Н.А. Дегтярева

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-практическое пособие

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ
ГУМАНИТАРНО-ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Н.А. Дегтярева

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-практическое пособие

Челябинск, 2018

УДК **519.2 (076)**
ББК **22.172 я 7**
Д **26**

Дегтярева, Н.А. Математическая статистика [Текст]: учебно-практическое пособие / Н.А. Дегтярева. – Челябинск: Изд-во Южно-Урал. гос. гуман.-пед. ун-та, 2018. – 122 с.

ISBN 978-5-6042129-5-0

Пособие содержит методические рекомендации, решение типовых задач и 20-30 вариантов заданий к лабораторно-практическим работам по следующим темам: вариационные ряды и их выборочные характеристики, эмпирическая функция распределения; статистическое оценивание: методы точечного оценивания; статистическое оценивание: интервальные оценки параметров распределения; проверка параметрических гипотез; проверка непараметрических гипотез, критерий Пирсона.

Рецензенты: В.Н. Павленко, д-р физ.-мат. наук, проф.
И.Д. Колмакова, д-р экон. наук, проф.

ISBN 978-5-6042129-5-0

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Математическая статистика – раздел математики, изучающий методы сбора, систематизации и обработки результатов массовых случайных наблюдений с целью выявления статистических закономерностей.

Математическая статистика опирается на теорию вероятностей. Если теория вероятностей изучает закономерности случайных явлений на основе абстрактного описания действительности, то математическая статистика оперирует непосредственно результатами наблюдений над случайными явлениями, представляющими выборку из некоторой конечной или гипотетической бесконечной генеральной совокупности.

Область применения математической статистики очень обширна, выделим несколько основных направлений ее методов:

1) *теория выборок,* посвященная методам формирования выборок из генеральной совокупности экспериментальных данных;

2) *теория оценок,* определяющая методы и способы оценки неизвестных параметров распределений совокупности или решения задачи предсказания, исходя из полученных экспериментальных данных;

3) *проверка статистических гипотез*, используемая, если нужно решить какое-то из предположений о распределении анализируемых данных более правдоподобно;

4) к*орреляционно-регрессионный анализ,* задачами которого являются выявление зависимостей и подбор математических формул, наилучшим образам описывающих статистические данные;

5) *дисперсионный анализ*, позволяющий оценить разброс экспериментальных данных и сопоставить его с конкретной ситуацией, к которой относятся данные.

Статистические программные пакеты сделали эти методы более доступными и наглядными, так как трудоемкую работу по расчету различных статистик, параметров, характеристик, построению таблиц и графиков в основном стал выполнять компьютер. Среди множества используемых для этих целей пакетов прикладных программ выделим универсальные и специализированные статистические пакеты: STADIA, STATISTIKA, SPSS, STATISTIK VIEW, Эвриста, Статистик-консультант, Олимп и др. Но, чтобы использовать эти программы, несомненно, необходимо знать основные идеи и методы математической статистики, условия их применения.

Настоящее учебно-практическое пособие призвано помочь студентам в изучении математической статистики, для решения практических задач математической статистики.

Важным достоинством данного пособия является то, что в нем по каждой рассматриваемой теме из математической статистики приводится краткая теоретическая справка (в виде методических рекомендаций), рассматривается решение типовых задач, а затем предлагаются задания (20–30 вариантов) для самостоятельной работы студентов (в виде лабораторно-практических работ). Приведенные задания могут быть использованы как для промежуточного контроля знаний, так и в качестве практических заданий на экзамене.

Пособие предназначено для студентов-бакалавров по направлению подготовки – педагогическое образование.

ЛАБОРАТОРНО-ПРАКТИЧЕСКАЯ РАБОТА 1
ВАРИАЦИОННЫЕ РЯДЫ И ИХ ХАРАКТЕРИСТИКИ. ЭМПИРИЧЕСКАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ

Цель работы: освоение приемов первичной обработки статистических данных. Вычисление выборочных характеристик. Определение эмпирической функции распределения и построение ее графика.

1.1. Теоретические вопросы

1. Определение математической статистики.
2. Основные задачи математической статистики.
3. Генеральная и выборочная совокупность. Выборочный метод.
4. Выборочный ряд распределения (дискретный и интервальный).
5. Геометрическое изображение статистических рядов. Полигон частот (полигон относительных частот). Гистограмма частот (гистограмма относительных частот).
6. Эмпирическая функция распределения и ее свойства. Аналитическое и графическое представление эмпирической функции распределения.
7. Выборочные характеристики и методика их расчета.

1.2. Методические рекомендации

Чтобы произвести первичную обработку данных, необходимо:

а) произвести группировку данных;

б) изобразить полученный ряд графически;

в) найти выборочные характеристики.

Группировку статистического материала начинают с определения типа группировки. Чтобы определить тип группировки, полезно придерживаться следующего *алгоритма:*

1. Определить число k различных вариант в выборке.

2. Определить тип случайной величины X; если X дискретно, и $k \leq 10$, то шаг 3, иначе – шаг 4.

3. Сгруппировать данные в ***дискретный ряд.***

4. Найти наибольшую x_{max} и наименьшую x_{min} варианты.

5. Вычислить ширину Δx интервала группировки по формуле:

$$\Delta x = \frac{x_{\max} - x_{\min}}{c},$$

где c – количество интервалов (с = 1+ 3,322lgn).

6. Сгруппировать данные в ***интервальный ряд***, взяв в качестве левой границы первого интервала величину:

$$a_0 = x_{\min} - \frac{\Delta x}{2}$$

Статистическим распределением выборки (**дискретным вариационным рядом**) называется ранжированная совокупность вариант x_i с соответствующими им частотами или относительными частотами.

Статистическим распределением выборки (**интервальным вариационным рядом)** называется упорядоченная последовательность интервалов варьирования случайной величины с соответствующими частотами или относительными частотами попаданий в каждый из них значений случайной величины (в качестве частоты интервала принимают сумму частот вариант, попавших в этот интервал).

В качестве геометрической интерпретации ряда используют полигон частот (относительных частот) для дискретного ряда и гистограмму частот (относительных частот) для интервального ряда.

Ломаная, соединяющая точки (x_i, m_i) (или (x_i, ω_i)) отрезками прямых, называется **полигоном частот (или полигоном относительных частот**) или **многоугольником распределения.**

Гистограммой частот (**гистограммой относительных частот**) называют ступенчатую фигуру, состоящую из прямоугольников, построенных на интервалах с шириной Δx и высотой $h_i = \frac{m_i}{\Delta x}$ ($h_i = \frac{\omega_i}{\Delta x}$).

Для вычисления выборочных характеристик, по *сгруппированным данным,* используются следующие формулы:

- ***Начальный выборочный момент порядка j***:

$$m_j^* = \frac{1}{n}\sum_{i=1}^{n} x_i^j m_i .$$

✓ *при j = 1,* $m_1^* = \overline{x}_в$ – ***выборочное среднее:***

$$\overline{x}_в = \frac{x_1 m_1 + .. + x_k m_k}{n} = \frac{1}{n}\sum_{i=1}^{k} x_i m_i = \sum_{i=1}^{k} x_i \omega_i ;$$

✓ *при j = 2,* $m_2^* = \overline{x^2} = \frac{1}{n}\sum x_i^2 m_i$.

- ***Центральный выборочный момент порядка j***:

$$\alpha_j^* = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x}_в)^j m_i ;$$

✓ если j = 2, $\alpha_2^* = D_в$ – ***выборочная дисперсия***:

$$D_в = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_в)^2 m_i$$

$$\text{или} \quad D_в = \overline{x^2} - (\bar{x})^2 .$$

- ***Выборочное среднее квадратическое отклонение***: $\sigma_в = \sqrt{D_в}$.

- ***«Исправленная» дисперсия:*** $S^2 = \frac{n}{n-1} D_в$.

- ***«Исправленное» среднее квадратическое отклонение:*** $S = \sqrt{S^2}$.

- ***Коэффициент вариации V:*** $V = \frac{\sigma_в}{\bar{x}_в} \cdot 100\%$.

- ***Для дискретного ряда – мода выборки M_0*** – варианта, имеющая максимальную частоту.

- ***Для интервального ряда – мода выборки M_0:***

$$M_o = x_o + k\frac{m_i - m_{i-1}}{(m_i - m_{i-1}) + (m_i - m_{i+1})}$$

где x_0 – начало модального интервала, т.е. интервала, имеющего максимальную частоту,
k – длина модального интервала,
m_i – частота модального интервала,
m_{i-1} *и* m_{i+1} – частоты соответственно предшествующего и последующего за модальным интервалов.

- ***Для дискретного ряда – медиана выборки m_e***:

✓ если число вариант *нечетное* $m_e = x_{k+1}$ – значение серединного элемента;

✓ если число вариант *четное*, то $m_e = \frac{x_k + x_{k+1}}{2}$ – среднее арифметическое двух серединных элементов.

▪ ***Для интервального ряда – медиана выборки m_e:***

$$m_e = x_o + k\frac{n/2 - T_{i-1}}{m_i},$$

где x_0 – начало медианного интервала, т.е. интервала, в котором содержится серединный элемент,

k – длина медианного интервала,

n – объем выборки,

T_{i-1} – сумма частот интервалов, предшествующих медианному,

m_i – частота медианного интервала.

Эмпирической функцией распределения (***функцией распределения выборки***) называют функцию $F^*(x)$, определяющую для каждого значения x *относительную частоту* события $X < x$:

$$F^*(x) = \omega(X < x) = \frac{m_x}{n}, \quad \text{или} \quad F^*(x) = \omega_i^{нак} = \frac{m_i^{нак}}{n},$$

где m_x – число вариант меньших x; n – объем выборки.

Эмпирическая (выборочная) функция распределения $F^*(x)$ случайной величины X является *приближением* теоретической функции распределения $F(x)$.

Свойства эмпирической функции распределения

1. $0 \le F^*(x) \le 1$.
2. $P(x_1 \le X \le x_2) = F^*(x_2) - F^*(x_1)$.

3. $F^*(x)$ есть неубывающая функция на всей числовой оси, т.е. $F^*(x_1) \le F^*(x_2)$ при $x_1 < x_2$.

4. Если x_1 – наименьшая варианта, то $F^*(x) = 0$ при $x \le x_1$. Иначе – $F^*(-\infty) = 0$.

Если x_k – наибольшая варианта, то $F^*(x) = 1$ при $x > x_k$. Иначе – $F^*(+\infty) = 1$.

График эмпирической функции $F^*(x)$ распределения *дискретной случайной величины* – кусочно-постоянная функция. Она имеет скачки в точках, которые соответствуют имеющимся вариантам. Ее аналитический вид:

$$F^*(x) = \begin{cases} 0, & x \le x_1 \\ \dfrac{m_x}{n}, & x_1 < x < x_2 \\ 1 & \text{при } x > x_2 \end{cases}.$$

График эмпирической функции $F^*(x)$ распределения *непрерывной случайной величины* строят иначе. В этом случае в качестве представителя интервала берется его *середина*, или правый конец интервала. Объединяя отрезками точки, координатами которых являются середины (правые концы) интервалов и накопленные относительные частоты соответствующих интервалов, получаем ломаную линию, являющуюся хорошим приближением ***графика эмпирической функции*** *распределения непрерывной случайной величины*. Аналитический вид этой функции довольно сложен.

1.3. Решение типовых задач

Задача 1. Регистрация размеров, продаваемой в магазинах мужской обуви, дала следующие данные о 79 покупках:

39	40	38	43	41	42	40	38	41	42
36	43	41	42	38	41	40	42	41	42
42	40	40	39	41	39	38	40	41	41
37	40	42	43	42	38	40	40	41	41
43	41	40	43	41	42	42	39	43	41
41	40	42	39	41	41	42	42	40	40
41	39	40	40	39	42	40	43	41	41
43	42	42	39	42	41	42	40	41	

***1.** Провести статистическую обработку данных:*

a) составить вариационный ряд;

b) определить m_i, ω_i*;*

c) составить выборочный ряд распределения.

***2.** Изобразить полученный ряд графически.*

***3.** Вычислить выборочные характеристики.*

***4.** Найти эмпирическую функцию по распределению выборки и изобразить ее графически.*

Решение: Проведем первичную статистическую обработку данных:

1. Группировка данных

Случайная величина X – размер мужской обуви, дискретная. Полученные данные представляют собой выборку из n = 79 наблюдений.

Вначале составим ранжированный ряд:

36, 37, 38, 38, 38, 38, 38, 39, …, 39, 40, …, 40, 41, …, 41, 42, …, 42, 43, ..., 43.

Получено восемь различных значений случайной величины (восемь вариант x_i) – 36, 37, 38, 39, 40, 41, 42, 43 – вариационный ряд.

Подсчитаем частоту m_i значений варианты и соответствующую относительную частоту $\omega_i = \frac{m_i}{n}$:

$x_1 = 36$ $m_1 = 1$ $\omega_1 = 1/79$;

$x_2 = 37$ $m_2 = 1$ $\omega_2 = 1/79$;

$x_3 = 38$ $m_3 = 5$ $\omega_3 = 5/79$;

$x_4 = 39$ $m_4 = 8$ $\omega_4 = 8/79$;

$x_5 = 40$ $m_5 = 17$ $\omega_5 = 17/79$;

$x_6 = 41$ $m_6 = 21$ $\omega_6 = 21/79$;

$x_7 = 42$ $m_7 = 18$ $\omega_7 = 18/79$;

$x_8 = 43$ $m_8 = 8$ $\omega_8 = 8/79$.

Статистическое распределение выборки (***дискретный вариационный ряд***):

x_i	36	37	38	39	40	41	42	43
m_i	1	1	5	8	17	21	18	8
ω_i	1/79	1/79	5/79	8/79	17/79	21/79	18/79	8/79

2. Графическое представление ряда

Построим полигон частот. По оси x откладывают варианты x_i, а по оси y – соответствующие им частоты m_i.

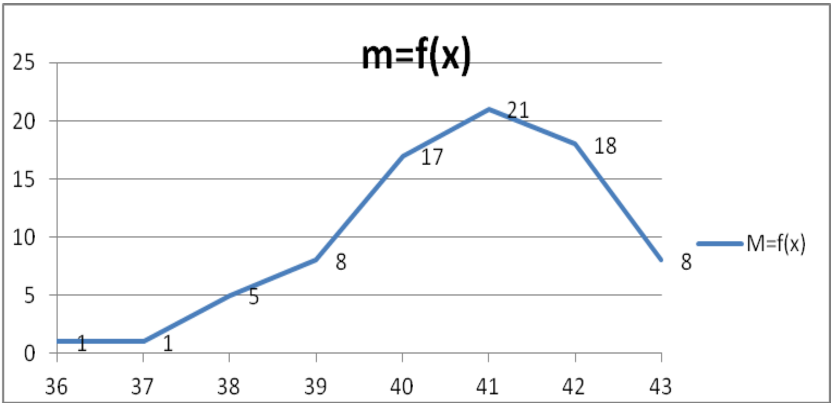

Рис. 1. Полигон частот

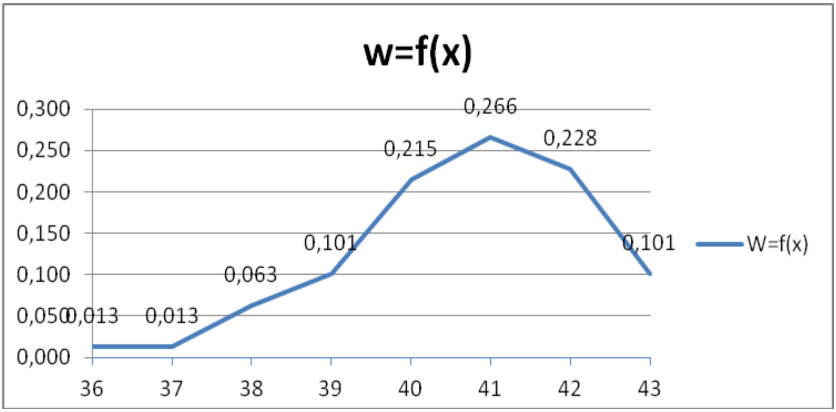

Рис. 2. Полигон относительных частот

3. Вычисление выборочных характеристик

Вычисление выборочных характеристик удобно вести с помощью расчетной таблицы:

Расчетная таблица

x_i	m_i	$x_i m_i$	x_i^2	$x_i^2 m_i$	$x_i - \overline{x}_в$	$(x_i - \overline{x}_в)m_i$	$(x_i - \overline{x}_в)^2$	$(x_i - \overline{x}_в)^2 m_i$
36	1	36	1296	1296	–4,71	–4,71	22,17	22,17
37	1	37	1369	1369	–3,71	–3,71	13,76	13,76
38	5	190	1444	7220	–2,71	–13,54	7,34	36,69
39	8	312	1521	12168	–1,71	–13,67	2,92	23,36
40	17	680	1600	27200	–0,71	–12,05	0,50	8,54
41	21	861	1681	35301	0,29	6,11	0,08	1,78
42	18	756	1764	31752	1,29	23,24	1,67	30,01
43	8	344	1849	14792	2,29	18,33	5,25	41,99
∑	79	3216	12524	131098	–9,67	0,00	53,69	178,30

$$m_j^* = \frac{1}{n}\sum_{i=1}^{k} x_i^j * m_i \quad .$$

$$m_1^* = \frac{1}{n}\sum_{i=1}^{k} x_i m_i = \overline{x}_в$$

$$m_1^* = \frac{1}{79}\cdot 3216 = 40.7 = \overline{x}_в$$

$$m_2^* = \overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 * m_i = 1659,47$$

$$m_2^* = \frac{1}{79} * 131098 = 1659,47$$

$$\alpha_j^* = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x}_в)^j m_i \quad .$$

$$\alpha_1^* = \frac{1}{n}\sum_{i=1}^{k} (x_i - \overline{x}_в) m_i$$

$$\alpha_1^* = \frac{1}{79}\cdot 0.00 = 0.00$$

$$\alpha_2^* = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x}_в)^2 m_i = D_в$$

$$\alpha_2^* = \frac{1}{79}178,30 = 2,26 = D_в$$

или

$$D_в = \overline{x^2} - (\overline{x})^2 = \frac{131098}{79} - (40.71)^2 = 2.26$$

$$\sigma_В = \sqrt{D_В} = \sqrt{2,26} = 1,50$$

$$S^2 = \frac{n}{n-1}D_в = \frac{79}{78}\cdot 2.26 = 2.29$$

$$S = \sqrt{S^2} = \sqrt{2,29} = 1,51$$

$$M_0 = 41$$

$$m_e = \frac{39+40}{2} = 39.5$$

$$V = \frac{\delta_в}{\overline{x}_в}\cdot 100\%,$$

$$V = \frac{\sqrt{2.26}}{40.71} = 3.69.$$

1. Эмпирическая функция распределения

Для нахождения аналитического выражения эмпирической функции $F^*(x)$ и построения ее графика составим таблицу:

x_i	36	37	38	39	40	41	42	43	
$m_i^{нак}$	0	1	2	7	15	32	53	71	79
$\omega_i^{нак}$	0	1/79	2/79	7/79	15/79	32/79	53/79	71/79	1

где $\omega_i^{нак}$ – относительная накопленная частота.

