МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«Курганский государственный университет»

Кафедра фундаментальной математики и методики преподавания математики

Математическая статистика
Часть 1

Материалы для практических занятий и самостоятельной работы
для студентов очной, очно-заочной и заочной форм обучения
37.03.01 «Психология»,37.05.02 «Психология служебной деятельности»

Курган 2018

Кафедра: «фундаментальной математики и методики преподавания математики»

Дисциплина: «Математическая статистика» 37.03.01 «Психология», 37.05.02«Психология служебной деятельности»

Составил: ст. преподаватель Е. А. Лукерьянова

Утверждены на заседании кафедры «31»августа 2018 г.

Рекомендованы методическим советом университета «20»декабря 2017 г.

Содержание

Введение

Математические методы широко применяются в различных областях науки и техники, поэтому изучение дисциплины «Математическая статистика» важно для студентов, обучающихся по специальности «Психология».

Предметом данной дисциплины являются математические методы обработки количественной информации. Наиболее распространенным методом обработки экспериментальных данных является выборочный метод.

Цель курса «Математическая статистика» – познакомить студентов с методами обработки экспериментальных данных, полученных в результате наблюдений над случайными явлениями или в результате специально поставленных экспериментов; привить им практические навыки анализа данных с целью получения выводов, характеризующих изучаемые случайные явления.

Основной задачей курса должно стать овладение методами организации сбора, обработки данных статистического наблюдения.

Особая роль в подготовке студентов к профессиональной деятельности принадлежит самостоятельной работе, организуемой в процессе обучения. Настоящие материалы предназначены для организации самостоятельной работы по изучению курса «Математическая статистика».

Материалы для практических занятий и самостоятельной работы включают в себя планы занятий, вопросы для экзамена. Планы занятий содержат вопросы для повторения, задачи для решения в аудитории различного уровня сложности, задачи для самостоятельного решения.

Краткое содержание дисциплины.

1 выборки и их характеристики.

2 статистическая оценка параметров распределения.

3 проверка статистических гипотез.

4 корреляционнорегрессионный анализ.

Материалы для практических занятий и самостоятельной работы составлено в соответствии с учебным планом по дисциплине «Математическая статистика.

Раздел 1. Выборки и их характеристики

Тема 1. Статистическое распределение выборки

Теоретическая справка

Задачи математической статистики

Установление закономерностей, которым подчинены массовые случайные явления, основано на изучении методами теории вероятностей статистических данных — результатов наблюдений.

Первая задача математической статистики — указать способы сбора и группировки статистических сведений, полученных в результате наблюдений или в результате специально поставленных экспериментов.

Вторая задача математической статистики — разработать методы анализа статистических данных в зависимости от целей исследования.

К ним относятся:

а) оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения, вид которого известен; оценка зависимости случайной величины от одной или нескольких случайных величин и др.
б) проверка статистических гипотез о виде неизвестного распределения или о величине параметров распределения, вид которого известен.

Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, если имеется партия деталей, то качественным признаком может служить стандартность детали, а количественным — контролируемый размер детали.

Иногда проводят сплошное обследование, т. е. обследуют каждый из объектов совокупности относительно признака, которым интересуются. На практике, однако, сплошное обследование применяют сравнительно редко. Например, если совокупность содержит очень большое число объектов, то провести сплошное обследование физически невозможно.

Если обследование объекта связано с его уничтожением или требует больших материальных затрат, то проводить сплошное обследование практически не имеет смысла. В таких случаях случайно отбирают из всей совокупности ограниченное число объектов и подвергают их изучению.

Выборочной совокупностью или просто *выборкой* называют совокупность случайно отобранных объектов.

Генеральной совокупностью называют совокупность объектов, из которых производится выборка.

Объемом совокупности (выборочной или генеральной) называют число объектов этой совокупности. Например, если из 1000 деталей отобрано для обследования 100 деталей, то объем генеральной совокупности N=1000, а объем выборки n =100.

