_1 : a \neq a_0$ критическая область двусторонняя, поэтому в качестве аргумента функции СТЬЮДРАСПОБР указываем α (не делить на два!):

$$t_{\alpha/2,n-1} = \text{СТЬЮДРАСПОБР}(\alpha; n-1),$$

$$t_{0,05/2;19} = \text{СТЬЮДРАСПОБР}(0{,}05; 19) = 2{,}09.$$

Так как $|t_0| < t_{0,05/2;19}$ (1,89<2,09), то нет оснований отвергать нулевую гипотезу.

Рассмотренный пример является поучительным. Мы видим, что в случае альтернативных гипотез вида $H_1 : a > a_0 = 35{,}0$ и $H_1 : a \neq a_0 = 35{,}0$ получены противоположные статистические выводы. Прояснить ситуацию можно, рассчитав *P*-критическое значение. Для этого используем наблюдаемое значение критерия $t_0 = 1{,}89$ и функцию Excel

СТЬЮДРАСП($t_{набл}$; *df*; «хвосты»),

где *df* – количество степеней свободы; «хвосты» – число возвращаемых хвостов распределения. Если «хвосты» = 1, функция СТЬЮДРАСП возвращает одностороннее распределение. Если «хвосты» = 2, функция СТЬЮДРАСП возвращает двустороннее распределение.

Имеем для двусторонней критической области:

СТЬЮДРАСП(1,89; 19; 2)=0,074,

что свидетельствует об очень малом «запасе прочности» статистического вывода на уровне значимости α=0,05. Ведь уже на уровне значимости α=0,075 нулевую гипотезу можно будет отвергнуть. В подобных случаях рекомендуется нарастить объем выборочной совокупности, если это возможно.

Ответ: полученные статистические выводы ненадежны. Рекомендуется увеличить объем выборки.◄

Задача 3.6. Из нормальной генеральной совокупности с некой дисперсией σ^2 извлечена выборка объемом $n = 21$, и по ней найдена исправленная выборочная дисперсия $s^2 = 16{,}2$. Требуется при уровне значимости $\alpha = 0{,}01$ проверить нулевую гипотезу $H_0: \ \sigma^2 = \sigma_0^2 = 15{,}0$, приняв в качестве альтернативной гипотезы: 1) $H_1 : \sigma^2 > \sigma_0^2 = 15{,}0$; 2) $H_1 : \sigma^2 \neq \sigma_0^2 = 15{,}0$.

►В качестве критерия используется случайная величина

$$X_0^2 = \frac{(n-1)S^2}{\sigma_0^2},$$

имеющая распределение «хи-квадрат» с количеством степеней свободы $df = n - 1$.

Наблюдаемое значение критерия:

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(21-1)\cdot 16{,}2}{15{,}0} = 21{,}6.$$

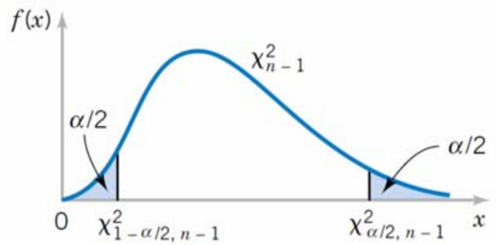

Рис. 3.4. Двусторонняя критическая область $H_1: \sigma^2 \neq \sigma_0^2$

В качестве существенного отличия от распределений (нормальное и распределение Стьюдента) отметим, что распределение «хи-квадрат» не является симметричным (рис. 3.4-3.6).

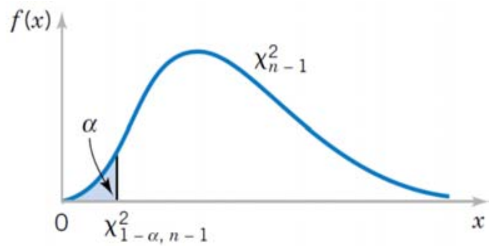

Рис. 3.5. Левосторонняя критическая область $H_1: \sigma^2 < \sigma_0^2$

Этот факт надо учитывать при определении критических областей. По условию альтернативная гипотеза имеет вид $\sigma^2 > \sigma_0^2 = 15{,}0$, поэтому критическая область – правосторонняя (рис. 3.5).