$m_i^{нак}$ – накопленная частота.

Накопленные частоты *получают путем последовательного суммирования частот, начиная с первой.*

Эмпирическая функция $F^*(x)$ распределения дискретной случайной величины:

$$F^*(x) = \begin{cases} 0, & x \le 36 \\ \frac{1}{79}, & 36 < x \le 37 \\ \frac{2}{79}, & 37 < x \le 38 \\ \frac{7}{79}, & 38 < x \le 39 \\ \frac{15}{79}, & 39 < x \le 40 \\ \frac{32}{79}, & 40 < x \le 41 \\ \frac{53}{79}, & 41 < x \le 42 \\ \frac{71}{79}, & 42 < x \le 43 \\ 1, & x > 43 \end{cases}$$

График эмпирической функции $F^*(x)$ распределения дискретной случайной величины (рис. 3):

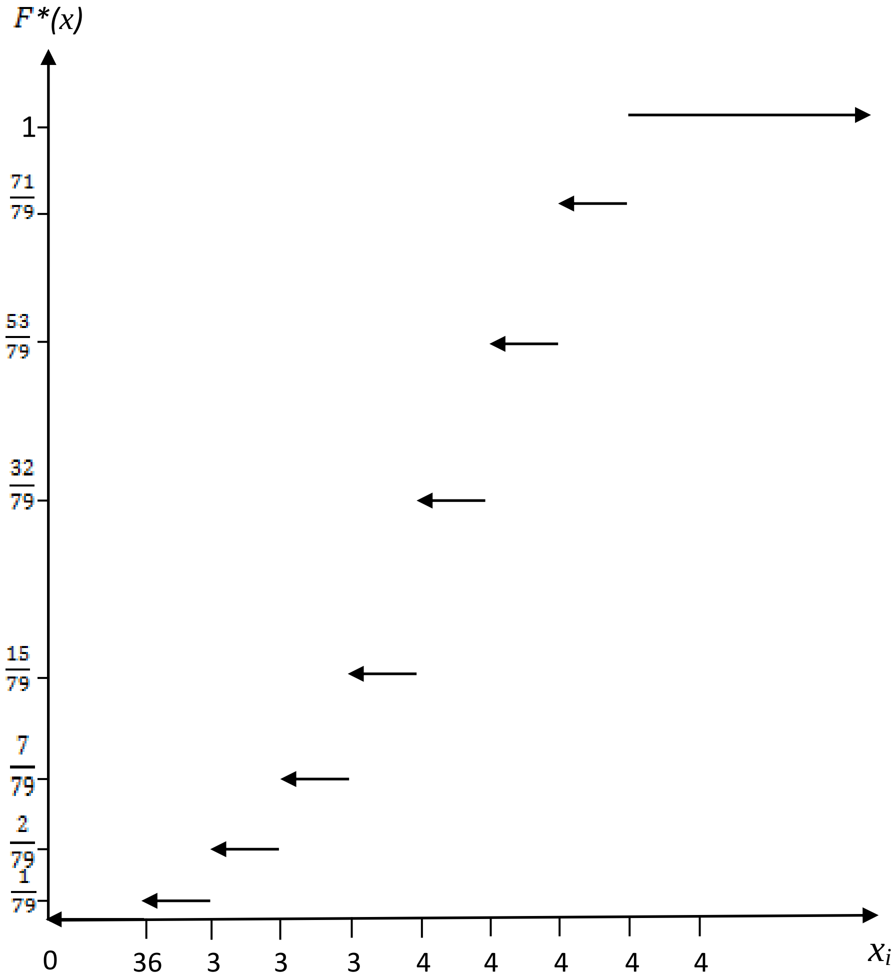

Рис. 3. График эмпирической функции $F^*(x)$ распределения дискретной случайной величины

Задача 2. При измерении роста 56 новобранцев получены следующие данные:

188	170	76	178	76	180	76	185	184	174	68	174	189	172
175	167	179	176	169	178	169	171	170	179	176	179	174	176
188	178	172	176	167	166	180	183	176	182	178	172	185	183
175	174	180	166	169	171	178	169	170	179	171	178	173	177

1. *Провести статистическую обработку данных:*

a) составить вариационный ряд;

b) определить m_i, ω_i*;*

c) составить выборочный ряд распределения.

2. *Изобразить полученный ряд графически.*

3. *Вычислить выборочные характеристики.*

4. *Найти эмпирическую функцию по распределению выборки и изобразить ее графически.*

Решение: Проведем первичную статистическую обработку данных:

2. Группировка данных

Случайная величина X – рост новобранца, непрерывная величина. Сгруппируем данные в ***непрерывный ряд.***

Найдем наибольшую x_{max} = 189 см и наименьшую x_{min} = 166 см варианты.

Вычислим ширину Δx интервала группировки по формуле:

$$\Delta x = \frac{x_{\max} - x_{\min}}{c} = \frac{189 - 166}{1 + 3{,}32 \lg 56} \approx 4 \text{ см},$$

где c – количество интервалов (с = 1+ 3,322lgn).

Сгруппируем данные в ***интервальный ряд***, взяв в качестве левой границы первого интервала величину:

$$a_0 = x_{\min} - \frac{\Delta x}{2} = 166 - \frac{4}{2} = 164 \ см. \ .$$

Проведем разноску вариант по интервалам с помощью следующей таблицы:

<table>
<tr><th>№</th><th>Интервалы</th><th>Середины интервалов</th><th>Разноска вариант</th><th>Частоты</th></tr>
<tr><td>1</td><td>[164;168]</td><td>166</td><td>| | | | |</td><td>5</td></tr>
<tr><td>2</td><td>[168;172]</td><td>170</td><td>| | | | | | | |
| | | | |</td><td>13</td></tr>
<tr><td>3</td><td>[172;176]</td><td>174</td><td>| | | | | | | |
| | | | | | | | |</td><td>16</td></tr>
<tr><td>4</td><td>[176;180]</td><td>175</td><td>| | | | | | | |
| | | | |</td><td>13</td></tr>
<tr><td>5</td><td>[180;184]</td><td>182</td><td>| | | |</td><td>4</td></tr>
<tr><td>6</td><td>[184;188]</td><td>186</td><td>| | | |</td><td>4</td></tr>
<tr><td>7</td><td>[188;192]</td><td>190</td><td>|</td><td>1</td></tr>
</table>

Итак, получаем интервальный ряд следующего вида:

№	Интервалы	Середины интервалов x_i	Частоты m_i	Относительные частоты w_i	Высота h_i
1	[164;168]	166	5	0,09	1,25
2	[168;172]	170	13	0,23	3,25
3	[172;176]	174	16	0,29	4,00
4	[176;180]	175	13	0,23	3,25
5	[180;184]	182	4	0,07	1,00
6	[184;188]	186	4	0,07	1,00
7	[188;192]	190	1	0,02	0,25
	Всего		56	1	

3. Графическое представление ряда

В качестве геометрической интерпретации ряда используют гистограмму частот (относительных частот) для интервального ряда.

Рис. 4. Гистограмма частот

4. Вычисление выборочных характеристик

Вычисление выборочных характеристик удобно вести с помощью расчетной таблицы:

Расчетная таблица

№	Интервалы	x_i	m_i	$x_i m_i$	$x_i^2 m_i$	$x_i - \bar{x}_в$	$(x_i - \bar{x}_в)m_i$	$(x_i - \bar{x}_в)^2 m_i$
1	[164;168]	166	5	830	137780	–9	–45	405
2	[168;172]	170	13	2210	375700	–5	–65	325
3	[172;176]	174	16	2784	484416	–1	–16	16
4	[176;180]	178	13	2314	411892	3	39	117
5	[180;184]	182	4	728	132496	7	28	196
6	[184;188]	186	4	744	138384	11	44	484
7	[188;192]	190	1	190	36100	15	15	225
Итого			56	9800	1716768		0	1768

$$m_j^* = \frac{1}{n}\sum_{i=1}^{k} x_i^j * m_i$$

$$\bar{\boldsymbol{x}}_в = 175$$

$m_1^* = \frac{1}{56} * 9800 = 175,$

$m_2^* = \frac{1}{56}$*1716768=30656,57

$$\alpha_j^* = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x}_в)^j * m_i$$

$\alpha_1^* = \frac{1}{56}$*0 = 0,

$\alpha_2^* = \frac{1}{56} * 1768 = 31{,}57 = D_в$

$S^2 = \frac{n}{n-1} * D_в = 1{,}02 * 31{,}57 = 32{,}2$

$S = \sqrt{S^2} = \sqrt{32{,}2} = 5{,}67$

$D = \overline{x^2} - (\bar{x}\,)^2,\ \overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i{}^2 * m_i = 30656{,}57$

$D = 30656{,}57 - 30625 = 31{,}57$

$\sigma_в = \sqrt{D_в} = \sqrt{31{,}57} = 5{,}62$

M_e = 178, M_0 = 174

V = $\frac{\sigma_b}{\bar{\bar{x}}_b}$ 100% = $\sigma_в / \bar{\boldsymbol{x}}_в * 100\%$=5,62/175*100% = 3,21%

5. Эмпирическая функция распределения

Для построения графика эмпирической функции $F*(x)$ составим таблицу:

x_i	∞<x<164	168	172	176	180	184	188	192	x >192
$\omega_i^{нак}$	0	0,16	0,32	0,46	0,78	0,85	0,94	0,99	1

График эмпирической функции $F^*(x)$ распределения непрерывной случайной величины представлен на рис. 6:

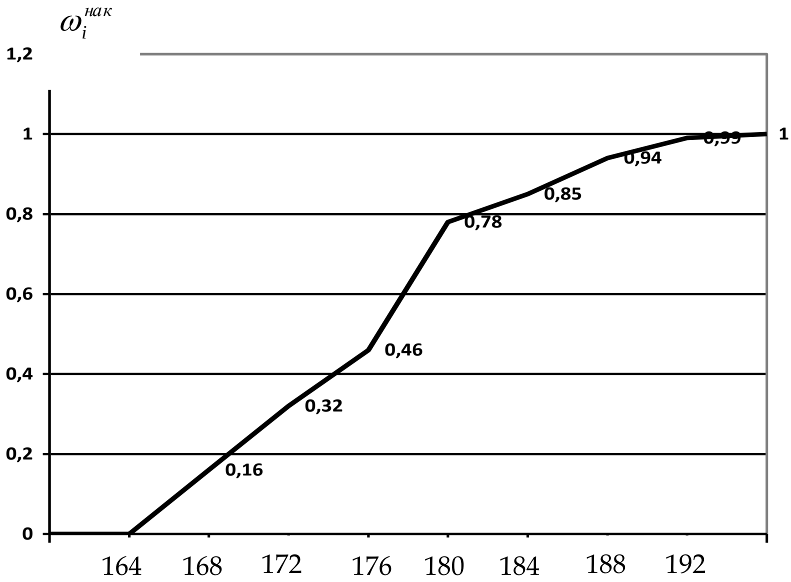

Рис. 6. График эмпирической функции $F^*(x)$ распределения непрерывной случайной величины

1.4. Задания к лабораторно-практической работе

Во всех указанных ниже вариантах заданий требуется:

1. Провести статистическую обработку данных:

a) составить вариационный ряд;

b) определить m_i, ω_i;

c) составить выборочный ряд распределения.

2. Изобразить полученный ряд графически.

3. Вычислить выборочные характеристики.

4. Найти эмпирическую функцию по распределению выборки и изобразить ее графически.

Вариант 1

Анализируются объемы ежедневных продаж некоторого товара за 50 дней. Получены следующие данные:

20	15	15	10	11	15	11	11
22	20	15	11	12	14	9	22
14	12	11	12	25	22	20	14
25	22	20	5	14	12	1	9
15	14	22	25	15	22	20	20

Вариант 2

Анализируется продолжительность телефонных разговоров с клиентами некоторой справочной телефонной службы. Случайным образом отобраны 60 телефонных разговоров и зафиксирована их длительность (в секундах):

39, 60, 40, 52, 32, 68, 77, 61, 68, 60, 47, 49, 70, 55, 66, 80, 35, 67, 70, 55, 42, 52, 60, 82, 70, 55, 47, 39, 50, 58, 45, 50, 53, 33, 49, 54, 55, 70, 62, 60, 60, 40, 59, 64, 70, 55, 54, 35, 48, 52, 57, 55, 82, 70, 51, 35, 49, 60, 55, 47.

Вариант 3

В случайном порядке отобрано 50 личных карточек студентов и выписаны их экзаменационные оценки по математике:

3	3	3	4	4	3	2	4	5	4
2	5	2	5	4	2	3	3	3	3
3	3	3	4	3	4	3	3	5	3
5	3	2	2	3	4	4	4	3	4
4	3	4	5	3	3	2	4	5	2

Вариант 4

Получены сведения от 50 случайно отобранных студентов о затратах времени (в часах) на самостоятельную работу в течение недели:

5,0	5,5	4,0	4,8	6,0	10,0	6,0	15,0	2,0	0,0	1,0	10,5	2,2
4,0	7,0	6,0	3,8	9,0	3,5	1,2	5,0	6,5	4,2	1,0	0,5	11,0
3,0	8,0	3,5	4,5	12,0	10,0	5,5	6,0	7,5	8,5	8,5	18,0	6,0
8,0	9,0	4,5	6,0	8,5	1,5	5,0	14,0	10,0	4,0			

Вариант 5

В супермаркете проводились наблюдения над числом X покупателей, обратившихся в кассу за один час. Наблюдения в течение 30 часов дали следующие результаты:

70, 75, 100, 120, 75, 60, 100, 120, 70, 60, 65, 100, 65, 100, 70, 75, 60, 100, 100, 120, 70, 75, 70, 120, 65, 70,75, 70, 100, 100.

Вариант 6

Измерялось время выполнения определенной операции. Наблюдения проводились за 50 операторами и были получены следующие результаты (в секундах):

38	60	41	51	33	42	45	21	53	60
68	52	47	46	49	49	41	57	54	59
77	47	28	48	58	32	42	58	61	30
61	35	47	72	41	45	44	55	30	40
67	65	39	48	43	60	54	42	59	50

Вариант 7

Выкопано 50 кустов картофеля и сосчитано количество клубней в каждом кусте. Получены следующие результаты:

7	11	10	9	12	9	7	11	10	12
3	5	6	7	9	6	3	5	6	10
7	9	8	9	8	9	7	8	10	8
9	8	7	8	5	7	9	9	7	7
4	4	10	10	7	8	4	5	8	10

Вариант 8

Изучается время безотказной работы устройств определенного типа. С этой целью произведена случайная выборка объёма 50, которая дала следующее результаты (в часах):

13,4	14,7	14,7	15,2	13,0	8,8	14,0	17,9	15,1	16,5
14,2	16,3	14,6	11,7	16,4	15,1	17,6	14,1	18,8	11,6
18,0	12,4	17,2	14,5	16,3	13,7	15,5	16,2	8,4	14,7
11,3	10,7	16,9	15,8	16,1	12,3	14,0	17,7	14,7	16,2
10,1	15,8	18,3	17,5	12,7	20,7	13,5	14,0	15,7	21,9

Вариант 9

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.

Вариант 10

Измерена техническая длина стебля (в см) у 50 растений льна и получены следующие результаты:

90,1	76,2	79,9	45,4	70,7	79,1	98,2	92,1	89,4	85,4
109,9	82,2	81,4	59,1	68,5	67,0	77,0	76,1	91,5	93,1
99,1	80,0	84,0	60,1	80,7	100,4	78,0	88,1	75,7	79,0
115,2	69,4	83,3	78,2	84,4	72,4	83,9	94,1	78,5	93,0
68,0	74,4	81,7	97,0	77,0	69,0	81,3	82,0	84,4	81,4

Вариант 11

Взвешены 50 клубней картофеля, случайно отобранных, и получены следующие результаты (в граммах):

93	159	253	152	80	96	156	188	208	200
77	180	144	206	191	180	140	213	53	81
48	220	216	200	117	118	145	181	77	120
135	135	185	145	110	134	138	251	188	131
109	111	150	150	197	206	113	109	142	120

Вариант 12

Анализируется продолжительность телефонных разговоров с клиентами некоторой справочной телефонной службы. Случайным образом отобраны 50 телефонных разговоров и зафиксированы их длительность (в секундах):

60	40	59	64	70	55	54	35	48	52

57	55	82	70	51	35	49	60	55	39
60	52	32	68	77	61	68	60	47	49
70	55	66	80	35	67	60	82	70	55
47	58	45	50	33	49	54	55	70	62

Вариант 13

Среднемесячная заработанная плата случайно выбранных 50 работников предприятия за истекший год составила (ден. ед.):

218	220	170	276	162	204	144	102	144	185
205	199	250	158	192	253	155	189	169	175
130	180	276	110	122	138	201	170	171	185
188	191	184	101	143	150	157	102	266	182
190	177	180	197	125	280	280	160	168	115

Вариант 14

При измерении роста 50 новобранцев получены следующие данные (в см):

177	173	171	179	170	185	182	175	166	170
178	172	182	176	182	180	164	167	172	177
180	176	174	176	175	174	170	171	165	178
168	175	179	164	170	172	182	174	165	175
183	185	176	180	174	177	176	170	180	172

Вариант 15

На 50 одинаковых по площади участках, случайно выбранных на большом массиве гречихи, определена следующая урожайность (в ц/га):

4,3	5,1	6,8	4,1	8,3	7,1	10,5	12,1	7,5	19,2
6,9	8,5	6,8	7,1	8,3	9,2	6,9	7,1	8,3	7,9

7,7	8,4	8,8	6,9	9,3	12,3	8,4	7,9	7,6	8,1
9,2	7,3	7,1	7,5	8,1	8,6	9,9	9,4	8,5	8,4
8,7	7,1	7,9	7,1	7,5	13,2	14,3	12,1	11,5	12,4

Вариант 16

На 50 одинаковых по площади участках определена урожайность яровой пшеницы (в ц/га):

13,9	11,6	15,1	10,7	14,0	11,7	11,2	10,9	15,0	11,3
12,4	10,5	11,7	8,2	12,5	10,6	15,0	10,6	11,6	15,1
13,1	10,4	11,3	10,2	13,2	10,5	11,6	8,1	10,5	11,7
6,3	10,6	10,2	15,1	6,4	10,5	11,2	9,5	10,4	11,3
11,8	11,3	11,0	9,6	11,9	10,7	10,1	10,1	10,6	10,2

Вариант 17

В ходе проведения эксперимента получен следующий набор данных:

32, 26, 16, 44, 28, 40, 30, 31, 17, 30, 37, 32, 42, 31, 36, 49,35, 21, 25, 40, 27, 25, 33, 34, 27, 43, 19, 23, 36, 48, 31, 35, 43, 32, 26, 35, 33, 45, 19, 22, 28, 49, 23, 32, 33, 27, 43, 35, 23, 44.