Для того чтобы по данным выборки можно было достаточно уверенно судить об интересующем признаке генеральной совокупности, необходимо, чтобы объекты выборки правильно его представляли. Другими словами, выборка должна правильно представлять пропорции генеральной совокупности. Это требование коротко формулируют так: выборка должна быть *репрезентативной (представительной)*.

При составлении выборки можно поступать двумя способами: после того, как объект отобран и над ним произведено наблюдение, он может быть возвращен либо не возвращен в генеральную совокупность. В соответствии со сказанным выборки подразделяют на повторные и бесповторные.

Повторной называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность.

Бесповторной называют выборку, при которой отобранный объект в генеральную совокупность не возвращается.

На практике обычно пользуются бесповторным случайным отбором.

Пусть из генеральной совокупности извлечена выборка, причем x_1наблюдалось n_1раз, x_2— n_2раз, x_k—n_k раз и $\sum n_i=n$ объем выборки. Наблюдаемые значения x_iназывают *вариантами,* а последовательность вариант, записанных в возрастающем порядке, — *вариационным рядом.* Числа наблюдений называют *частотами,* а их отношения к объему выборки $n_i/n=W_i$ — *относительными частотами.*

Статистическим распределением выборки называют перечень вариант и соответствующих им частот или относительных частот. Статистическое распределение можно задать также в виде последовательности интервалов и соответствующих им частот (в качестве частоты, соответствующей интервалу, принимают сумму частот, попавших в этот интервал).

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

Пример 1. Исходные данные представлены в таблице 1.Написать распределение относительных частот.

Таблица 1 – Распределение частот выборки объёма n = 20

x_i2	6	12
n_i3	10	7

Решение. Найдем относительные частоты, для чего разделим частоты на объем выборки:

W_1 = 3/20=0,15, W_2= 10/20 = 0,50, W_3 = 7/20 = 0,35. Напишем распределение относительных частот (таблица 2).

Таблица 2 – Распределение относительных частот:

x_i	2	6	12
W_i	0,15	0,50	0,35

Контроль: 0,15+0,50 + 0,35=1.

Пусть известно статистическое распределение частот количественного признака X. Введем обозначения: n_x — число наблюдений, при которых наблюдалось значение признака, меньшее *x; n* — общее число наблюдений (объем выборки).

Ясно, что относительная частота события X<*x* равна n_x /n. Если *x* изменяется, то, вообще говоря, изменяется и относительная частота, т. е. относительная частота n_x /n есть функция от *x*. Так как эта функция находится эмпирическим (опытным) путем, то ее называют эмпирической.

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

Итак, по определению, $F^*(x)=n_x/n$, где n_x — число вариант, меньших x; n— объем выборки. Таким образом, для того чтобы найти, например, $F^*(x_2)$, надо число вариант, меньших x_2, разделить на объем выборки: $F^*(x_2)= n_{x2}/n$.

Пример 2. Построить эмпирическую функцию по данному распределению выборки (таблица 3).

Таблица 3 – Распределение частот

варианты x_i	2	6	10
частоты n_i	18	12	30

Решение. Найдем объем выборки: 12+18+30=60. Наименьшая варианта равна 2, следовательно, $F^*(x)=0$ при $x<2$.

Значение $X<6$, а именно $x_1=2$, наблюдалось 12 раз, следовательно, $F^*(x)=12/60=0{,}2$ при $2<x\le 6$. Значения $X<10$, а именно $x_1=2$ и

$x_2=12+18=30$ раз, следовательно, $F^*(x)=30/60=0{,}5$ при $6<x$.

Так как $x=10$ — наибольшая варианта, то $F^*(x)=1$ при $x>10$.

Для наглядности статистическое распределение представляют в виде полигона или гистограммы.

Полигоном частот называют ломаную, отрезки которой соединяют точки $(x_1;n_1)$, $(x_2;n_2)$, ... $(x_к;n_к)$. Для построения полигона частот на оси абсцисс откладывают варианты x_i, а на оси ординат — соответствующие им частоты n_i. Точки $(x_i;n_i)$ соединяют отрезками прямых и получают полигон частот.