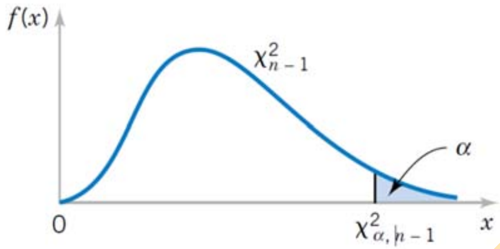

Рис. 3.6. Правосторонняя критическая область $H_1: \sigma^2 > \sigma_0^2$

Найдем по уровню значимости $\alpha = 0{,}01$ и числу степеней свободы $df = n - 1 = $ $= 21 - 1 = 20$ критическую точку правосторонней критической области при $\alpha = 0{,}01$:

$$\chi^2_{0.01;21-1} = \text{ХИ2ОБР}\left(0.01;21-1\right) = 37{,}57\,.$$

Так как $\chi_0^2 < \chi^2_{крит,пс}$, то нет оснований отвергать нулевую гипотезу о равенстве генеральной дисперсии гипотетическому значению $\sigma_0^2 = 15{,}0$.

Ответ: различие между выборочной исправленной дисперсией (16,2) и гипотетической генеральной дисперсией (15,0) незначимо $(\alpha = 0{,}01)$. ◄

Задача 3.7. В результате продолжительного хронометража времени работы опытных специалистов над техническим заданием установлено, что дисперсия этого времени $\sigma_0^2 = 2{,}0$ мин². Результаты 20 наблюдений за работой молодого инженера представлены в табл. 3.10.

Таблица 3.10

Исходные данные к задаче 3.7

Время работы, мин	x_i	56	58	60	62	64
Частота	n_i	1	4	10	3	2

Можно ли на уровне значимости $\alpha = 0{,}05$ считать, что молодой инженер работает ритмично (в том смысле, что дисперсия затрачиваемого им времени значимо не отличается от дисперсии времени опытных специалистов)?

► Для нахождения исправленной дисперсии дополним заданную табл. 3.10 и получим табл. 3.11.

Таблица 3.11

Расчетная таблица к задаче 3.7

x_i	56	58	60	62	64	Σ	$\frac{1}{n}\Sigma$
n_i	1	4	10	3	2	20	-
$x_i n_i$	56	232	600	6	128	1202	60,1
$x_i^2 n_i$	3136	13456	36000	11532	8192	72316	3615,8

Выборочная средняя:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{s} x_i n_i = 60{,}1\,.$$

Исправленная выборочная дисперсия:

$$s^2 = \frac{n}{n-1}\left(\overline{x^2} - \bar{x}^2\right) = \frac{20}{20-1}\left(3615{,}8 - 60{,}1^2\right) = 4{,}0\,.$$

В качестве нулевой гипотезы примем $H_0: \sigma^2 = \sigma_0^2 = 2{,}0$, при этом альтернативная гипотеза будет $H_1: \sigma^2 \neq \sigma_0^2 = 2{,}0$.

Найдем наблюдаемое значение критерия:

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(20-1)\cdot 4{,}0}{2{,}0} = 38{,}0 .$$

Так как конкурирующая гипотеза имеет вид $\sigma^2 \neq \sigma_0^2$, то критическая область – двусторонняя. Найдем по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $df = n-1 = 20-1 = 19$ критические точки (рис. 3.4):

$$\chi^2_{0,05/2;19} = \text{ХИ2ОБР}\left(\frac{\alpha}{2}; df\right) = \left(\frac{0{,}05}{2}; 19\right) = 32{,}85 ,$$

$$\chi^2_{1-0,05/2;19} = \text{ХИ2ОБР}\left(1-\frac{\alpha}{2}; df\right) = \left(1-\frac{0{,}05}{2}; 19\right) = 8{,}91 .$$

Так как условие $\chi^2_{1-0,05/2;19} < \chi_0^2 < \chi^2_{0,05/2;19}$ не выполняется, то нулевая гипотеза отвергается.

Ответ: молодой инженер работает неритмично $(\alpha = 0{,}05)$. ◄

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Montgomery, Douglas C. Applied statistics and probability for engineers / Douglas C. Montgomery, George C. Runger. – 3rd ed. John Wiley & Sons., 2003. 821 p.
2. Основные задачи математической статистики: учеб. пособие / Т.В. Довжик, М.Е. Ильин; Рязан. гос. радиотехн. ун-т. – Рязань, 2023. 112 с.

ОГЛАВЛЕНИЕ

Математическая статистика. Часть 1

Составители: Е л к и н а Наталья Викторовна
К о н ю х о в Алексей Николаевич

Редактор И.В. Черникова
Корректор С.В. Макушина
Подписано в печать 30.06.24. Формат бумаги 60×84 1/16.
Бумага писчая. Печать трафаретная. Усл. печ. л. 2,5.
Тираж 60 экз. Заказ
Рязанский государственный радиотехнический университет.
390005, Рязань, ул. Гагарина, 59/1.
Редакционно-издательский центр РГРТУ.