Вариант 18

В городе А для определения сроков гарантийного обслуживания проведено исследование величины среднего пробега автомобилей, находящихся в эксплуатации в течение двух лет с момента продажи автомобиля магазином. Получен следующий результат (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	18,8
11,2	9,9	25,3	4,2	29,6					

Вариант 19

В таблице приведена выборка результатов измерения роста 105 юношей. Измерения проводились с точностью до 1 см.

155	170	185	180	188	152	173	178	178	168	185
173	170	183	175	173	170	183	175	180	175	193
178	183	180	197	178	181	187	168	174	179	184
183	178	189	178	163	166	178	175	182	190	167
170	178	183	170	178	181	173	168	185	175	170
155	169	186	179	189	155	174	179	179	169	186
174	171	184	175	193	178	184	180	196	175	181
188	168	179	178	183	184	178	181	177	163	166
178	175	183	190	167	170	178	183	170	178	182
173	168	186	176	171	188					

Вариант 20

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

0	0	0	1	0	5	0	10	0	0
0	1	1	1	0	1	0	1	0	0
5	0	5	0	1	1	10	1	5	1
1	0	0	0	5	1	0	1	1	0
0	5	0	0	0	0	1	0	10	1

ЛАБОРАТОРНО-ПРАКТИЧЕСКАЯ РАБОТА 2
СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ. МЕТОДЫ ТОЧЕЧНОГО ОЦЕНИВАНИЯ

Цель работы: изучить основные методы получения точечных оценок: метод моментов Пирсона; метод наибольшего правдоподобия.

2.1 Теоретические вопросы

1. Основные принципы точечного оценивания.
2. Метод моментов Пирсона.
3. Метод наибольшего правдоподобия.

2.2 Методические рекомендации

Оценки параметров генеральной совокупности, полученные на основе выборки, называются *статистическими*. Если статистическая оценка характеризуется одним числом, то она называется ***точечной.*** К числу таких оценок относятся выборочная средняя и выборочная дисперсии.

Выборочная средняя является оценкой математического ожидания случайной величины и представляет собой *несмещенную оценку*. Выборочная дисперсия оценивает дисперсию генеральной совокупности и является *смещенной оценкой*. «Исправленная» выборочная дисперсия является несмещенной оценкой.

При заданном виде закона распределения случайной величины X неизвестные параметры этого распределения можно оценить ***методом моментов.*** Этот метод состоит в том, что приравниваются соответствующие теоретические и эмпирические моменты, и из полученных уравнений находятся оценки параметров.

В случае одного параметра в теоретическом распределении для его оценки достаточно составить одно уравнение, приравняв начальный теоретический момент первого порядка ν_1 к начальному эмпирическому моменту первого порядка M_1, т.е. $\nu_1 = M_1$. Так как ν_1 = M(X) – математическое ожидание случайной величины X, а эмпирический момент первого порядка $M_1 = \bar{x}_в$, то получаем уравнение: $\boldsymbol{M(X) = \bar{x}_в}$.

Для оценки двух параметров закона распределения необходимо составить два уравнения относительно этих параметров. Приравниваем начальный теоретический момент первого порядка к начальному эмпирическому моменту первого порядка и центральный теоретический момент второго порядка к центральному эмпирическому моменту второго порядка, получаем два уравнения: $\nu_1 = M_1$ *и* $\mu_2 = m_2$.

Так как $M_1 = \bar{x}_в$, $m_2 = D_в$, а $\nu_1 = M(X)$, $\mu_2 = D(X)$, то получим систему уравнений с двумя неизвестными параметрами:

$$\begin{cases} M(X) = \bar{x}_в \\ D(X) = D_в \end{cases},$$

Решив эту систему относительно неизвестных параметров, получим их точечные оценки.

В качестве точечной оценки параметра Θ, найденной ***методом наибольшего правдоподобия***, принимают такие значения $\Theta^* = \Theta^*(x_1, x_2, ..., x_n)$, при которых **функция**

правдоподобия достигает максимума. При этом оценку Θ^* называют оценкой наибольшего правдоподобия.

Функцией правдоподобия *дискретной случайной величины X* называют функцию аргумента Θ:

$$L(x_1, x_2, ..., x_n, \Theta) = p(x_1; \Theta)*p(x_2, \Theta)*...*p(x_n, \Theta),$$

где $x_1, x_2, ..., x_n$ – варианты выборки;

Θ – параметр, для которого находится оценка;

$p(x_i, \Theta)$ – вероятность события X = x_i, зависящая от параметра Θ.

Функцией правдоподобия *непрерывной случайной величины X* называют функцию аргумента Θ:

$$L(x_1, x_2, ..., x_n, \Theta) = f(x_1; \Theta)*f(x_2, \Theta)*...*f(x_n, \Theta).$$

где $f(x_i, \Theta)$ – заданная функция плотности вероятности в точках x_i.

Так как функции *L* и *lnL* достигают максимума при одном и том же значении Θ, то обычно точки экстремума находятся для *lnL*. Для этого определяется производная $\frac{d \ln L}{d\Theta}$ и приравнивается к нулю. На основании достаточного условия (вторая производная должна быть отрицательна) можно убедиться, что полученная точка является точкой максимума.

2.3 Решение типовых задач

Задача 1. На предприятии изготавливается определенный вид продукции. Ежемесячный объем выпуска этой продукции является случайной величиной, для характеристики которой принят показательный закон распределения:

$$f(x) = \lambda e^{-\lambda x} \quad (x \geq 0).$$

В течение шести месяцев проводился замер объемов выпуска продукции, после которого получены следующие данные:

Месяц	1	2	3	4	5	6
Объем выпуска	20	24	25	28	27	32

Найти оценку параметра λ.

Решение: Так как закон распределения содержит лишь один параметр λ, то для его оценки требуется составить одно уравнение: ν_1 = M$_1$ или M(X) = $\bar{x}_в$.

Находим выборочную среднюю:

$$\overline{x_в} = \frac{1}{n}\sum_{1}^{n} x_i = 1/6\ (20+24+25+28+27+32) = 26.$$

Определяем математическое ожидание:

$$M(X) = \int_0^\infty xf(x)dx = \lambda \int_0^\infty xe^{-\lambda x}dx = \frac{1}{\lambda}.$$

Следовательно, $\frac{1}{\lambda} = \overline{x_в}$. Это равенство является приближенным, так как правая часть его является случайной величиной, поэтому из этого уравнения получается не точное значение λ, а его оценка λ^*:

$$\frac{1}{\lambda^*} = \overline{x_в}.$$

Итак, $\frac{1}{\lambda^*} = 26$, откуда $\lambda^* = \frac{1}{26}$.

Задача 2. Найти методом наибольшего правдоподобия оценку параметров а и σ нормального распределения с плотностью:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-a)^2}{2\sigma^2}}.$$

Решение: Закон распределения непрерывной случайной величины содержит два параметра *a* и σ.

Для непрерывной случайной величины функция правдоподобия принимает вид:

$$L=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_1-a)^2}{2\sigma^2}}*\ldots*\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_n-a)^2}{2\sigma^2}}=$$

$$=\frac{1}{\sigma(\sqrt{2\pi})^n}e^{-\frac{\sum(x_i-a)^2}{2\sigma^2}}.$$

Найдем функцию *lnL*:

$$\ln L=\ln(\sigma^{-n}(\sqrt{2\pi})^{-n})+\ln e^{-\frac{\sum(x_i-a)^2}{2\sigma^2}}=-n\ln\sigma-n\ln\sqrt{2\pi}-\frac{\sum(x_i-a)^2}{2\sigma^2}.$$

Для отыскания ее максимума составляют и решают систему:

$$\begin{cases}\dfrac{d\ln L}{da}=0\\[2mm]\dfrac{d\ln L}{d\sigma}=0\end{cases}$$

$$\frac{d\ln L}{da}=-\left(\frac{\sum(x^2{}_i-2x_ia+a^2)}{2\sigma^2}\right)'=\frac{1}{2\sigma^2}\left(-\sum(2x_i-2a)\right)=\frac{1}{\sigma^2}\sum(x_i-a)=\frac{\sum x_i-na}{\sigma^2}$$

.

$$\frac{d\ln L}{d\sigma}=-\frac{n}{\sigma}+\frac{\sum(x_ia)^2}{\sigma^3}.$$

$$\begin{cases} \dfrac{\sum (x_i - a)^2}{\sigma^3} - \dfrac{n}{\sigma} = 0 \\ \dfrac{\sum x_i - na}{\sigma^2} = 0 \end{cases} \qquad \left\{ \dfrac{\sum x_i - na}{\sigma^2} = 0 \Rightarrow \sum x_i = na \Rightarrow a = \dfrac{\sum x_i}{n} = x_в \right. .$$

$$\frac{\sum (x_i - \bar{x}_в)^2}{\sigma^3} - \frac{n}{\sigma} = 0 \qquad \frac{\sum (x_i - \bar{x}_в)^2 - n\sigma^2}{\sigma^3} = 0 .$$

$$\sum (x_i - \bar{x}_в)^2 = n\sigma^2 . \qquad \sigma^2 = \frac{\sum (x_i - \bar{x}_в)^2}{n} = D_в . \qquad \sigma = \sqrt{D_в} .$$

$$\begin{cases} \dfrac{\sum (x_i - a)^2}{\sigma^3} - \dfrac{n}{\sigma} = 0 \\ \dfrac{\sum x_i - na}{\sigma^2} = 0 \end{cases}$$

Таким образом, получены оценки: $\begin{cases} a^* = \bar{x}_в \\ \sigma^* = \sqrt{D_в} \end{cases}$.

2.4 Задания к лабораторно-практической работе

Вариант 1

При условии показательного распределения случайной величины X

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{если } x \geq 0 \\ 0, & \text{если } x < 0 \end{cases}.$$

Произведена выборка:

x_i	4	3	10	12	15
m_i	3	3	6	4	4

Найти методом моментов оценку параметра λ .

Вариант 2

Случайная величина X задана функцией распределения:

$$F(X) = 1 - e^{-\lambda x} \quad (x \geq 0).$$

Произведена выборка:

x_i	3	5	6	8	10
m_i	2	3	5	10	10

Найти методом моментов оценку параметра λ .

Вариант 3

При условии равномерного распределения случайной величины X:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{если } x \in (a;b) \\ 0, & \text{если } x \notin (a;b) \end{cases}$$

Произведена выборка:

x_i	2	3	4	5	6
m_i	4	6	5	12	8

Найти методом моментов оценку параметров *a* и *b*.

Вариант 4

При условии равномерного распределения случайной величины X произведена выборка:

x_i	3	5	7	9	11	13	15	17	19	21
m_i	21	16	15	26	22	14	21	22	18	25

Найти методом моментов оценку параметров *a* и *b*.

Вариант 5

Случайная величина подчиняется нормальному закону распределения:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-a)^2}{2\sigma^2}}.$$

Произведена выборка:

x_i	3	5	7	9	11	13	15
m_i	6	9	16	25	20	16	8

Найти методом моментов оценку параметров a и σ.

Вариант 6

Случайная величина X (число семян сорняков в пробе зерна) распределена по закону Пуассона: $P_m(x_i)=\frac{\lambda^{x_i}e^{-\lambda}}{x_i!}$ (x_i – число появления события в i-ом опыте; m – число испытаний, произведенных в одном опыте). Ниже приведено эмпирическое распределение семян сорняков в 1000 пробах зерна (в первой строке указано количество x_i сорняков в одной пробе; во второй строке указана частота m_i – число проб, содержащих x_i семян сорняков):

x_i	0	1	2	3	4	5	6
m_i	405	366	175	40	8	4	2

Найти методом моментов точечную оценку параметра λ распределения Пуассона.

Вариант 7

Случайная величина X (число нестандартных изделий в партии изделий) распределена по закону Пуассона. Ниже приведено распределение нестандартных изделий в 200 партиях (в первой строке указано количество x_i нестандартных изделий в одной партии; во второй строке указана частота m_i – число партий, содержащих x_i нестандартных изделий):

x_i	0	1	2	3	4
m_i	132	43	20	3	2

Найти методом моментов точечную оценку параметра λ распределения Пуассона.

Вариант 8

Случайная величина X (число появлений события A в r независимых испытаниях) подчинена биномиальному закону распределения с неизвестным параметром *p:*

$$P_r(x) = C_r^{x_i} p^{x_i} (1-p)^{r-x_i}.$$

Ниже приведено эмпирическое распределение числа появления события в 10 опытах по 5 испытаний в каждом (в первой строке указано число x_i появления события A в одном опыте; во второй строке указана частота m_i – количество опытов, в которых наблюдалось x_i появлений события A):

x_i	0	1	2	3	4
m_i	5	2	1	1	1

Найти методом моментов точечную оценку параметра p биномиального распределения.

Вариант 9

Случайная величина X (время работы элемента) имеет показательное распределение:

$$f(x) = \lambda e^{-\lambda x} \quad (x \geq 0).$$

Ниже приведено эмпирическое распределение среднего времени работы 200 элементов (в первой строке приведено среднее время x_i работы элемента в часах; во второй строке указана частота m_i – количество элементов, проработавших в среднем x_i часов):

x_i	2,5	7,5	12,5	17,5	22,5	27,5
m_i	133	45	15	4	2	1

Найти методом моментов точечную оценку неизвестного параметра показательного распределения.

Вариант 10

Случайная величина X (отклонение контролируемого размера изделия от номинала) подчинена нормальному закону распределения с неизвестными параметрами a и σ:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-a)^2}{2\sigma^2}}.$$

Ниже приведено эмпирическое распределение отклонения от номинала 200 изделий (в первой строке указано отклонение x_i (мм); во второй строке приведена частота m_i – количество изделий, имеющих отклонение x_i):

x_i	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,2	2,3
m_i	6	9	26	25	30	26	21	24	20	8	5

Найти методом моментов точечные оценки неизвестных параметров a и σ нормального распределения.

Вариант 11

Случайная величина X (ошибка измерения дальности радиодальномером) подчинена равномерному закону распределения с неизвестными параметрами *a и b:*

$$f(x) = \frac{1}{b-a}, \ (b>a).$$

Ниже приведено эмпирическое распределение средней ошибки 200 измерений дальности (в первой строке указана средняя ошибка x_i; во второй строке указана частота m_i – количество измерений, имеющих среднюю ошибку x_i):

x_i	3	5	7	9	11	13	15	17	19	21
m_i	21	16	15	26	22	14	21	22	18	25

Найти методом моментов точечные оценки неизвестных параметров a и b равномерного распределения.

Вариант 12

Случайная величина X распределена по показательному закону. Статистическое распределение выборки представлено в таблице:

x_i	5	15	25	35	45	55	65
m_i	365	245	150	100	70	45	25

Найдите методом наибольшего правдоподобия точечную оценку параметра λ.

Вариант 13

Стеклянные однородные изделия отправлены для реализации из Москвы в Новосибирск в 1000 контейнерах. После поступления товара было выявлено количество разбитых изделий в каждом контейнере. Результаты представлены в таблице:

x_i	0	1	2	3	4
m_i	785	163	32	16	4

Считая, что число разбитых изделий описывается законом Пуассона, найти методом наибольшего правдоподобия точечную оценку параметра λ.

Вариант 14

Случайная величина X распределена по закону Пуассона с неизвестным параметром λ. Статистическое распределение выборки представлено в таблице:

x_i	0	1	2	3	4	5	6	7
m_i	199	169	87	31	9	3	1	1

Найдите методом наибольшего правдоподобия точечную оценку параметра λ.

Вариант 15

Случайная величина X распределена по биномиальному закону:

$$P_r(x) = C_r^{x_i} p^{x_i} q^{r-x_i}.$$

Статистическое распределение выборки представлено в таблице:

x_i	0	1	2	3	4	5	6	7
m_i	2	3	10	22	26	20	12	5

Найдите методом наибольшего правдоподобия точечную оценку параметра p указанного закона распределения случайной величины (r = 10 – количество испытаний в одном опыте; x_i – число появлений события в i-ом опыте).

Вариант 16

Случайная величина X (число появлений события A в r независимых испытаниях) подчинена биномиальному закону распределения с неизвестным параметром *p:*

$$P_r(x) = C_r^{x_i} p^{x_i} (1-p)^{r-x_i}.$$

Ниже приведено эмпирическое распределение числа появления события A в 1000 испытаниях (в первой строке указано число x_i появления события в одном опыте из r = 10 испытаний; во второй строке приведена частота m_i – число опытов, в которых наблюдалось x_i появлений события A):

x_i	0	1	2	3	4	5	6	7
m_i	2	3	10	22	26	20	12	5

Найти методом наибольшего правдоподобия точечную оценку неизвестного параметра p биномиального распределения.

Вариант 17

Случайная величина X (число поврежденных стеклянных изделий в одном контейнере) распределена по закону Пуассона с неизвестным параметром λ :

$$P_m(x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}.$$

(x_i – число появления события в i-ом опыте; m – число испытаний, произведенных в одном опыте). Ниже приведено эмпирическое распределение числа поврежденных изделий в 500 контейнерах (в первой строке указано количество x_i поврежденных изделий в одном контейнере; во второй строке приведена частота m_i – число контейнеров, содержащих x_i поврежденных изделий):

x_i	0	1	2	3	4	5	6	7
m_i	199	169	87	31	9	3	1	1

Найти методом наибольшего правдоподобия точечную оценку неизвестного параметра λ распределения Пуассона.

Вариант 18

Случайная величина X (время безотказной работы элемента) имеет показательное распределение:

$$f(x) = \lambda e^{-\lambda x} \quad (x \geq 0).$$

Ниже приведено статистическое распределение среднего времени работы 1000 элементов (в первой строке указано среднее время x_i безотказной работы одного элемента в часах; во второй строке указана частота m_i – количество элементов, проработавших в среднем x_i часов):

x_i	5	15	25	35	45	55	65
m_i	365	245	150	100	70	45	25

Найдите методом наибольшего правдоподобия точечную оценку неизвестного параметра λ показательного распределения.