Полигоном относительных частот называют ломаную, отрезки которой соединяют точки $(x_1;W_1)$, $(x_2;W_2)$,...,$(x_к;W_k)$. Для построения полигона относительных частот на оси абсцисс откладывают варианты x_iа на оси ординат — соответствующие им относительные частоты W_i. Точки $(x_i;W_i)$ соединяют отрезками прямых и получают полигон относительных частот.

В случае непрерывного признака целесообразно строить гистограмму, для чего интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько частичных интервалов длиной h и находят для каждого частичного интервала n_i — сумму частот вариант, попавших в i-й интервал.

Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению n_i/h (плотность частоты).

Для построения гистограммы частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии n_i/h.

Площадь i-го частичного прямоугольника равна $hn_i/h=n_i$ — частоте вариант i-го интервала, следовательно, площадь гистограммы частот равна сумме всех частот, т. е. объему выборки .

Для графического изображения кумулятивных рядов пользуются кумулятивными кривыми и огивами. Кумулятивная кривая строится следующим образом: на оси абсцисс отмечают точки, соответствующие границам интервалов или значениям признака, и в каждой такой точке восстанавливают перпендикуляр, длина которого пропорциональна накопительной частоте, концы соседних перпендикуляров соединяют отрезками. Полученная ломаная линия называется *кумулятивной кривой*. Если на горизонтальной оси откладывать накопительные частоты, а на вертикальной — значения признака, то полученная при этом кривая носит название *огивы*.

Вопросы для повторения

1 Что изучает математическая статистика, предмет математической статистики.
2 Основные задачи математической статистики.
3 Генеральная и выборочная совокупности.
4 Объем генеральной совокупности, объем выборочной совокупности.
5 Повторные и бесповторные выборки.
6 Репрезентативность выборки.
7 Вариационный ряд.
8 Статистическое распределение выборки.
9 Эмпирическая функция распределения и ее свойства.
10 Полигон и гистограмма.

Задачи для решения в аудитории

1 Для проведения демографических исследований выбрали 50 семей и получили следующие данные о количестве членов семьи:

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

Записать вариационный ряд исследуемого признака. Составить статистическое распределение выборки. Написать эмпирическую функцию распределения; построить полигон частот, график эмпирической функции.

2 Выборка задана в виде распределения частот (таблица 4).

Таблица 4 – Статистическое распределение выборки

xi	2	5	7
ni	1	3	6

Найти распределение относительных частот.

3 Известны данные о посевных площадях картофеля (тыс. гектаров) по районам Курганской области:

1,5; 1,5; 0,5; 1,2; 0,9; 0,9; 0,8; 0,5; 1,2; 1,1; 0,6; 1,1; 1,2; 0,9; 1,5; 1,2; 0,8; 0,4; 1; 0,1; 1,1; 0,8; 0,1; 1,5; 0,5; 1,2; 0,8; 0,3; 0,4; 1,3; 0,7; 0,1; 0,3; 1,6; 0,8.

Записать вариационный ряд. Составить статистическое распределение относительных частот, построить полигон относительных частот.

4 В 1889-1890 гг. был измерен рост 1000 взрослых мужчин (рабочих московских фабрик). Результаты измерений указаны в таблице 5

Таблица 5 − Рост взрослых мужчин

Рост (см)	Число мужчин	Рост (см)	Число мужчин
143-146	1	167-170	170
146-149	2	170-173	120
149-152	8	173-176	64
152-155	26	176-179	28
155-158	65	179-182	10
158-161	120	182-185	3
161-164	181	185-188	1
164-167	201		Всего 1000

Построить полигон частот, гистограмму частот.

5 Построить гистограмму относительных частот следующего распределения (таблица 6).