Вариант 19

В таблице собраны результаты наблюдений над промежутками времени между последовательными прибытиями судов дальнего следования в порт А (в первой строке указаны промежутки времени (час) x_i прибытия судов; во второй строке указана частота m_i – число случаев):

x_i	0 – 4	4 – 8	8 – 12	12 – 16	16 – 20	20 – 24	24 – 28	28 – 32
m_i	67	43	30	18	11	7	5	4

Предполагается, что случайная величина X подчинена нормальному закону распределения с неизвестными параметрами a и σ:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-a)^2}{2\sigma^2}}.$$

Найти методом максимального правдоподобия точечные оценки неизвестных параметров a и σ нормального распределения.

Вариант 20

В таблице собраны результаты наблюдений о годовом удое 100 коров, отобранных случайным образом из генеральной совокупности X (в первой строке указаны группы x_i по удою молока (тыс. л.); во второй строке указана частота m_i – число коров):

x_i	1,6 – 2,2	2,2 – 2,8	2,8 – 3,4	3,4 – 4,0	4,0 – 4,6	4,6 – 5,2	5,2 – 5,8	5,8 – 6,4
m_i	4	14	17	37	15	6	4	3

Предполагается, что случайная величина X подчинена нормальному закону распределения с неизвестными параметрами a и σ:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x_i-a)^2}{2\sigma^2}}.$$

Найти методом максимального правдоподобия точечные оценки неизвестных параметров a и σ нормального распределения.

ЛАБОРАТОРНО-ПРАКТИЧЕСКАЯ РАБОТА 3
СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

Цель работы: освоить основные методы получения интервальных оценок параметров распределения.

3.1. Теоретические вопросы

1. Интервальная оценка.
2. Доверительный интервал. Доверительная вероятность.
3. Доверительный интервал для математического ожидания нормально распределенной случайной величины с параметрами a и σ при известном σ.
4. Доверительный интервал для математического ожидания нормально распределенной случайной величины с параметрами a и σ при неизвестном σ.
5. Доверительный интервал для среднего квадратического отклонения σ нормально распределенной случайной величины по «исправленному» выборочному среднему квадратическому отклонению s.
6. Доверительный интервал для дисперсии σ^2 нормально распределенной случайной величины с известными параметрами a и σ.

3.2. Методические рекомендации

Статистической оценкой θ_n^* неизвестного параметра θ теоретического распределения называют всякую функцию результатов наблюдений $x_1, x_2, \ldots, x_n$ исследуемой случайной величины X.

Точечной называют статистическую оценку, которая определяется одним числом θ^*.

Интервальная оценка – это оценка, которая определяется двумя числами – концами интервала. Она позволяет ответить на вопрос: внутри какого интервала, и с какой вероятностью находится неизвестное значение оцениваемого параметра θ генеральной совокупности?

Доверительной вероятностью (надежностью) или уровнем доверия оценки неизвестного параметра θ по θ^* называют вероятность $\gamma = 1-\alpha$, с которой осуществляется неравенство $|\theta-\theta^*|<\delta$ (где $(\delta>0)$ δ – *точность оценки*), т.е. $P(|\theta-\theta^*|<\delta)=\gamma$ или $P(\theta^*-\delta<\theta<\theta^*+\delta)=\gamma$.

Интервал $(\theta^*-\delta;\theta^*+\delta)$, в который попадает оцениваемый параметр θ с заданной надежностью γ, называется ***доверительным интервалом*** для параметра θ.

Построение интервальных оценок (доверительных интервалов)

Как же конкретно построить по выборочным данным $x_1, \ldots, x_n$ такой случайный интервал, который с наперед заданной доверительной вероятностью γ накрывал бы неизвестной значение параметра θ?

Очевидно, что этот интервал должен конструироваться вокруг *точечной* оценки θ^* параметра θ, а её точечный вид и ширина определяется характером закона распределения случайной величины θ^*, в частности её функцией распределения, которая, к сожалению, тоже зависит от неизвестного значения параметра θ.

Поэтому необходимо подобрать такую функцию от результатов наблюдения $x_1, x_2, \ldots, x_n$ (т.е. такую статистику), закон распределения вероятностей которой обладал бы одновременно следующими свойствами:

а) не зависит от оцениваемого параметра θ;

б) описывается одним из стандартных, затабулированных распределений (стандартным нормальным, χ^2, F-, Стьюдента);

в) из того факта, что значения данной статистики заключены в определенных пределах с заданной вероятностью, можно сделать вывод, что оцениваемый параметр тоже должен лежать между некоторыми границами с той же самой вероятностью.

Этот подход, если его удается реализовать, приводит к построению *точных* (при каждом конечном объеме выборки n) *доверительных интервалов* (θ_1, θ_2).

Интервальной оценкой (с надежностью γ) *математического ожидания а* нормально распределенной случайной величины с параметрами *а и* σ по выборочной средней $\bar{x}$ ***при известном*** *среднем квадратическом отклонении* σ генеральной совокупности служит доверительный интервал:

$$\bar{x} - \frac{t_\gamma \sigma}{\sqrt{n}} < a < \bar{x} + \frac{t_\gamma \sigma}{\sqrt{n}},$$

где $\delta = \dfrac{t_\gamma \sigma}{\sqrt{n}}$ – точность оценки;

n – объем выборки;

t_γ – значение аргумента функции Лапласа $\Phi(t_\gamma)$ (приложение 3), при котором $2\Phi(t_\gamma) = \gamma \Rightarrow \Phi(t_\gamma) = \frac{\gamma}{2}$.

Интервальной оценкой (с надежностью γ) *математического ожидания a* нормально распределенной случайной величины с параметрами *а и* σ по выборочной средней $\bar{x}$ ***при неизвестном*** *среднем квадратическом отклонении* σ генеральной совокупности служит доверительный интервал*:*

$$\bar{x} - \frac{t_\gamma S}{\sqrt{n}} < a < \bar{x} + \frac{t_\gamma S}{\sqrt{n}},$$

где $\delta = \frac{t_\gamma S}{\sqrt{n}}$ – точность оценки;

n – объем выборки;

t_γ – по *таблице значений* t_γ = *(γ;n)* (приложение 7), по заданным значениям n и γ находится квантиль t_γ.

Интервальной оценкой (с надежностью γ) *среднего квадратического отклонения* σ нормально распределенной случайной величины с параметрами *а и* σ по «исправленному» выборочному среднему квадратическому отклонению *s* служит доверительный интервал:

$$s(1- q) < \sigma < s(1 + q) \qquad (\text{при } q < 1),$$

или

$$0 < \sigma < s(1 + q) \qquad (\text{при } q > 1),$$

где величину q находят по таблице значений q = q(γ, *n*) (приложение 8), по заданным *n* и γ.

Интервальной оценкой (с надежностью γ) *дисперсии* σ^2 нормально распределенной случайной величины с параметрами *а и* σ с известным параметром *а* служит доверительный интервал:

$$\left\{\frac{S^2(n-1)}{\chi_2^2} < \sigma^2 < \frac{S^2(n-1)}{\chi_1^2}\right\},$$

где значения χ_1^2 и χ_2^2 находят по таблице значений $\chi^2_{\alpha,k=n-1}$ критерия Пирсона (приложение 4), из равенств: $P(\chi^2 > \chi_1^2) = \frac{1+\gamma}{2}, P(\chi^2 > \chi_2^2) = \frac{1+\gamma}{2};$

S^2 – исправленная дисперсия;
n – объем выборки.

3.3. Решение типовых задач

Задача 1. Найти доверительный интервал с надежностью 0,95 для оценки математического ожидания нормально распределенной случайной величины X, если известны ее среднее квадратическое отклонение σ = 4, выборочная средняя $\bar{x}$ = 16 и объем выборки *n* = 16.

Решение: С надежностью γ доверительный интервал: $(\bar{x} - \frac{t_\gamma \sigma}{\sqrt{n}}; \bar{x} + \frac{t_\gamma \sigma}{\sqrt{n}})$ покрывает неизвестный параметр *а*.

По надежности γ = 0,95 из соотношения $\Phi(t_\gamma)$ = $\frac{\gamma}{2}$ находим значение функции Лапласа: Ф(*t*) = 0,475.

По таблице значений функции Лапласа (приложение 3), находим t = 1,96.

Используя неравенства для интервальной оценки математического ожидания *при известном значении* σ, получаем:

$$16 - 1{,}96 \cdot \frac{4}{4} < a < 16 + 1{,}96 \cdot \frac{4}{4},$$

или

$$14{,}04 < a < 17{,}96.$$

Задача 2. По данным девяти независимых равноточных измерений некоторой физической величины найдены среднее арифметическое результатов измерений $\bar{x}$ = 30,1 и «исправленное» среднее квадратическое отклонение s = 6. Оценить истинное значение измеряемой величины с помощью доверительного интервала с надежностью γ = 0,99. Предполагается, что результаты измерений распределены нормально.

Решение: С надежностью γ доверительный интервал: $(\bar{x} - \frac{t_\gamma s}{\sqrt{n}}; \bar{x} + \frac{t_\gamma s}{\sqrt{n}})$ покрывает неизвестный параметр a.

По *таблице значений* t_γ = *(γ;n)* (приложение 7), по заданным значениям n = 9 и γ = 0,99 находим квантиль t_γ = 2,36.

Используя неравенства для интервальной оценки математического ожидания *при неизвестном значении* σ, получаем:

$$30{,}1 - 2{,}36 \cdot \frac{6}{3} < a < 30{,}1 + 1{,}36 \cdot \frac{6}{3}$$

или $$25{,}38 < a < 34{,}82.$$

Задача 3. По данным выборки объема n = 16 из генеральной совокупности найдено «исправленное» среднее квадратическое отклонение s = 1 нормально распределенного количественного признака. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надежностью 0,95.

Решение: С надежностью γ доверительный интервал: s(1 – q) < σ < s(1 + q), если q < 1, или 0 < σ < s(1 + q), если q > 1, покрывает неизвестный параметр σ.

По таблице значений q = q(γ, n) (приложение 8), по заданным n = 16 и γ = 0,95 находят величину q = 0,44. Так как q < 1, то подставив s = 1, q = 0,44, получим искомый доверительный интервал:

$$1(1 - 0{,}44) < \sigma < 1(1 + 0{,}44) \text{ или } 0{,}56 < \sigma < 1{,}44.$$

Задача 4. На основании выборочных наблюдений производительности труда 20 работниц было установлено, что среднее квадратическое отклонение суточной выработки составляет 15 м ткани в час. Предполагая, что производительность труда работницы имеет нормальное распределение, найти границы, в которых с надежностью 0,9 заключены генеральные дисперсия и среднее квадратическое отклонение суточной выработки работниц.

Решение: С надежностью γ доверительный интервал: $\left\{\dfrac{S^2(n-1)}{\chi_2^2} < \sigma^2 < \dfrac{S^2(n-1)}{\chi_1^2}\right\}$ покрывает неизвестный параметр σ^2.

По условию γ = 0,9, тогда $\frac{1+\gamma}{2} = 0{,}05, \frac{1+\gamma}{2} = 0{,}95.$ При числе свободы $k = n - 1$= 20 – 1 = 19 определим значения χ_1^2 и χ_2^2 по таблице значений χ^2 критерия Пирсона (приложение 4) для вероятностей:

$$P(\chi^2 > \chi_1^2) = \frac{1+\gamma}{2} = 0{,}05, P(\chi^2 > \chi_2^2) = \frac{1+\gamma}{2} = 0{,}95,$$

т.е. χ_1^2 = 10,1 и χ_2^2 = 30,1. Тогда доверительный интервал для σ^2 и σ для примет вид:

$$\frac{15^2 \cdot 19}{30{,}1} < \sigma^2 < \frac{15^2 19}{10{,}1} \quad \text{или} \quad 149{,}5 < \sigma^2 < 445{,}6,$$

$$\sqrt{149{,}5} < \sigma < \sqrt{445{,}6} \quad \text{или} \quad 12{,}2 < \sigma < 21{,}1.$$

3.4. Задания к лабораторно-практической работе

Вариант 1

Найти доверительный интервал для оценки с надежностью 0,95 неизвестного математического ожидания a нормально распределенного признака X генеральной совокупности, если генеральное среднее квадратическое отклонение σ = 5, выборочная средняя $\bar{x}$ = 14 и объем выборки n = 25.

Вариант 2

По наблюдениям на 13 разрезах установлено, что изучаемый вид почвы имеет среднюю мощность горизонта 21 см при среднем квадратическом отклонении 3,2 см. Указать 95% доверительный интервал для средней мощности горизонта нормально распределенного признака.

Вариант 3

Найти доверительный интервал для оценки с надежностью 0,99 неизвестного математического ожидания *a* нормально распределенного признака X генеральной совокупности, если генеральное среднее квадратическое отклонение σ = 4, выборочная средняя $\bar{x}$ = 10,2 и объем выборки n = 16.

Вариант 4

По результатам испытаний 16 двигателей были определены точечные оценки среднего ресурса 3000 ч. И среднее квадратическое отклонение 45 ч. Известно, что распределение ресурса нормальное. Построить доверительный интервал для генерального среднего с доверительной вероятностью 0,95.

Вариант 5

Одним и тем же прибором со средним квадратическим отклонением случайных ошибок измерений σ = 40 м произведено пять равноточных измерений расстояния от орудия до цели. Найти доверительный интервал для оценки истинного расстояния *a* до цели с надежностью γ = 0,95, зная среднее арифметическое результатов измерений $\bar{x}$ = 2000 м. Предполагается, что результаты измерений распределены нормально.

Вариант 6

Выборка из большой партии электроламп содержит 100 ламп. Средняя продолжительность горения лампы выборки оказалась равной 1000 ч. Найти с надежностью γ = 0,95 доверительный интервал для средней продолжительности a горения лампы всей партии, если известно, что среднее квадратическое отклонение продолжительности горения лампы σ = 40 ч. Предполагается, что продолжительность горения ламп распределена нормально.

Вариант 7

Станок-автомат штампует валики. По выборке объема n = 100 вычислена выборочная средняя диаметров изготовленных валиков. Найти с надежностью 0,95 точность δ, с которой выборочная средняя оценивает математическое ожидание диаметров изготовляемых валиков, зная, что их среднее квадратическое отклонение σ = 2 мм. Предполагается, что диаметры валиков распределены нормально.

Вариант 8

Из большой партии электроламп было отобрано в случайном порядке 400 штук для определения продолжительности горения. Выборочная средняя продолжительности горения оказалась равной $\bar{x}$ = 1220 ч. Найти с надежностью γ = 0,99 доверительный интервал для средней продолжительности a горения лампы всей партии, если известно, что среднее квадратическое отклонение продолжительности горения лампы σ = 35 ч. Предполагается, что продолжительность горения ламп распределена нормально.

Вариант 9

Найти доверительный интервал для оценки с надежностью 0,8 неизвестного математического ожидания a нормально распределенного признака X генеральной совокупности, если генеральное среднее квадратическое отклонение $\sigma = 5$, выборочная средняя $\bar{x} = 20$ и объем выборки $n = 25$.

Вариант 10

Найти доверительный интервал для оценки с надежностью 0,99 неизвестного математического ожидания a нормально распределенного признака X генеральной совокупности, если генеральное среднее квадратическое отклонение $\sigma = 5$, выборочная средняя $\bar{x} = 16{,}8$ и объем выборки $n = 25$.

Вариант 11

Из генеральной совокупности извлечена выборка объема $n = 10$:

Варианта x_i	– 2	1	2	3	4	5
Частота m_i	2	1	2	2	2	1

Оценить с надежностью 0,95 математическое ожидание a нормально распределенного признака генеральной совокупности по выборочной средней при помощи доверительного интервала.

Вариант 12

Из генеральной совокупности извлечена выборка объема $n = 12$:

Варианта x_i	– 0,5	– 0,4	– 0,2	0	0,2	0,6	0,8	1	1,2	1,5
Частота m_i	1	2	1	1	1	1	1	1	2	1

Оценить с надежностью 0,95 математическое ожидание a нормально распределенного признака генеральной совокупности по выборочной средней при помощи доверительного интервала.

Вариант 13

По данным 16 независимых равноточных измерений некоторой физической величины найдены среднее арифметическое результатов измерений $\bar{x}$ = 42,8 и «исправленное» среднее квадратическое отклонение s = 8 нормально распределенного количественного признака. Оцените истинное значение измеряемой величины с надежностью 0,999.

Вариант 14

На овцеводческой ферме из стада произведена выборка для взвешивания 36 овец. Их средний вес оказался равным 50 кг. Предположив распределение веса нормальным и определив несмещенную оценку выборочной дисперсии s^2 = 16, найти доверительный интервал для оценки математического ожидания с надежностью 0,8.

Вариант 15

Для определения средней урожайности пшеницы на площади 500000 га. Производилось выборочное измерение урожайности на 2500 га. Результаты измерений приведены в таблице:

Урожайность с 1 га (в ц)	15 – 17	17 – 19	19 – 21	21 – 23	23 – 25	25 – 27
Число гектаров	100	300	500	700	600	300

Считая распределение нормальным, указать доверительный интервал для средней урожайности по всему массиву, взяв $\gamma = 0,95$

Вариант 16

Из генеральной совокупности извлечена выборка объема $n = 16$ и найдена выборочная средняя, равная 30. Получено также несмещенное значение выборочной дисперсии $s^2 = 9$. Предположив распределение случайной величины X нормальным, найти доверительный интервал для оценки математического ожидания с надежностью 0,8.

Вариант 17

На овцеводческой ферме из стада произведена выборка для взвешивания 36 овец. Их средний вес оказался равным 50 кг. Предположив распределение веса нормальным и определив несмещенную оценку выборочной дисперсии $s^2 = 16$, найти доверительный интервал для оценки математического ожидания с надежностью 0,9.

Вариант 18

Из генеральной совокупности извлечена выборка объема $n = 16$ и найдена выборочная средняя, равная 30. Получено также несмещенное значение выборочной дисперсии $s^2 = 9$. Предположив распределение случайной величины X нормальным, найти доверительный интервал для оценки математического ожидания с надежностью 0,9.

Вариант 19

На овцеводческой ферме из стада произведена выборка для взвешивания 36 овец. Их средний вес оказался равным 50 кг. Предположив распределение веса нормальным и определив несмещенную оценку выборочной дисперсии $s^2 = 16$, найти доверительный интервал для оценки математического ожидания с надежностью 0,8.

Вариант 20

В сельхозпредприятии 30000 овец. В результате выборочного настрига шерсти с 1000 овец были получены следующие результаты:

Настриг шерсти в кг	3	4	5	6	7
Число овец	101	146	302	354	97

На уровне доверия $\gamma = 0{,}95$ построить доверительный интервал для генеральной средней, предполагая, что настриг шерсти имеет нормальное распределение.

Вариант 21

На контрольных испытаниях 16 осветительных ламп были получены оценки математического ожидания и среднеквадратического отклонения их срока службы $\bar{x} = 3000$ ч., S = 20 ч. Считая, что срок службы лампы является нормальной случайной величиной, указать на уровне доверия $\gamma = 0{,}9$ интервал, накрывающий неизвестное значение дисперсии σ^2 и σ.

Вариант 22

Для определения точности вольтметра, систематическая ошибка которого равна нулю, проведено 12 измерений напряжения и получены точечные оценки S = 0,6 мВ. Построить доверительный интервал для дисперсии на 95% уровне.

Вариант 23

По данным выборки объема n = 10 из генеральной совокупности нормально распределенного количественного признака найдено «исправленное» среднее квадратическое отклонение S = 5,1. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надежностью 0,999.

Вариант 24

По данным выборки объема n = 50 из генеральной совокупности нормально распределенного количественного признака найдено «исправленное» среднее квадратическое отклонение S = 14. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надежностью 0,999.

Вариант 25

Произведено 12 измерений одним прибором (без систематической ошибки) некоторой физической величины, причем «исправленное» среднее квадратическое отклонение S случайных ошибок оказалось равным 0,6. Найти точность прибора с надежностью 0,99. Предполагается, что результаты измерений распределены нормально.

Вариант 26

В нескольких мелких магазинах проведена проверка качества 100 изделий, после чего осуществлена обработка полученных данных. В результате получено несмещенное значение выборочного среднего квадратического отклонения $S = 4$. Считая распределение качественных изделий нормальным, найти с надежностью 0,95 доверительный интервал для оценки среднего квадратического отклонения.

Вариант 27

Случайная величина X распределена по нормальному закону. Статистическое распределение выборки представлено в таблице:

Варианта x_i	3	5	7	8	10	12	14
Частота m_i	3	7	4	6	7	5	8

Найти с надежностью 0,97 доверительный интервал для оценки математического ожидания и с надежностью 0,95 для оценки среднего квадратического отклонения.

Вариант 28

Случайная величина X распределена по нормальному закону. Статистическое распределение выборки представлено в таблице:

Варианта x_i	1	3	5	7	9
Частота m_i	2	5	4	6	3

Найти с надежностью 0,95 доверительный интервал для оценки математического ожидания и с надежностью 0,99 для оценки среднего квадратического отклонения.

Вариант 29

По данным выборки объема $n = 25$ найдено несмещенное значение выборочного среднего квадратического отклонения $S = 3$ нормально распределенной случайной величины X. Найти с надежностью 0,99 доверительный интервал для оценки среднего квадратического отклонения случайной величины.

Вариант 30

Произведено 10 измерений одним прибором (без систематической ошибки) некоторой физической величины, причем «исправленное» среднее квадратическое отклонение S случайных ошибок оказалось равным 0,8. Найти точность прибора с надежностью 0,95. Предполагается, что результаты измерений распределены нормально.

ЛАБОРАТОРНО-ПРАКТИЧЕСКАЯ РАБОТА 4
ПРОВЕРКА ПАРАМЕТРИЧЕСКИХ ГИПОТЕЗ

Цель работы: изучить основные методы проверки статистических гипотез параметрического типа.

4.1. Теоретические вопросы

1. Параметрические и непараметрические статистические гипотезы.
2. Ошибки 1-го и 2-го рода. Уровень значимости. Мощность критерия.
3. Схема проверки нулевой гипотезы.
4. Алгоритмы проверки гипотез о числовых значениях параметров нормального распределения.
5. Алгоритмы проверки гипотез о сравнении двух вероятностей биномиальных распределений.

4.2. Методические рекомендации

Статистической гипотезой называется всякое высказывание о генеральной совокупности, проверяемое по выборке.

Статистические гипотезы делятся на:

а) *параметрические* – это гипотезы, сформулированные относительно параметров распределения известного типа;

б) *непараметрические* – это гипотезы, сформулированные относительно вида неизвестного распределения.

Основную выдвигаемую гипотезу называют *нулевой (основной)* H_0. Гипотезу H_1, которая противоречит нулевой, называют *конкурирующей (альтернативной).*

Выбор между гипотезами H_0 и H_1 может сопровождаться ошибками двух родов. *Ошибка первого рода* a означает вероятность принятия H_1, если верна гипотеза H_0: $a = P(H_1/ H_0)$. *Ошибка второго рода* β означает вероятность принятия H_0, если верна гипотеза H_1: $\beta = P(H_0/ H_1)$. Существует правильное решение двух видов: вероятность принять верную гипотезу равна $P(H_0/ H_0) = 1 - a$, а вероятность отвергнуть неверную нулевую гипотезу равна $P(H_1/ H_1) = 1 - \beta$.

Правило, по которому принимается решение о том, верна или не верна гипотеза H_0, называется ***критерием***, где:

- $a = P(H_1/ H_0)$ – *уровень значимости критерия;*
- М $= 1 - \beta = P(H_1/ H_1)$ – *мощность критерия.*

Схема проверки нулевой гипотезы:

1. Располагая выборочными данными $x_1, x_2, \ldots, x_n$ и учитывая конкретные условия задачи, формируют нулевую гипотезу H_0 и конкурирующую гипотезу H_1.

2. Задают уровень значимости a (обычно 0,1; 0,01; 0,05; 0,001).

3. Рассматривается выборочная статистика наблюдений (статистический критерий) – некоторая функция К, зависящая от условий решаемой статистической задачи. Эта функция, являясь случайной величиной, подчинена некоторому известному, затабулированному закону распределения (нормальному распределению, t – распределению Стьюдента, F – распределению Фишера, χ^2 – распределению Пирсона).

4. На основании выборки $x_1, x_2, \ldots, x_n$ определяют значение критерия (статистики) К. В зависимости от вида альтернативной гипотезы выбирают по соответствующей

таблице квантили критерия для двусторонней ($K_{1-\frac{\alpha}{2}}, K_{\frac{\alpha}{2}}$) или односторонней области ($K_{1-\alpha}$ или K_{α}). Если значения критерия попадают в критическую область, то H_0 отвергается и принимается гипотеза H_1; в противном случае принимается гипотеза H_0 и считается, что H_0 не противоречит выборочным данным (при этом существует возможность ошибки с вероятностью, равной α).

Сравнение выборочной средней с гипотетической генеральной средней нормальной совокупности

а) Дисперсия генеральной совокупности известна

Правило 1. Для того чтобы при заданном уровне значимости α проверить нулевую гипотезу:

$$H_0: a = a_0$$

(о равенстве генеральной средней a нормальной совокупности с ***известной дисперсией*** σ^2 гипотетическому (предполагаемому) значению a_0) при конкурирующей гипотезе:

$$H_1: a \neq a_0,$$

надо вычислить наблюдаемое значение критерия:

$$U_{набл} = \frac{(\bar{x} - a_0)\sqrt{n}}{\sigma}$$

и по таблице функции Лапласа (приложение 3), найти критическую точку $u_{кр}$ двусторонней критической области из равенства:

$$\Phi(u_{кр}) = (1 - \alpha)/2.$$

Если $|U_{набл}| < u_{кр}$ – нет оснований отвергнуть нулевую гипотезу.

Если | $U_{набл}$ | > $u_{кр}$ – нулевую гипотезу отвергают.

***Правило* 2**. При конкурирующей гипотезе:

$$H_1: \boldsymbol{a > a_0}$$

критическую точку правосторонней критической области находят из равенства:

$$\Phi(u_{кр}) = (1 - 2\alpha)/2.$$

Если $U_{набл} < u_{кр}$ – нет оснований отвергнуть нулевую гипотезу.

Если $U_{набл} > u_{кр}$ – нулевую гипотезу отвергают.

***Правило* 3.** При конкурирующей гипотезе:

$$H_1: \boldsymbol{a < a_0}$$

сначала находят вспомогательную критическую точку $u_{кр}$ по правилу 2, а затем полагают границу левосторонней критической области: $u'_{кр} = -u_{кр}$.

Если $U_{набл} > -u_{кр}$ – нет оснований отвергнуть нулевую гипотезу.

Если $U_{набл} < -u_{кр}$ – нулевую гипотезу отвергают.

Мощность критерия проверки нулевой гипотезы H_0: $\boldsymbol{a = a_0}$ о равенстве генеральной средней гипотетическому значению a_0 при известном среднем квадратическом отклонении σ находят в зависимости от вида конкурирующей гипотезы.

При конкурирующей гипотезе H_1: $\boldsymbol{a > a_0}$ для гипотетического значения генеральной средней $\boldsymbol{a = a_1 > a_0}$ мощность правостороннего критерия:

$$1 - \beta = 0{,}5 - \Phi(u_{кр} - \lambda),$$

где $u_{кр}$ находят из равенства Ф ($u_{кр}$) = (1 – 2α)/2, λ = (a_1 – – a_0)$\sqrt{n}/\sigma$.

При различных значениях a_1 функция мощности одностороннего критерия:

$$\pi_1\ (a_1) = 0{,}5 - Ф\ (u_{кр} - \lambda).$$

При конкурирующей гипотезе H_1: $\boldsymbol{a \neq a_0}$ для гипотетического значения генеральной средней $\boldsymbol{a = a_1}$ мощность двустороннего критерия:

$$-\beta = 1 - [Ф(u_{кр} - \lambda) + Ф(u_{кр} + \lambda)]$$

где $u_{кр}$ находят из равенства Ф ($u_{кр}$) = (1 – α)/2, λ = (a_1 – – a_0)$\sqrt{n}/\sigma$.

При различных значениях $\boldsymbol{a_1}$ функция мощности двустороннего критерия:

$$\pi_1\ (a_1) = 1 - [Ф(u_{кр} - \lambda) + Ф(u_{кр} + \lambda)].$$

В формулах:

$$Ф(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$ – функция Лапласа.

б) Дисперсия генеральной совокупности неизвестна

Если дисперсия генеральной совокупности ***неизвестна*** (например, в случае малых выборок), то в качестве критерия проверки нулевой гипотезы принимают случайную величину:

$$T = (\overline{X} - a_0)\sqrt{n}/S,$$

где $S = \sqrt{\dfrac{\sum n_i x_i^2 - [\sum n_i x_i]^2/n}{n-1}}$ исправленное среднее квадратическое отклонение. Величина Т имеет распределение Стьюдента с $k = n - 1$ степенями свободы.

Правило 1. Для того чтобы при заданном уровне значимости α проверить нулевую гипотезу:

$$H_0: \boldsymbol{a = a_0}$$

о равенстве неизвестной генеральной средней *a* гипотетическому значению a_0 при конкурирующей гипотезе:

$$H_1: \boldsymbol{a \neq a_0,}$$

надо вычислить наблюдаемое значение критерия:

$$T_{набл} = (\bar{x} - a_0)\sqrt{n} / S$$

и по таблице критических точек распределения Стьюдента, по заданному уровню значимости α, помещенному в верхней строке таблицы, и числу степеней свободы $k = n - 1$ найти критическую точку $t_{двуст.кр}(\alpha, k)$.

Если $|T_{набл}| < t_{двуст.кр}(\alpha, k)$ – нет оснований отвергнуть нулевую гипотезу.

Если $|T_{набл}| > t_{двуст.кр}(\alpha, k)$ – нулевую гипотезу отвергают.

Правило 2. При конкурирующей гипотезе:

$$H_1: \boldsymbol{a > a_0}$$

по уровню значимости α, помещенному в нижней строке таблицы приложения, и числу степеней свободы $k = n - 1$ находят критическую точку $t_{правост.кр}(\alpha, k)$ правосторонней критической области.

Если $|T_{набл}| < t_{правост.кр}(\alpha, k)$ – нет основании отвергнуть нулевую гипотезу.

Если $|T_{набл}| > t_{правост.кр}(\alpha, k)$ – нулевую гипотезу отвергают.

Правило 3. При конкурирующей гипотезе:

$$H_1: \boldsymbol{a < a_0}$$

сначала находят вспомогательную критическую точку (по правилу 2) $t_{правост.кр}(\alpha\ , k)$ и полагают границу левосторонней критической области и полагают $t_{левостр.кр} = - t_{правостр.кр}$.

Если $|T_{набл}| > - t_{правост.кр}(\alpha, k)$ – нет оснований отвергнуть нулевую гипотезу.

Если $|T_{набл}| < - t_{правост.кр}(\alpha, k)$ – нулевую гипотезу отвергают.

Сравнение исправленной выборочной дисперсии с гипотетической генеральной дисперсией нормальной совокупности

Обозначим через *n* объем выборки, по которой найдена исправленная дисперсия *S2*.

Правило 1. Для того чтобы при заданном уровне значимости α проверить нулевую гипотезу:

$$H_0:\ \sigma^2 = \sigma_0^{\ 2}$$

о равенстве неизвестной генеральной дисперсии σ^2 гипотетическому (предполагаемому) значению $\sigma_0^{\ 2}$ при конкурирующей гипотезе:

$$H_1:\ \sigma^2 > \sigma_0^{\ 2}$$

надо вычислить наблюдаемое значение критерия:

$$\chi^2{}_{набл} = \frac{(n-1)S^2}{\sigma^2{}_0},$$

и по таблице критических точек распределения $\chi 2$, по заданному уровню значимости α и числу степеней свободы $k = n - 1$ найти критическую точку $\chi_{кр}^{\ 2}\ (\alpha\ ; k)$.

Если $\chi^2_{набл} < \chi_{кр}^{\ 2}$ – нет оснований отвергнуть нулевую гипотезу.

Если $\chi^2_{набл} > \chi_{кр}^{\ 2}$ – нулевую гипотезу отвергают.

Правило 2. При конкурирующей гипотезе:

$$H_1\text{:}\ \sigma^2 \neq \sigma_0^{\ 2}$$

находят левую $\chi^2_{лев.кр}$ $(1 - \alpha/2; k)$ и правую $\chi^2_{правкр}$ $(\alpha/2; k)$ критические точки.

Если $\chi^2_{лев.кр} < \chi^2_{набл} < \chi^2_{правкр}$ – нет оснований отвергнуть нулевую гипотезу.

Если $\chi^2_{набл} < \chi^2_{лев.кр}$ или $\chi^2_{набл} > \chi^2_{правкр}$ – нулевую гипотезу отвергают.

Правило 3. При конкурирующей гипотезе:

$$H_1\text{:}\ \sigma^2 < \sigma_0^{\ 2}$$

находят критическую точку $\chi_{кр}^{\ 2}$ $(1 - \alpha, k)$.

Если $\chi^2_{набл} > \chi_{кр}^{\ 2}$ $(1 - \alpha, k)$ – нет оснований отвергнуть нулевую гипотезу.

Если $\chi^2_{набл} < \chi_{кр}^{\ 2}$ $(1 - \alpha, k)$ – нулевую гипотезу отвергают.

Замечание. Если число степеней свободы $k > 30$, то критическую точку можно найти из равенства *Уилсона – Гильферти:*

$$\chi_{кр}^{\ 2}\ (\alpha, k) = k[\,1\text{-}2/9k + z_\alpha\sqrt{2/(9k)}\,]3,$$

где z_α находят, используя функцию Лапласа (приложение 3), из равенства: $\Phi(z_\alpha) = (1 - 2\alpha) / 2.$

Сравнение двух вероятностей биномиальных распределений

Пусть в двух генеральных совокупностях производятся независимые испытания: в результате каждого испытания событие А может появиться в первой совокупности с неизвестной вероятностью p_1, а во второй – с неизвестной вероятностью p_2;. По выборкам, извлеченным из первой и второй совокупностей, найдены соответственные частоты:

$$\omega_1(A) = m_1 / n_1 \text{ и } \omega_2(A) = m_2 / n_2,$$

где m_1, m_2 – числа появлений события А;

n_1, n_2 – количества испытаний.

В качестве оценок неизвестных вероятностей примем относительные частоты: $p_1 \approx \omega_1$ *и* $p_2 \approx \omega_2$. Требуется при заданном уровне значимости α проверить нулевую гипотезу, состоящую в том, что вероятности p_1 и p_2 равны между собой: H_0: $\boldsymbol{p_1 = p_2}$.

Другими словами, требуется установить, значимо или незначимо различаются относительные частоты ω_1 и ω_2.

Предполагается, что выборки имеют достаточно большой объем.

Правило 1. Для того чтобы при заданном уровне значимости α проверить нулевую гипотезу:

$$H_0: \boldsymbol{p_1 = p_2 = p}$$

о равенстве вероятностей появления события в двух генеральных совокупностях (имеющих биномиальные распределения) при конкурирующей гипотезе:

$$H_1: p_1 \neq p_2$$

надо вычислить наблюдаемое значение критерия:

$$U_{набл} = \frac{m_1/n_1 - m_2/n_2}{\sqrt{\frac{m_1+m_2}{n_1+n_2}\left(1-\frac{m_1+m_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}},$$

и по таблице функции Лапласа (приложение 3) найти критическую точку $u_{кр}$ по равенству $\Phi(u_{кр}) = (1 - \alpha)/2$.

Если $|U_{набл}| < u_{кр}$ – нет оснований отвергнуть нулевую гипотезу.

Если $|U_{набл}| > u_{кр}$ – нулевую гипотезу отвергают.

Правило 2. При конкурирующей гипотезе:

$$H_1: p_1 > p_2$$

находят критическую точку правосторонней критической области по равенству $\Phi(u_{кр}) = (1 - 2\alpha)/2$.

Если $U_{набл} < u_{кр}$ – нет оснований отвергнуть нулевую гипотезу.

Если $U_{набл} > u_{кр}$ – нулевую гипотезу отвергают.

Правило 3. При конкурирующей гипотезе:

$$H_1: p_1 < p_2$$

находят критическую точку $u_{кр}$ по правилу 2, а затем полагают границу левосторонней критической области

$$u'_{кр} = -u_{кр}.$$

Если $U_{набл} > -u_{кр}$ – нет оснований отвергнуть нулевую гипотезу.

Если $U_{набл} < -u_{кр}$ – нулевую гипотезу отвергают.