Таблица6 – Статистическое распределение выборки

Частичный интервал длиной h	Сумма частот вариант частичного интервала ni
2-5	9
5-8	10
8-11	25
11-14	6

6 Средняя месячная зарплата (тыс. рублей) за год каждого из пятидесяти случайно отобранных работников хозяйства такова:

317,304, 230, 285, 290, 320, 262, 279, 205, 180, 234, 221, 241, 270, 257, 290, 258, 296, 301, 150, 160, 210, 235 ,308, 240, 370, 180, 244 ,365, 130, 170, 250, 370, 267, 288, 231, 253, 315, 370, 180, 244, 365, 130, 170, 250, 370, 267, 288, 231, 253, 315, 201, 256, 279, 285, 226, 367, 247, 252, 320, 160, 215, 350.

Построить интервальное статистическое распределение частот. Построить гистограмму частот.

7 Средняя температура мая в г. Москве измерялась в течение 40 лет:

12; 12; 12; 12; 12,8; 13,8; 13,1; 13; 10,8; 11,3; 12; 13; 10,9; 10; 11,5; 13; 13,9; 14,2; 14; 14; 13,9; 15; 14,9; 14,9; 10,1; 10; 10; 12; 12,4; 11; 13; 14,2; 15; 15; 15,5; 15,9; 16; 15,9; 16; 16,9.

Записать вариационный ряд, составить статистическое распределение частот. Построить полигон частот. Найти эмпирическую функцию распределения, построить ее график.

8 По сведениям автоинспекции количество дорожных происшествий на улицах города N в октябре было таким:

6,8,10,7,6,11,9,8,7,11,7,8,10,6,11,9,7,7,8,12,10,6,6,6,5,4,9,9,8,10,7.

Записать вариационный ряд, составить статистическое распределение относительных частот. Построить полигон относительных частот. Записать эмпирическую функцию ее график.

Тема 2. Характеристики вариационного ряда

Теоретическая справка

Пусть изучается генеральная совокупность относительно количественного признака X.

Генеральной средней называют среднее арифметическое значений признака генеральной совокупности.

Если все значения признака различны, то

$$X_Г=\frac{x_1+x_2+\cdots+x_N}{N}.$$

Если значения признака имеют частоты $N_1, N_2, ..., N_k$, где $N_1, +N_2+...+N_k=N$, то

$$X_Г=\frac{X_1N_1+X_2N_2+\cdots X_KN_K}{N}.$$

Пусть для изучения генеральной совокупности относительно количественного признака X извлечена выборка объема n.

Выборочной средней называют среднее арифметическое значение признака выборочной совокупности.

Если все значения признака выборки различны, то

$$X_В=\frac{X_1+X_2+\cdots+X_n}{n}$$

если же все значения имеют частоты $n_1, n_2, ..., n_k$, то

$$X_В=\frac{X_1n_1+X_2n_2+\cdots X_Kn_K}{n}.$$

Выборочная средняя является несмещенной и состоятельной оценкой генеральной средней.

Если выборка представлена интервальным вариационным рядом, то за x_i принимают середины частичных интервалов.

Для того чтобы охарактеризовать рассеяние значений количественного признака *X* генеральной совокупности вокруг своего среднего значения, вводят сводную характеристику – генеральную дисперсию.

Генеральной дисперсией Dг называют среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения.

Если все значения признака генеральной совокупности объема N различны, то

$$D_Г=\frac{\sum_{i=1}^{N}\left(x_i-\overline{x_Г}\right)^2}{N}.$$

Если же значения признака имеют соответственно частоты $N_1, N_2, ..., N_k$, где $N_1, +N_2+...+N_k=N$, то

$$D_Г=\frac{\sum_{i=1}^{N}N_i\left(x_i-\overline{x_Г}\right)^2}{N}.$$

Кроме дисперсии для характеристики рассеяния значений признака генеральной совокупности вокруг своего среднего значения пользуются сводной характеристикой – средним квадратическим отклонением.

Генеральным средним квадратическим отклонением (стандартом) называют квадратный корень из генеральной дисперсии: $\sigma_{г}=\sqrt{D_{г}}$.

Для того чтобы наблюдать рассеяние количественного признака значений выборки вокруг своего среднего значения, вводят сводную характеристику – выборочную дисперсию.

Выборочной дисперсией называют среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения.

Если значения признака выборки имеют частоты $n_1, n_2, ..., n_k$, то

$$D_B = \frac{\sum_{i=1}^{k}(x_i - \bar{x}_B)^2 \cdot n_i}{n}.$$

Для характеристики рассеивания значений признака выборки вокруг своего среднего значения пользуются сводной характеристикой – средним квадратическим отклонением. Выборочным средним квадратическим отклонением называют квадратный корень из выборочной дисперсии:

$$\sigma_{в}=\sqrt{D_{в}}.$$

Вычисление дисперсии (выборочной или генеральной) можно упростить, используя следующую теорему.

Теорема. Дисперсия равна среднему квадратов значений признака минус квадрат общей средней:

$$D = \overline{x^2} - \left[\overline{x}\right]^2.$$

Если выборка представлена интервальным вариационным рядом, то за x_i принимают середины частичных интервалов.

Модой M_0 называют варианту, которая имеет наибольшую частоту.

Например:

x_i	1	4	7	9
n_i	5	1	20	6

$M_0 = 7$

Медианой m_0 называют варианту, которая делит вариационный ряд на две части, равные по числу вариант.

Если $n = 2k + 1$, то $m_e = x_{k+1}$.

Если $n = 2k$, то $m_e = \frac{x_k + x_{k+1}}{2}$.

Например: для вариационного ряда 2, 3, 5, 6, 7 медиана равна 5, а для вариационного ряда 2, 3, 5, 6, 7, 9 медиана равна 5,5.

Размахом варьирования R называют разность между наибольшей и наименьшей вариантами.

$R = x_{max} - x_{min}$.

Так, для вариационного ряда 1,3, 4, 5, 6, 10 размах варьирования равен 9.

Средним абсолютным отклонением Θ **называют среднее арифметическое абсолютных** отклонений.

$$\Theta = \frac{\sum_{i=1}^{k} n_i |x_i - \bar{x}_B|}{n}.$$

Так, если

x_i	1	3	6	16
n_i	4	10	5	1

то $\bar{x}_B = \frac{4\cdot 1+10\cdot 3+5\cdot 6+1\cdot 16}{20}$=4, $\Theta = \frac{4\cdot|1-4|+10\cdot|3-4|+5\cdot|6-4|+1\cdot|16-4|}{20}$=2,2.

Коэффициентом вариации Vназывают выраженное в процентах отношение выборочного среднего квадратического отклонения к выборочной средней.

$$V = \frac{\sigma_B}{\bar{x}_B} \cdot 100\%.$$

Коэффициент вариации используется для сравнения рассеяния двух вариационных рядов: тот из рядов имеет большее рассеяние, у которого коэффициент вариации больше.

Для предыдущего примера V=71,25% , так как $\bar{x}_В = 4, \sigma_В = \sqrt{8,1} \approx 2,85$.

$D_B = \frac{\sum_{i=1}^{4}(x_i - \bar{x}_B)^2 \cdot n_i}{n}$,$D_B = \frac{-12+10+20+144}{20} = 8,1$.

$\sigma_B = \sqrt{D_B}$, $v = \frac{2,85}{4} \cdot 100\% \approx 71,25\%$.

При построении интервального вариационного ряда необходимо помнить:

1) число интервалов k можно определить по формулам:

$$k \approx 1 + 3,3 \cdot \ln n,$$
$$k \approx 2 \cdot \ln n,$$
$$k \approx log_2 n + 1;$$

2) длина интервала $h \approx \frac{2}{k}$, при этом h икподбирают так, чтобы $h \cdot k = R$;

3) за начало первого интервала можно взять x_{min} или $x_{min} - \frac{h}{2}$ (x_{min} принимают за середину первого интервала).

Образцы решения типовых задач

1 По данному распределению выборки (таблица 7) дать характеристику распределения признака, вычислить для этого:

1) $\overline{x}$, 2) медиану, 3) дисперсию, среднее квадратическое отклонение, размах варьирования, среднее абсолютное отклонение, коэффициент вариации, 4) моду. Построить полигон и гистограмму частот.