4.3. Решение типовых задач

Задача 1. Из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma = 5{,}2$ извлечена выборка объема $n = 100$ и по ней найдена выборочная средняя $\overline{x} = 27{,}56$. Требуется при уровне значимости 0,05 проверить нулевую гипотезу H_0: $a = a_0 = 26$ при конкурирующей гипотезе H_1: $a \neq 26$.

Решение: Исходя из условий задачи, сформулируем основную и альтернативную гипотезы:

$$H_0: \boldsymbol{a = a_0 = 26}$$

$$H_1: \boldsymbol{a \neq 26.}$$

Так как *генеральная дисперсия* σ^2 *известна,* то выбираем статистику:

$$U = \frac{(\overline{x} - a_0)\sqrt{n}}{\sigma}.$$

Вычисляем наблюдаемое значение критерия (статистики):

$$U_{набл} = \frac{(\overline{x} - a_0)\sqrt{n}}{\sigma} = \frac{27{,}56 - 26}{5{,}2} \cdot \sqrt{100} = 3.$$

По таблице функции Лапласа (приложение 3) находим критическую точку $u_{кр} = 1{,}96$ двусторонней критической области из равенства:

$$\Phi(u_{кр}) = (1 - \alpha)/2 = (1 - 0{,}05) / 2 = 0{,}475.$$

Так как $|U_{набл}| = 3 > u_{кр} = 1{,}96$, то нулевую гипотезу отвергаем. Другими словами, выборочная и гипотетическая генеральная средние различаются значимо.

Задача 2. На основании сделанного прогноза средняя дебиторская задолженность однотипных предприятий региона должна составить $a_0 = 120$ ден. ед. Выборочная проверка десяти предприятий дала среднюю задолженность $\bar{x} = 135$ ден. ед., а среднее квадратическое отклонение задолженности $s = 20$ ден. ед. На уровне значимости 0,05 выяснить, можно ли принять данный прогноз.

Решение: Исходя из условий задачи, сформулируем основную и альтернативную гипотезы:

$$H_0: \boldsymbol{a = (a_0 = 120)}$$

$$H_1: \boldsymbol{a > 120}$$

Так как *генеральная дисперсия* σ^2 *неизвестна,* то выбираем статистику:

$$T = (\bar{x} - a_0)\sqrt{n} / S .$$

Величина T имеет распределение Стьюдента с $k = n - 1$ степенями свободы.

Вычисляем наблюдаемое значение критерия (статистики):

$$T_{набл} = (\bar{x} - a_0)\sqrt{n} / S = (135 - 120)\sqrt{10} / 20 = 2{,}25.$$

По уровню значимости $\alpha = 0{,}05$, помещенному в нижней строке таблицы приложения 1, и числу степеней свободы $k = n - 1 = 10 - 1 = 9$ находят критическую точку

$t_{правост.кр}(\alpha = 0{,}05, k = 9) = 1{,}83$ правосторонней критической области.

Так как $|T_{набл}| = 2{,}25 > t_{0,05;9} = 1{,}83$, то нулевая гипотеза отвергается, т.е. на 5%-м уровне значимости сделанный прогноз должен быть отвергнут.

Задача 3. По паспортным данным автомобильного двигателя расход топлива на 100 км составляет 10 л. В результате изменения конструкции двигателя ожидается, что расход топлива уменьшится. Для проверки проведены испытания 25 случайно отобранных автомобилей с новым двигателем. По результатам испытаний получено, что на 100 км в среднем затрачивается 9,3 л.

Считая, что расход топлива – X, на 100 км имеет нормальное распределение со средним квадратическим отклонением $\sigma = 2$ л., проверить гипотезу, что изменение конструкции не повлияло на расход топлива (на 5%-м уровне значимости).

Решение: Исходя из условий задачи, сформулируем основную и альтернативную гипотезы:

$$H_0: \boldsymbol{a = (a_0 = 10)}$$

$$H_1: \boldsymbol{a < 10}$$

Так как *генеральная дисперсия* σ^2 *известна,* то выбираем статистику:

$$U = \frac{(\bar{x} - a_0)\sqrt{n}}{\sigma}.$$

Вычисляем наблюдаемое значение критерия (статистики):

$$U_{набл} = \frac{(\bar{x}-a_0)\sqrt{n}}{\sigma} = \frac{9{,}3-10}{2}\cdot\sqrt{25} = -1{,}75\,,$$

и по таблице функции Лапласа (приложение 3), находим критическую точку $u_{кр}$ правосторонней критической области из равенства:

$$\Phi(u_{кр}) = (1 - 2\alpha)/2 = (1 - 2\cdot 0{,}05) / 2 = 0{,}45,$$

а затем полагают границу левосторонней критической области: $$u'_{кр} = - u_{кр.} = - 1{,}65.$$

Так как $U_{набл} = - 1{,}75 < - u_{кр} = - 1{,}65$, то на 5%-м уровне значимости нулевую гипотезу отвергают. Таким образом, можно утверждать (с риском ошибиться в 5% случаев), что автомобиль с новым двигателем более экономичен.

Задача 4. Из нормальной генеральной совокупности извлечена выборка объема n = 21, и по ней найдена исправленная выборочная дисперсия s^2 = 16,2. Требуется при уровне значимости 0,01 проверить нулевую гипотезу: H_0: $\sigma^2 = \sigma_0^2 = 15$, приняв в качестве конкурирующей гипотезы H_1: $\sigma_0^{2.} > 15$.

Решение: Исходя из условий задачи, сформулируем основную и альтернативную гипотезы:

$$\mathbf{H_0}\text{:}\ \sigma^2 = \sigma_0^2 = 15,$$

$$\mathbf{H_1}\text{:}\ \sigma_0^{2.} > 15.$$

Выбираем статистику:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2_0}.$$

Вычисляем наблюдаемое значение критерия:

$$\chi^2_{набл} = \frac{(n-1)S^2}{\sigma^2_0} = \frac{(21-1)\cdot 16{,}2}{15} = 21{,}6.$$

По условию, конкурирующая гипотеза имеет вид $\sigma_0^{2.} > 15.$, поэтому критическая область – правосторонняя. По таблице критических точек распределения $\chi 2$ (приложение 4), по заданному уровню значимости α = 0,01 и числу степеней свободы $k = n - 1$ = 21 – 1 = 20 находим критическую точку $\chi_{кр}{}^2$ (0,01;20) = 37,6.

Так как $\chi^2{}_{набл}$ = 21,6 < $\chi_{кр}{}^2$ = 37,6, то нет оснований отвергнуть нулевую гипотезу о равенстве генеральной дисперсии гипотетическому значению $\sigma_0^2 = 15$. Другими словами, различие между исправленной дисперсией (s^2 = 16,2) и гипотетической генеральной дисперсией ($\sigma_0^2 = 15$) незначимо.

Задача 5. За смену отказали 15 элементов устройства 1, состоящего из 800 элементов и 25 элементов устройства 2, состоящего из 1000 элементов. При уровне значимости 0,05 проверить нулевую гипотезу H_0: $p_1 = p_2 = p_o$ равенстве вероятностей отказа элементов обоих устройств при конкурирующей гипотезе H_1: $p_1 \neq p_2$.

Решение: Исходя из условий задачи, сформулируем основную и альтернативную гипотезы:

$$H_0: \boldsymbol{p_1 = p_2 = p_o}$$

$$H_1: \boldsymbol{p_1 \neq p_2}.$$

Выбираем статистику:

$$U = \frac{m_1/n_1 - m_2/n_2}{\sqrt{\frac{m_1+m_2}{n_1+n_2}\left(1-\frac{m_1+m_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}.$$

Вычисляем наблюдаемое значение критерия:

$$U_{набл} = \frac{m_1/n_1 - m_2/n_2}{\sqrt{\frac{m_1+m_2}{n_1+n_2}\left(1-\frac{m_1+m_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} =$$

$$= \frac{15/800 - 25/1000}{\sqrt{\frac{15+25}{800+1000}\cdot(1-\frac{15+25}{800+1000})}\cdot\sqrt{\frac{1}{800}+\frac{1}{1000}}}.$$

$$U_{набл} = -0{,}89.$$

По таблице функции Лапласа (приложение 3) находим критическую точку $u_{кр} = 1{,}96$ по равенству

$$\Phi(u_{кр}) = (1 - \alpha)/2 = (1 - 0{,}05) / 2 = 0{,}475.$$

Так как $|U_{набл}| = 0{,}89 < u_{кр} = 1{,}96$, то нет оснований отвергнуть нулевую гипотезу. Другими словами, вероятности отказа элемента обоих устройств различаются незначимо.

4.4. Задания к лабораторно-практической работе

Вариант 1

Из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma = 40$ извлечена выборка объема $n = 64$ и по ней найдена выборочная средняя $\bar{x} = 136{,}5$. Требуется при уровне значимости 0,01 проверить нулевую гипотезу H_0: $a = a_0 = 130$ при конкурирующей гипотезе H_1: $a \neq 130$.

Вариант 2

Из нормальной генеральной совокупности с известным средним квадратическим отклонением $\sigma = 40$ извлечена выборка объема $n = 64$ и по ней найдена выборочная средняя $\bar{x} = 136{,}5$. Требуется при уровне значимости 0,01 проверить нулевую гипотезу H_0: $a = a_0 = 130$ при конкурирующей гипотезе H_1: $a > 130$.

Вариант 3

Установлено, что средний вес таблетки лекарства сильного действия должен быть равен $a_0 = 0{,}50$ мг. Выборочная проверка 121 таблетки полученной партии лекарства показала, что средний вес таблетки этой партии $\bar{x} = 0{,}53$ мг. Требуется при уровне значимости 0,01 проверить нулевую гипотезу H_0: $a = a_0 = 0{,}50$ при конкурирующей гипотезе H_1: $a > 0{,}50$. Многократными предварительными опытами по взвешиванию таблеток, поставляемых фармацевтическим заводом, было установлено, что вес таблеток распределен нормально со средним квадратическим отклонением $\sigma = 0{,}11$ мг.

Вариант 4

Из нормальной генеральной совокупности сельскохозяйственных предприятий, рассматриваемых по показателю урожайности пшеницы, с известным средним квадратическим отклонением $\sigma = 9{,}4$ и генеральной средней $a_0 = 38{,}1$, извлечена выборка объема $n = 50$. По ней найдена выборочная средняя $\bar{x}$ = 42. Требуется при 5%-м уровне значимости проверить нулевую гипотезу: H_0: $a = a_0 = 38{,}1$ при конкурирующей гипотезе H_1: $a < 38{,}1$.

Вариант 5

Составлена случайная выборка из 64 покупателей, которые интересовались товаром А. Из них товар А купили 16 человек. Поставщик утверждает, что данный товар должен привлечь в среднем $\bar{x}$ = 21 человек, а среднее квадратическое отклонение σ_x равно одному человеку. Проверить нулевую гипотезу H_0: $a = a_0 = 21$ при конкурирующей гипотезе H_1: $a > 21$, при 5%-м уровне значимости.

Вариант 6

Составлена случайная выборка из 64 покупателей, которые интересовались товаром А. Из них товар А купили 16 человек. Поставщик утверждает, что данный товар должен привлечь в среднем $\bar{x}$ = 21 человек, а среднее квадратическое отклонение σ_x равно одному человеку. Проверить нулевую гипотезу H_0: $a = a_0 = 21$ при конкурирующей гипотезе H_1: $a < 21$, при 5%-м уровне значимости.

Вариант 7

Средний размер подшипников должен составлять 35 мм. Однако для выборки из 82 подшипников он составил 35,3 мм при выбранном среднем квадратическом отклонении 0,1 мм. При 5%-м уровне значимости проверить гипотезу о том, что станок, на котором изготавливают подшипники, не требует подналадки.

Вариант 8

Поставщик удобрений утверждает, что применение новой партии удобрений обеспечивает урожайность пшеницы в 60 ц/га. Удобрения распространили на площади в 37 га и получили урожай 55 ц/га при выборочном среднем квадратичном отклонении 3 ц/га. При 5%-м уровне значимости оценить справедливость утверждения поставщика.

Вариант 9

Среднесуточная продажа хлеба в течение многих лет для данного магазина составила 6 т при среднем квадратичном отклонении 0,05 т. Сегодня магазином было продано 7 т хлеба. Можно ли при 5%-м уровне значимости предполагать, что и завтра будет продано 7 т хлеба?

Вариант 10

Из нормальной генеральной совокупности сельскохозяйственных предприятий, рассматриваемых по показателю урожайности пшеницы, с известным средним квадратическим отклонением $\sigma = 9{,}4$ и генеральной средней $a_0 = 38{,}1$, извлечена выборка объема $n = 50$. По ней найдена выборочная средняя $\bar{x} = 42$. Требуется при 5%-м уровне значимости проверить нулевую гипотезу: H_0: $a = a_0 = 38{,}1$ при конкурирующей гипотезе H_1: $a > 38{,}1$.

Вариант 11

Фирма – изготовитель женских украшений, выпустив новый товар, утверждает, что 40% покупателей купят эти украшения. В ходе 10-дневной рекламной распродажи в среднем приобрели украшения 29,5% покупателей, выборочное (исправленное) среднее квадратичное отклонение составило 16,5%. При 5%-м уровне значимости оценить утверждение изготовителя товара.

Вариант 12

Поставщик двигателей утверждает, что срок их службы равен 800 ч. Для выборки из 17 двигателей средний срок службы оказался равным 865 ч при выборочном (исправленном) среднем квадратичном отклонении 120 ч. Проверить нулевую гипотезу при уровне значимости 1%.

Вариант 13

Из большой партии ананасов одного размера случайным образом отобрано 36 штук. Выборочная средняя масса одной штуки оказалась равной 930 г. Используя двусторонний критерий при α = 0,05, проверить гипотезу, что средняя масса одного ананаса составляет 1 кг, если среднее квадратичное отклонение неизвестно, а выборочное (исправленное) составило 250 г.

Вариант 14

По выборке объема n = 16, извлеченной из нормальной генеральной совокупности, найдены выборочная средняя $\bar{x}$ = 118,2 и «исправленное» среднее квадратическое отклонение S = 3,6. Требуется при уровне значимости 0,05 проверить нулевую гипотезу H_0: $a = a_0 = 120$ при конкурирующей гипотезе H_1: $a \neq 120$.

Вариант 15

По выборке объема n = 16, извлеченной из нормальной генеральной совокупности, найдены выборочная средняя $\bar{x}$ = 118,2 и «исправленное» среднее квадратическое отклонение S = 3,6. Требуется при уровне значимости 0,05 проверить нулевую гипотезу H_0: $a = a_0 = 120$ при конкурирующей гипотезе H_1: $a < 120$.

Вариант 16

Проектный контролируемый размер изделий, изготовляемых станком-автоматом, $a = a_0 = 35$ мм. Измерения 20 случайно отобранных изделий дали следующие результаты:

Контролируемый размер x_i	34,8	34,9	35,0	35,1	35,3
Частота (число изделий m_i)	2	3	4	6	5

Требуется при уровне значимости 0,05 проверить нулевую гипотезу $a = a_0 = 35$ при конкурирующей гипотезе H_1: $a \neq 35$.

Вариант 17

Проверить нулевую гипотезу о том, что заданное значение $a_0 = 70$ является математическим ожиданием нормально распределенной случайной величины при 5%-м уровне значимости для двусторонней критической области, если в результате обработки выборки объема n = 10 получено выборочное среднее $\bar{x}$ = 72, а выборочное среднее квадратическое отклонение s = 5.

Вариант 18

Проверить нулевую гипотезу о том, что заданное значение $a_0 = 90$ является математическим ожиданием нормально распределенной случайной величины при 5%-м уровне значимости для двусторонней критической области, если в результате обработки выборки объема n = 10 получено выборочное среднее $\bar{x} = 88$, а выборочное среднее квадратическое отклонение s = 6.

Вариант 19

Проверить нулевую гипотезу о том, что заданное значение $a_0 = 60$ является математическим ожиданием нормально распределенной случайной величины при 5%-м уровне значимости для двусторонней критической области, если в результате обработки выборки объема $n = 10$ получено выборочное среднее $\bar{x} = 54$, а выборочное среднее квадратическое отклонение s = 2.

Вариант 20

Проверить нулевую гипотезу о том, что заданное значение $a_0 = 84$ является математическим ожиданием нормально распределенной случайной величины при 5%-м уровне значимости для двусторонней критической области, если в результате обработки выборки объема $n = 10$ получено выборочное среднее $\bar{x} = 80$, а выборочное среднее квадратическое отклонение s = 6.

Вариант 21

Из нормальной генеральной совокупности извлечена выборка объема $n = 17$ и по ней найдена исправленная выборочная дисперсия $S^2 = 0{,}24$. Требуется при уровне значимости

0,05 проверить нулевую гипотезу H_0: $\sigma^2 = \sigma_0^2 = 0,18$ при конкурирующей гипотезе H_1: $\sigma^2 > 0,18$.

Вариант 22

Из нормальной генеральной совокупности извлечена выборка объема $n = 31$:

Варианты x_i	10,1	10,3	10,6	11,2	11,5	11,8	12,0
Частоты m_i	1	3	7	10	6	3	1

Требуется при уровне значимости 0,05 проверить нулевую гипотезу H_0: $\sigma^2 = \sigma_0^2 = 0,18$ при конкурирующей гипотезе H_1: $\sigma^2 > 0,18$.

Вариант 23

Точность работы станка-автомата проверяется по дисперсии контролируемого размера изделий, которая должна превышать $\sigma_0^2 = 0,1$. Взята проба из 25 случайных отобранных изделий, причем получены следующие результаты измерений:

Контролируемый размер изделий пробы x_i	3,0	3,5	3,8	4,4	4,5
Частота m_i	2	6	9	7	1

Требуется при уровне значимости 0,05 проверить, обеспечивает ли станок требуемую точность.

Вариант 24

В результате длительного хронометража времени сборки узла различными сборщиками установлено, что дисперсия этого времени $\sigma_0^2 = 2$ мин². Результаты 20 наблюдений за работой новичка таковы:

Время сборки одного узла в минутах x_i	56	58	60	62	64
Частота m_i	1	4	10	3	2

Можно ли при уровне значимости 0,05 считать, что новичок работает ритмично (в том смысле, что дисперсия затрачиваемого им времени существенно не отличается от дисперсии времени остальных сборщиков)?

Вариант 25

Партия изделий принимается, если дисперсия контролируемого размера значимо не превышает 0,2. Исправленная выборочная дисперсия, найденная по выборке объема n = 121, оказалась равной $S_X^2 = 0,3$.

Требуется при уровне значимости 0,01 проверить нулевую гипотезу H_0: $\sigma^2 = \sigma_0^2 = 0,2$ при конкурирующей гипотезе H_1: $\sigma^2 > 0,2$. Можно ли принять партию при уровне значимости 0,01?

Вариант 26

В партии из 500 деталей, изготовленных первым станком-автоматом, оказалось 60 нестандартных; из 600 деталей второго станка 42 нестандартных. При уровне значимости 0,01 проверить нулевую гипотезу H_0: $p_1 = p_2 = p$ о равенстве вероятностей изготовления нестандартной детали обоими станками при конкурирующей гипотезе H_1: $p_1 \neq p_2$.

Вариант 27

Для оценки качества изделий, изготовленных двумя заводами, взяты выборки n_1 = 200 и n_2 = 300 изделий. В этих выборках оказалось соответственно m_1 =20 и m_2 = 15 бракованных изделий. При уровне значимости 0,05 проверить нулевую гипотезу H_0: $p_1 = p_2 = p$ о равенстве вероятностей изготовления бракованных изделий обоими заводами при конкурирующей гипотезе H_1: $p_1 > p_2$.

Вариант 28

Из 100 выстрелов по цели каждым из двух орудий зарегистрировано соответственно $m_1 = 12$ и $m_2 = 8$ промахов. При уровне значимости 0,05 проверить нулевую гипотезу H_0: $p_1 = p_2 = p$ о равенстве вероятностей промаха обоих орудий при конкурирующей гипотезе H_1: $p_1 \neq p_2$.

Вариант 29

Ниже приведены результаты выборочного обследования двух партий изделий:

№ партии	Число изделий		Сумма
	бракованных	небракованных	
1	8	92	100
2	13	287	300
Сумма	21	379	

При уровне значимости 0,05 проверить нулевую гипотезу H_0: $p_1 = p_2 = p$ о равенстве вероятностей при конкурирующей гипотезе H_1: $p_1 \neq p_2$. Можно ли считать, что доля брака в обеих партиях одна и та же?

Вариант 30

На основе выборки в 100 семян, взятой случайным образом из партии с однородной смесью семян пшеницы, при помощи стандартной методики установлена всхожесть 92%. При уровне значимости 0,01 проверить нулевую гипотезу, что всхожесть всей партии равна 90%, т.е. H_0: $p = p_0 = 0,9$ при конкурирующей гипотезе H_1: $p < p_0 = 0,9$.

ЛАБОРАТОРНО-ПРАКТИЧЕСКАЯ РАБОТА 5
ПРОВЕРКА НЕПАРАМЕТРИЧЕСКИХ ГИПОТЕЗ. КРИТЕРИЙ ПИРСОНА

Цель работы: освоение алгоритма применения критерия Пирсона для проверки статистических гипотез о виде закона распределения.

5.1. Теоретические вопросы

1. Алгоритм проверки гипотезы о виде закона распределения. Критерий согласия Пирсона.

2. Проверка гипотезы о нормальном законе распределения генеральной совокупности по критерию Пирсона.

5.2. Методические рекомендации

Критерии, с помощью которых проверяется нулевая гипотеза о неизвестном законе распределения, называются ***критериями согласия.***

Критерий согласия χ^2 «хи-квадрат» Пирсона – наиболее часто употребляемый критерий для проверки гипотезы о законе распределения.

Пусть необходимо *проверить нулевую гипотезу H_0 о том, что исследуемая случайная величина X подчиняется определенному закону распределения*.

Схема применения критерия χ^2 для проверки гипотезы H_0 сводится к следующему:

1. Определяется мера расхождения эмпирических и теоретических частот χ^2:

$$\chi^2 = \sum \frac{(m_i - np_i)^2}{np_i},$$

где m_i – *эмпирические* (опытные) *частоты* случайной величины X;

$np_i = m_i^|$ – *теоретические частоты,* представляющие произведение числа наблюдений *n* на вероятности p_i, рассчитанные по предполагаемому теоретическому распределению.

2. Выборочная статистика χ^2 при $n \to \infty$ имеет χ^2 – распределение с $k = m - s - 1$ степенями свободы, где *m* – число интервалов эмпирического распределения, *s* – число параметров теоретического распределения, определяемых по опытным данным.

Для выбранного уровня значимости α по таблице χ^2 – распределения находят критическое значение $\chi^2_{кр.}$ при числе степеней свободы *k.*

3. Если фактически наблюдаемое значение χ^2 больше критического, т.е. если $\chi^2_{набл.} > \chi^2_{кр.}$, то гипотеза H_0 отвергается; если $\chi^2_{набл.} \leq \chi^2_{кр.}$, то гипотеза H_0 не противоречит опытным данным.

Рассмотрим ***проверку гипотезы о нормальном законе распределения*** генеральной совокупности **по критерию Пирсона.**

1. **Пусть** эмпирическое распределение задано в виде последовательности ***равноотстоящих вариант*** и соответствующих им частот:

x_i	x_1	x_2	…	x_n
m_i	m_1	m_2	…	m_n

Требуется, ***используя критерий Пирсона, проверить гипотезу о том, что генеральная совокупность X распределена нормально.***

Правило 1. Для того чтобы при заданном уровне значимости α проверить гипотезу о нормальном распределении генеральной совокупности, надо:

1. Вычислить непосредственно (при малом числе наблюдений) или упрощенным методом (при большом числе наблюдений), например, методом произведений, выборочную среднюю $\bar{x}_в$ и выборочное среднее квадратическое отклонение $\sigma_в$.

2. Вычислить теоретические частоты:

$$m_i = \frac{nh}{\sigma_в}\varphi(u_i),$$

где *n* – объем выборки,
h – шаг (разность между двумя соседними вариантами),

$$u_i = \frac{x_i - \bar{x}_в}{\sigma_в}, \quad \varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

3. Сравнить эмпирические и теоретические частоты с помощью критерия Пирсона. Для этого:

а) составить расчетную таблицу, по которой находят наблюдаемое значение критерия:

$$\chi^2_{набл} = \sum \frac{(m_i - m_i^|)^2}{m_i^|};$$

б) по таблице критических точек распределения χ^2 (приложение 4), по заданному уровню значимости α и числу степеней свободы $k = s - 3$ (*s* –число групп выборки) находят критическую точку, $\chi^2_{кр}(\alpha; k)$ правосторонней критической области.

Если $\chi^2_{набл} < \chi_{кр}^2$ – нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности. Другими словами, эмпирические и теоретические частоты различаются незначимо (случайно).

Если $\chi^2_{набл} > \chi_{кр}^2$ – гипотезу отвергают. Другими словами, эмпирические и теоретические частоты различаются значимо.

Замечание 1. Малочисленные частоты (m_i < 30) следует объединить; в этом случае и соответствующие им теоретические частоты также надо сложить. Если производилось объединение частот, то при определении числа степеней свободы по формуле $k = s - 3$ следует в качестве s принять число групп выборки, оставшихся после объединения частот.

2. **Пусть** эмпирическое распределение задано в виде ***последовательности интервалов*** (x_i, x_{i+1}) и соответствующих им частот m_i, (m_i – сумма частот, которые попали в i-й интервал).

Требуется, ***используя критерии Пирсона, проверить гипотезу о том, что генеральная совокупность X распределена нормально.***

Правило 2. Для того чтобы при уровне значимости α проверить гипотезу о нормальном распределении генеральной совокупности, надо:

1. Вычислить, например, методом произведений, выборочную среднюю $\bar{x}$ и выборочное среднее квадратическое

отклонение $\sigma_в$, причем в качестве вариант x_i^* принимают среднее арифметическое концов интервала:

$$x_i = (x_i + x_{i+1})/2.$$

2. Пронормировать X, т.е. перейти к случайной величине Z = (X – $\bar{x}^*$)/ σ^*, и вычислить концы интервалов: $z_i = (x_i - \bar{x}^*)/\sigma^*$, $z_{i+1} = (x_{i+1} - \bar{x}^*)/\sigma^*$, причем наименьшее значение Z, т.е. $z_1 = -\infty$, а наибольшее, т.е. $z_{s+1} = \infty$.

3. Вычислить теоретические частоты: $m_i = nP_i$,

где $P_i = \Phi(z_{i+1}) - \Phi(z_i)$ – вероятности попадания X в интервалы $(x_i + x_{i+1})$;

Ф (Z) – функция Лапласа.

4. Сравнить эмпирические и теоретические частоты с помощью критерия Пирсона. Для этого:

а) составляют расчетную таблицу, по которой находят наблюдаемое значение критерия Пирсона:

$$\chi^2_{набл} = \sum \frac{(m_i - m_i^|)^2}{m_i^|};$$

б) по таблице критических точек распределения χ^2, по заданному уровню значимости α и числу степеней свободы $k = s - 3$ (s – число интервалов выборки) находят критическую точку правосторонней критической области $\chi_{кр}^2\ (\alpha; k)$.

Если $\chi^2_{набл} < \chi_{кр}^2$ – нет оснований отвергнуть гипотезу о нормальном распределении генеральной совокупности.

Если $\chi^2_{набл} > \chi_{кр}^2$ – гипотезу отвергают.

Замечание 2. Интервалы, содержащие малочисленные эмпирические частоты (m_i < 5), следует объединить, а частоты этих интервалов сложить. Если производилось объединение интервалов, то при определении числа степеней свободы по формуле $k = s - 3$ следует в качестве s принять число интервалов, оставшихся после объединения интервалов.

5.3. Решение типовых задач

Задача 1. Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с эмпирическим распределением выборки объема n = 200:

x_i	5	7	9	11	13	15	17	19	21
m_i	15	26	25	30	26	21	24	20	13

Решение:

1. Используя метод произведений, найдем выборочную среднюю $\bar{x}_в$ = 12,63 и выборочное среднее квадратическое отклонение $\sigma_в$ = 4,695.

2. Вычислим теоретические частоты, учитывая, что n = 200, h = 2, $\sigma_в$ = 4,695, по формуле:

$$m_i^| = \frac{nh}{\sigma_в}\varphi(u_i) = \frac{200*2}{4,695}\varphi(u_i) = 85,2\varphi(u_i).$$

Составим расчетную табл. 1:

Таблица 1

i	x_i	$u_i = \dfrac{x_i - \overline{x}_в}{\sigma_в}$	$\varphi(u_i)$	$m_i^{\vert} = 85{,}2\varphi(u_i)$
1	5	– 1,62	0,1074	9,1
2	7	– 1,20	0,1942	16,5
3	9	– 0,77	0,2966	25,3
4	11	– 0,35	0,3752	32,0
5	13	0,08	0,3977	33,9
6	15	0,51	0,3503	29,8
7	17	0,93	0.2589	22,0
8	19	1,36	0,1582	13,5
9	21	1,78	0,0818	7,0

3. Сравним эмпирические и теоретические частоты.

а) Составим расчетную табл. 2, из которой найдем наблюдаемое значение критерия: $\chi^2_{набл} = \sum \dfrac{(m_i - m_i^{\vert})^2}{m_i^{\vert}}$.

Таблица 2

i	m_i	$m_i^{\vert}$	$m_i - m_i^{\vert}$	$(m_i - m_i^{\vert})^2$	$(m_i - m_i^{\vert})^2/m_i^{\vert}$
1	15	9,1	5,9	34,81	3,8
2	26	16,5	9,5	90,25	5,5
3	25	25,3	– 0,3	0,09	0,0
4	30	32,0	– 2,0	4,00	0,1
5	26	33,9	– 7,9	62,41	1,8
6	21	29,8	– 8,8	77,44	2,6
7	24	22,0	2,0	4,00	0,2
8	20	13,5	6,5	42,25	3,1
9	13	7,0	6,0	36,00	5,1
$\sum$	200				$\chi^2_{набл} = 22{,}2$

Из табл. 2 находим $\chi^2_{набл} = 22{,}2$.

б) По таблице критических точек распределения χ^2 (приложение 4), по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = s - 3 = 9 - 3 = 6$ находим критическую точку правосторонней критической области:

$$\chi_{кр}^2\ (0.05;\ 6) = 12{,}6.$$

Так как $\chi^2_{набл} > \chi_{кр}^2$ – гипотезу о нормальном распределении генеральной совокупности отвергаем. Другими словами, эмпирические и теоретические частоты различаются значимо.

Задача 2. Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с эмпирическим распределением выборки объема $n = 100$, приведенным в табл. 1.

Таблица 1

Номер интервала i	Граница интервала		Частота m_i
	x_i	x_{i+1}	
1	3	8	6
2	8	13	8
3	13	18	15
4	18	23	40
5	23	28	16
6	28	33	8
7	33	38	7
			$n = 100$

Решение:

1. Вычислим выборочную среднюю и выборочное среднее квадратическое отклонение методом произведений. Для этого перейдем от заданного интервального распределения к распределению равноотстоящих вариант, приняв в качестве варианты x_i^* среднее арифметическое концов интервала: $x^*_1 = (x_i + x_{i+1})/2$. В итоге получим распределение:

x_i^*	5,5	10,5	15,5	20,5	25,5	30,5	35,5
m_i	6	8	15	40	16	8	7

Выполнив выкладки по методу произведений, найдем выборочную среднюю и выборочное среднее квадратическое отклонение: $\bar{x}^* = 20{,}7$, $\sigma^* = 7{,}28$.

3. Найдем интервалы (z_i, z_{i+1}), учитывая, что $\bar{x}^* = 20{,}7$, $\sigma^* = 7{,}28$, $1/\sigma^* = 0{,}137$. Для этого составим расчетную табл. 2 (левый конец первого интервала примем равным $-\infty$, а правый конец последнего интервала равен ∞).

Таблица 2

i	Границы интервала		$x_i - \bar{x}^\bullet$	$x_{i+1} - \bar{x}^\bullet$	Границы интервала	
	x_i	x_{i+1}			$z_i = (x_i - \bar{x}^\bullet)/\sigma^\bullet$	$z_{i+1} = (x_{i+1} - \bar{x}^\bullet)/\sigma^\bullet$
1	3	8	–	– 12,7	$-\infty$	– 1,74
2	8	13	– 12,7	– 7,7	– 1,74	– 1,06
3	13	18	– 7,7	– 2,7	– 1,06	– 0,37
4	18	23	– 2,7	2,3	– 0,37	0,32
5	23	28	2,3	7,3	0,32	1,00
6	28	33	7,3	12,3	1,00	1,69
7	33	38	12,3	–	1,69	∞

3. Найдем теоретические вероятности P_i и теоретические частоты $m_i^{|} = nP_i = 100 \cdot P_i$. Для этого составим расчетную табл. 3.

Таблица 3

i	Границы интервала		$\Phi(z_i)$	$\Phi(z_{i+1})$	Границы интервала	
	z_i	z_{i+1}			$P_i = \Phi(z_{i+1}) - \Phi(z_i)$	$n_i' = 100P_i$
1	–	– 1,74	– 0,5000	– 0,4591	0,0409	4,09
2	– 1,74	– 1,06	– 0,4591	– 0,3554	0,1037	10,37
3	– 1,06	– 0,37	– 0,3554	– 0,1443	0,2111	21,11
4	– 0,37	0,32	– 0,1443	0,1255	0,2698	26,98
5	0,32	1,00	0,1255	0,3413	0,2158	21,58
6	1,00	1,69	0,3413	0,4545	0,1132	11,32
7	1,69		0,4545	0,5000	0,0455	4,55
$\sum$					1	100

4. Сравним эмпирические и теоретические частоты, используя критерий Пирсона:

а) вычислим наблюдаемое значение критерия Пирсона. Для этого составим расчетную табл. 4. Столбцы 7 и 8 служат для контроля вычислений по формуле:

$$\chi^2_{набл} = \sum (m_i^2 / m_i^{|}) - n$$

Контроль: $\chi^2_{набл} = \sum (m_i^2 / m_i^{|}) - n = 113{,}22 - 100 = 13{,}22 =$

$= \chi^2_{набл}$. Вычисления произведены правильно.

Таблица 4

i	m_i	$m_i^{\vert}$	m_i - $m_i^{\vert}$	$(m_i - m_i^{\vert})^2$	$(m_i - m_i^{\vert})^2/m_i^{\vert}$	m_i^2	$m_i^2 / m_i^{\vert}$
1	6	4,09	1,91	3,6481	0,8920	36	8,8019
2	8	10,37	– 2,37	5,6169	0,5416	64	6,1716
3	15	21,11	– 6,11	37,3321	1,7684	225	10,6584
4	40	26,98	13,02	169,5204	6,2833	1600	59,3052
5	16	21,58	– 5,58	31,1364	1,4428	256	11,8628
6	8	11,32	– 3,32	11,0224	0,9737	64	5,6537
7	7	4,55	2,45	6,0025	1,3192	49	10,7692
$\sum$	100	100			$\chi^2_{набл}$ = 13.22		113,22

б) по таблице критических точек распределения χ^2 (приложение 4), по уровню значимости α = 0,05 и числу степеней свободы $k = s - 3$ = 7 – 3 = 4 (s – число интервалов) находим критическую точку правосторонней критической области $\chi_{кр}^2$ (0,05; 4) = 9,5.

Так как $\chi^2_{набл}$ = 113,22 > $\chi_{кр}^2$ = 9,5 – отвергаем гипотезу о нормальном распределении генеральной совокупности X; другими словами, эмпирические и теоретические частоты различаются значимо. Это означает, что данные наблюдений не согласуются с гипотезой о нормальном распределении генеральной совокупности.

5.4. Задания к лабораторно-практической работе

Вариант 1

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с эмпирическим распределением выборки объема n = 200:

x_i	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,1	2,3
m_i	6	9	26	25	30	26	21	24	20	8	5

Вариант 2

Дано распределение отклонений диаметров цапф передней оси от номинального размера:

x_i	26	29	32	35	38	41	44	47	50	53
m_i	1	4	13	23	22	29	29	16	11	2

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с эмпирическим распределением выборки объема n = 150.

Вариант 3

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X – число сделок на фондовой бирже за квартал с эмпирическим распределением выборки объема n = 400 (инвесторов):

x_i	0	1	2	3	4	5	6	7	8	9	10
m_i	146	97	73	34	23	10	6	3	4	2	2

Вариант 4

Экзаменационный билет по математике содержит 10 заданий. Пусть X – случайная величина числа задач, решенных абитуриентами на вступительном экзамене. Результаты сдачи экзамена по математике для 300 абитуриентов таковы:

x_i	0	1	2	3	4	5	6	7	8	9	10
m_i	13	17	15	35	10	9	40	51	45	33	32

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с эмпирическим распределением выборки.

Вариант 5

Результаты взвешивания 50 случайным образом отобранных пачек чая приведены ниже (в граммах):
150, 147, 152, 148, 149, 153, 151, 150, 149, 147, 153, 151, 152, 151, 149, 152, 150, 148, 152, 150, 152, 151, 148, 151, 152, 150, 151, 149, 148, 149, 150, 150, 151, 149, 151, 150, 151, 150, 149, 148, 147, 153, 147, 152, 150, 151, 149, 150, 151, 153.

Оценить закон распределения случайной величины X – массы пачки чая – для уровня значимости 5%.

Вариант 6

Дано следующее распределение успеваемости 125 студентов, сдавших три экзамена:

Число сданных экзаменов x_i	0	1	2	3
Число студентов m_i	3	5	47	70

Проверить гипотезу о биномиальном распределении числа сданных экзаменов при $\alpha = 0{,}05$.

Вариант 7

При принятии на работу фирма предлагает 4 теста. Результаты решения этих тестов десятью претендентами приведены ниже:

Число верно решенных тестов x_i	0	1	2	3	4
Число участников m_i	1	2	2	3	2

Проверить гипотезу о биномиальном распределении случайной величины X – числа успешно решенных тестов – при α = 0,05.

Вариант 8

Опыт, состоящий в одновременном подбрасывании четырех монет, повторили 100 раз. Эмпирическое распределение дискретной случайной величины X – числа появившихся «гербов» – оказалось следующим (в первой указано число x_i выпавших «гербов» в одном бросании монет; во второй строке – частоты m_i, т.е. число бросаний, при которых выпало x_i «гербов»):

x_i	0	1	2	3	4
m_i	8	20	42	22	8

Проверить гипотезу о биномиальном распределении случайной величины X при α = 0,05.

Вариант 9

В библиотеке случайно отобрано 200 выборок по 5 книг. Регистрировалось число поврежденных книг. В итоге получено следующее эмпирическое распределение книг (в первой указано число x_i поврежденных книг в одной выборке; во второй строке – частоты m_i, т.е. число выборок, содержащих x_i поврежденных книг):

x_i	0	1	2	3	4	5
m_i	72	77	34	14	2	1

Используя критерий Пирсона, при уровне значимости 0,05 проверить гипотезу о биномиальном распределении случайной величины X – число поврежденных книг – при α = 0,05.

Вариант 10

Произведено n = 100 опытов. Каждый опыт состоял из N = 10 испытаний, в каждом из которых вероятность p появления события A равна 0,4. В итоге получено следующее эмпирическое распределение (в первой указано число x_i появлений события A в одном опыте; во второй строке – частота m_i, т.е. число опытов, в которых наблюдалось x_i появлений события A):

x_i	0	1	2	3	4	5
m_i	3	10	27	34	23	8

Используя критерий Пирсона, при уровне значимости 0,05 проверить гипотезу о биномиальном распределении случайной величины X – число появлений события A – при α = 0,05.

Вариант 11

Масса (в граммах) произвольно выбранных 30 пачек полуфабриката «Геркулес» такова:
503, 509, 495, 493, 489, 485, 507, 511, 487, 495, 506, 504, 507, 511, 499, 491, 494, 518, 506, 515, 487, 509, 507, 488, 495, 490, 498, 497, 492, 495.

Можно ли при уровне значимости α = 0,05 утверждать, что случайная величина X – масса пачки – подчинена нормальному закону распределения?

Вариант 12

Результаты исследований удоя 100 коров на молочной ферме за лактационный период (в ц) (случайная величина X) представлены в виде сгруппированного статистического ряда:

x_i	4 – 6	6 – 8	8 – 10	10 – 12	12 – 14	14 – 16	16 – 18	18 – 20	20 – 22
m_i	3	3	6	11	15	20	14	18	10

Требуется проверить нулевую гипотезу о нормальном законе распределения удоя коров. Уровень значимости принять $\alpha = 0{,}05$.

Вариант 13

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с заданным эмпирическим распределением:

Номер интервала i	Граница интервала		Частота m_i
	x_i	x_{i+1}	
1	1	3	2
2	3	5	4
3	5	7	6
4	7	9	10
5	9	11	18
6	11	13	20
7	13	15	16
8	15	17	11
9	17	19	7
10	19	21	5
11	21	23	1
Сумма			$n = 100$

Вариант 14

Результаты исследований прочности на сжатие (случайная величина X) – 200 образцов бетона – представлены в виде сгруппированного статистического ряда:

Интервалы прочности, кг/см^2	Среднее значение интервала x_i	Частота m_i
190 – 200	195	10
200 – 210	205	26
210 – 220	215	56
220 – 230	225	64
230 – 240	235	30
240 – 250	245	14

Требуется проверить нулевую гипотезу о нормальном законе распределения прочности образцов бетона на сжатие. Уровень значимости принять α = 0,001.

Вариант 15

Результаты исследований 1000 экземпляров северной сосны по диаметру ствола (случайная величина X) представлены в виде сгруппированного статистического ряда:

x_i	14 – 18	18 – 22	22 – 26	26 – 30	30 – 34	34 – 38	38 – 42	42 – 46	46 – 50	50 –54	54 – 58
m_i	16	35	100	183	214	197	115	71	36	19	5

Требуется проверить нулевую гипотезу о нормальном законе распределения северной сосны по диаметру ствола. Уровень значимости принять α = 0,05.

Вариант 16

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с заданным эмпирическим распределением:

Номер интервала i	Граница интервала		Частота m_i
	x_i	x_{i+1}	
1	6	16	8
2	16	26	7
3	26	36	16
4	36	46	35
5	46	56	15
6	56	66	8
7	66	76	6
8	76	86	5
Сумма			n = 100

Вариант 17

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с заданным эмпирическим распределением:

Номер интервала i	Граница интервала		Частота m_i
	x_i	x_{i+1}	
1	5	10	7
2	10	15	8
3	15	20	15
4	20	25	18
5	25	30	23
6	30	35	19
7	35	40	14
8	40	45	10
9	45	50	6
Сумма			n = 120

Вариант 18

Результаты исследований 1000 нитей по крепости (случайная величина X) представлены в виде сгруппированного статистического ряда:

Крепость нити	Количество нитей
170 – 180	9
180 – 190	52
190 – 200	84
200 – 210	128
210 – 220	187
220 – 230	225
230 – 240	174
240 – 250	107
250 – 260	34
	$n = 1000$

Требуется проверить нулевую гипотезу о нормальном законе распределения нитей по крепости. Уровень значимости принять $\alpha = 0{,}05$.

Вариант 19

Результаты исследований 1000 женщин по росту (случайная величина X) представлены в виде сгруппированного статистического ряда:

Рост, см	Среднее значение интервала x_i	Частота m_i
134 – 137	135,5	1
137 – 140	138,5	4
140 – 143	141,5	16
143 – 146	144,5	53
146 – 149	147,5	121

149 – 152	150,5	193
152 – 155	153,5	229
155 – 158	156,5	186
158 – 161	159,5	121
161 – 164	162,5	53
164 – 167	165,5	17
167 – 170	168,5	5
170 – 173	171,5	1
		n = 1000

Требуется проверить нулевую гипотезу о нормальном законе распределения женщин по росту. Уровень значимости принять α = 0,05.

Вариант 20

Используя критерий Пирсона, при уровне значимости 0,05 проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности X с заданным эмпирическим распределением:

Номер интервала i	Граница интервала		Частота m_i
	x_i	x_{i+1}	
1	– 20	– 10	20
2	– 10	0	47
3	0	10	80
4	10	20	89
5	20	30	40
6	30	40	16
7	40	50	8
Сумма			n = 300

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Балдин, К.В. Теория вероятностей и математическая статистика: учебник / К.Б. Балдин, В.Н. Башлыков. – М.: Дашков и К, 2016. – 472 с.
2. Дегтярева, Н.А. Сборник задач по статистике: учебное пособие для студентов / Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 90 с.
3. Дегтярева, Н.А. Применение статистических методов исследования в сельском хозяйстве / Н.А. Дегтярева, Н.А. Берг // Известия высших учебных заведений. Уральский регион. – 2017. – № 1. – С. 42–47.
4. Мхитарян, В.С. Теория вероятностей и математическая статистика: учебник для студентов учреждений высшего профессионального образования / В.С. Мхитарян, В.Ф. Шишов, А.Ю. Козлов. – М.: ИЦ Академия, 2012. – 416 с.
5. Кибзун, А.И. Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами / А.И. Кибзун, Е.Р. Горяинова. – М.: Физматлит, 2013. – 232 с.
6. Дегтярева, Н.А. Эконометрические модели анализа и прогнозирования: монография / Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 170 с.
7. Ивченко, Г.И. Математическая статистика в задачах: около 650 задач с подробными решениями / Г.И. Ивченко, Ю.И. Медведев, А.В. Чистяков. – М.: Ленанд, 2015. – 320 с.

8. Ватутин, В.А. Теория вероятностей и математическая статистика в задачах / В.А. Ватутин, Г.И. Ивченко, Ю.И. Медведев. – М.: Ленанд, 2015. – 384 с.
9. Геворкян, П.С. Теория вероятностей и математическая статистика / П.С. Геворкян, А.В. Потемкин, И.М. Эйсымонт. – М.: Физматлит, 2016. – 176 с.
10. Ниворожкина, Л.И. Математическая статистика с элементами теории вероятностей в задачах с решениями: учебное пособие / Л.И. Ниворожкина, З.А. Морозова. – М.: Дашков и К, 2015. – 480 с.
11. Дегтярева, Н.А. Сборник задач по экономико-математическим методам и моделям: учебное пособие для студентов / Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 77 с.
12. Колемаев, В.А. Теория вероятностей и математическая статистика: учебник для вузов / В.А. Колемаев, В.Н. Калинина. – М.: ЮНИТИ-ДАНА, 2017. – 352 с.
13. Климов, Г.П. Теория вероятностей и математическая статистика: учебник / Г.П. Климов. – М.: Московский государственный университет имени М.В. Ломоносова, 2011. – 368 с.
14. Гмурман, В.Е. Теория вероятностей и математическая статистика: учебное пособие для вузов. – М.: Высш. шк., 2002. – 479 с.
15. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для вузов. – М.: Высш. шк., 2005. – 404 с
16. Прохоров, Ю.В. Лекции по теории вероятностей и математической статистике: учебник / Ю.В. Прохоров,

Л.С. Пономаренко. – М.: Московский государственный университет имени М.В. Ломоносова, 2012. – 254 с.
17. Щербакова, Ю.В. Теория вероятностей и математическая статистика: учебное пособие / Ю.В. Щербакова. – Саратов: Научная книга, 2012. – 159 с.
18. Гордеева, Д.С. Экономика образования: учебное пособие для студентов / Д.С. Гордеева, Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 101 с.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ 1

Значение *t*-критерия Стьюдента

v	Уровень значимости α			v	Уровень значимости α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	6,3137	12,7062	63,656	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9250	19	1,7291	2,0930	2,8609
3	2,3534	3,1824	5,8408	20	1,7247	2,0860	2,8453
4	2,1318	2,7765	4,6041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7970
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7531	2,1315	2,9467	60	1,6706	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6576	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758

ПРИЛОЖЕНИЕ 2

Значение *F*-критерия Фишера при уровне значимости 0,05

v_2	v_1							
	1	2	3	4	5	6	7	8
1	161	199	216	225	230	234	237	239
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59

ПРИЛОЖЕНИЕ 3

Таблица значений функции Лапласа

X	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	989	989	988	986	984	982	980	977	973
0,1	0,3970	965	961	956	951	945	939	932	925	918
0,2	0,3910	902	894	885	876	867	857	847	836	825
0,3	0,3814	802	790	778	765	752	739	726	712	697
0,4	0,3683	668	653	637	621	605	589	572	555	538
0,5	0,3521	503	485	467	448	429	410	391	372	352
0,6	0,3332	312	292	271	251	230	209	187	166	144
0,7	0,3123	101	079	056	034	011	2989	2966	2943	2920
0,8	0,2897	874	850	827	803	780	756	732	709	685
0,9	0,2661	637	613	589	565	541	516	492	468	444
1,0	0,2420	396	371	347	323	299	275	251	227	203
1,1	0,2179	155	131	107	083	059	036	012	1989	1965
1,2	0,1942	919	895	872	849	826	804	781	758	736
1,3	0,1714	691	669	647	626	604	582	561	539	518
1,4	0,1497	476	456	435	415	394	374	354	334	315
1,5	0,1295	276	257	238	219	200	182	163	145	127
1,6	0,1109	092	074	057	040	023	006	0989	0973	0957
1,7	0,0941	925	909	893	878	863	848	833	818	804
1,8	0,0790	775	761	748	734	721	707	694	681	669
1,9	0,0655	644	632	620	608	596	584	573	562	551
2,0	0,0541	529	519	508	498	488	478	468	459	449
2,1	0,0440	431	422	413	404	396	387	379	371	363
2,2	0,0355	347	339	332	325	317	310	303	297	290
2,3	0,0283	277	270	264	258	252	246	241	235	229
2,4	0,0224	210	213	208	203	198	194	189	184	180
2,5	0,0175	171	167	163	158	154	151	147	143	139
2,6	0,0136	132	129	126	122	119	116	113	110	107
2,7	0,0104	101	099	096	093	091	088	086	084	081

2,8	0,0079	077	075	073	071	069	067	065	063	061
2,9	0,0060	058	056	055	053	051	050	048	047	046
3,0	0,0044	043	042	040	039	038	037	036	035	034
3,1	0,0033	032	031	030	029	028	027	026	025	025
3,2	0,0024	023	022	022	021	020	020	019	018	018
3,3	0,0017	017	016	016	015	015	014	014	013	013
3,4	0,0012	012	012	011	011	010	010	010	009	009
3,5	0,0009	008	008	008	008	007	007	007	007	006
3,6	0,0006	006	006	006	005	005	005	005	005	004
3,7	0,0004	004	004	004	004	004	003	003	003	003
3,8	0,0003	003	003	003	003	002	002	002	002	002

ПРИЛОЖЕНИЕ 4

Таблица значений критерия χ^2-Пирсона

Число степеней свободы	Уровень значимости		Число степеней свободы	Уровень значимости	
	0,05	0,01		0,05	0,01
1	3,841	6,635	32	46,194	53,486
2	5,991	9,210	33	47,400	54,776
3	7,815	11,345	34	48,602	56,061
4	9,488	13,277	35	49,802	57,342
5	11,070	15,086	36	50,998	58,619
6	12,592	16,812	37	52,192	59,892
7	14,067	18,475	38	53,384	61,162
8	15,507	20,090	39	54,572	62,428
9	16,919	21,666	40	55,758	63,691
10	18,307	23,209	41	56,942	64,950
11	19,675	24,725	42	58,124	66,206
12	21,026	26,217	43	59,304	67,459
13	22,362	27,688	44	60,481	68,709
14	23,685	29,141	45	61,656	69,957
15	24,996	30,578	46	62,830	71,201
16	26,296	32,000	47	64,001	72,443
17	27,587	33,409	48	65,171	73,683
18	28,869	34,805	49	66,339	74,919
19	30,144	36,191	50	67,505	76,154
20	31,410	37,566	51	68,669	77,386
21	32,671	38,932	52	69,832	78,616
22	33,924	40,289	53	70,993	79,843
23	35,172	41,638	54	72,153	81,069
24	36,415	42,980	55	73,311	82,292

25	37,652	44,314	56	74,468	83,513
26	38,885	45,642	57	75,624	84,733
27	40,113	46,963	58	76,778	85,950
28	41,337	48,278	59	77,931	87,166
29	42,557	49,588	60	79,082	88,379
30	43,773	50,892	61	80,232	89,591
31	44,985	52,191	62	81,381	90,802

ПРИЛОЖЕНИЕ 5

Критические значения коэффициента корреляции Пирсона

$k = n - 2$	P		$k = n - 2$	P	
	0,05	0,01		0,05	0,01
5	0,75	0,87	27	0,37	0,47
6	0,71	0,63	28	0,36	0,46
7	0,67	0,80	29	0,36	0,46
8	0,63	0,77	30	0,35	0,45
9	0,60	0,74	за	0,33	0,42
10	0,66	0,71	40	0,30	0,39
11	0,65	0,66	45	0,29	0,37
12	0,63	0,66	50	0,27	0,35
13	0,51	0,64	60	0,25	0,33
14	0,50	0,62	70	0,23	0,30
15	0,48	0,61	80	0,22	0,28
16	0,47	0,59	90	0,21	0,27
17	0,46	0,58	100	0,20	0,25
18	0,44	0,66	125	0,17	0,23
19	0,43	0,55	160	0,16	0,21
20	0,42	0,54	200	0,14	0,18
21	0,41	0,53	300	0,11	0,15
22	0,40	0,52	400	0,10	0,13
23	0,40	0,51	500	0,09	0,12
24	0,39	0,50	700	0,07	0,10
25	0,38	0,49	900	0,06	0,09
26	0,37	0,48	1000	0,06	0,09

ПРИЛОЖЕНИЕ 6

Критические значения коэффициента корреляции рангов Спирмена

n	P		n	P		n	P	
	0,05	0,01		0,05	0,01		0 05	0,01
5	0,94	–	17	0,48	0,62	29	0,37	0,48
6	0,85	–	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,68	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,64	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,68	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,49	0,51	37	0,33	0,43
14	0,64	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,68	27	0,38	0,49	39	0,32	0,41
16	0,60	0,64	28	0,38	0,48	40	0,31	0,40

ПРИЛОЖЕНИЕ 7

Таблица значений $t_{\gamma} = t(\gamma, n)$

γ / n	0,95	0,99	0,999	γ / n	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3.50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3.25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3.11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3.01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,997	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

ПРИЛОЖЕНИЕ 8

Таблица значений $q = q(\gamma, n)$

γ n	0,95	0,99	0,999	γ n	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0.32	0.49	0.73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1.38	2,42	35	0.26	0.38	0.56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1.08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0.21	0,30	0,43
12	0,55	0.90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0.78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,211
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Учебное издание

Дегтярева Нина Адамовна
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
Учебно-практическое пособие

ISBN 978-5-6042129-5-0

Работа рекомендована РИС ЮУрГГПУ
Протокол № 17, 2018 г.
Издательство ЮУрГГПУ
454080 г. Челябинск, пр. Ленина, 69

Редактор Е.М. Сапегина
Технический редактор А.Г. Петрова
Эксперт Л.М. Свирская

Подписано в печать 17.09.2018
Объем 3,62 уч.-изд. л. (7,03 усл. п.л.)
Формат 60x84 1/16
Тираж 100 экз. Заказ №

Отпечатано с готового оригинал-макета
в типографии ЮУрГГПУ
454080 г. Челябинск, пр. Ленина, 69