Таблица 7 – Распределение выборки

Частичный интервал	1−5	5−9	9−13	13−17	17−21
Сумма частот вариант частичного интервала	10	20	50	12	8

Решение

Составим вспомогательную таблицу центров интервалов (таблица 8).

Таблица 8 – Центры интервалов

Середина интервала	3	7	11	15	19
Число интервалов	10	20	50	12	8

1) $\bar{x}_\beta = \frac{3\cdot10+7\cdot20+11\cdot50+15\cdot12+19\cdot8}{100} = \frac{1052}{100} = 10{,}52; \bar{x}_\beta = \frac{\sum_{i=1}^{k} n_i x_i}{n}$;

2) вычислим выборочную дисперсию,

$D_B = \frac{\sum_{i=1}^{k} n_i (x_i - x_B)^2}{n}$;

$D_B = \frac{1}{100}(10\cdot(3-10{,}52)^2 + 20\cdot(7-10{,}52)^2 + 50\cdot(11-10{,}52)^2 + 12\cdot(15-10{,}52)^2 + +8\cdot(19-10{,}52)^2) = 16{,}4096$;

Дисперсию можно вычислить по формуле: $D_B = \overline{x^2} - (\bar{x}_B)^2$.

$D_B = \frac{1}{100}(9\cdot10 + 49\cdot20 + 121\cdot50 + 225\cdot12 + 361\cdot8) - (10{,}52)^2 = 16{,}4096$.

3) $\sigma_B = \sqrt{D_B}$, $\sigma_B = \sqrt{16?4096} \approx 4{,}051;$

4) R=$x_{max} - x_{min}$, R= $19-3 = 16$;

5) $\Theta = \frac{\sum_{i=1}^{k} n_i\,|x_i - \bar{x}_B|}{n}$;

$\Theta = \frac{1}{100}(10 \cdot |3 - 10.52| + 20 \cdot |7 - 10.52| + 50 \cdot |11 - 10.52| + 12 \cdot |15 - 10.52| + 8 \cdot |19 - 10.52|) = 3.092$;

6) $= \frac{\sigma_{\text{В}}}{\bar{x}_{\text{в}}} \cdot 100\%$, $V = \frac{4{,}051}{10{,}52} \cdot 100\% = 38{,}5076\%$, $V \approx 38{,}51\%$;

7) $M_0 = 11$;

8) $m_e = 11$.

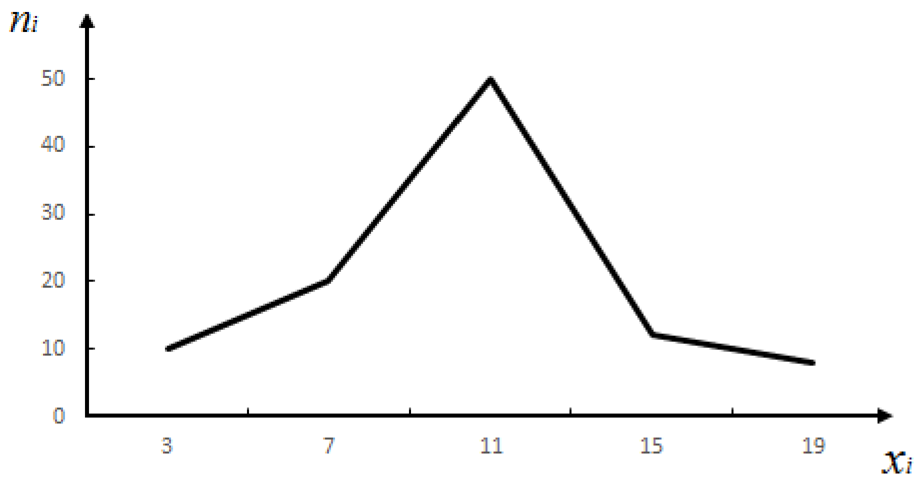

Рисунок 1 – Полигон частот

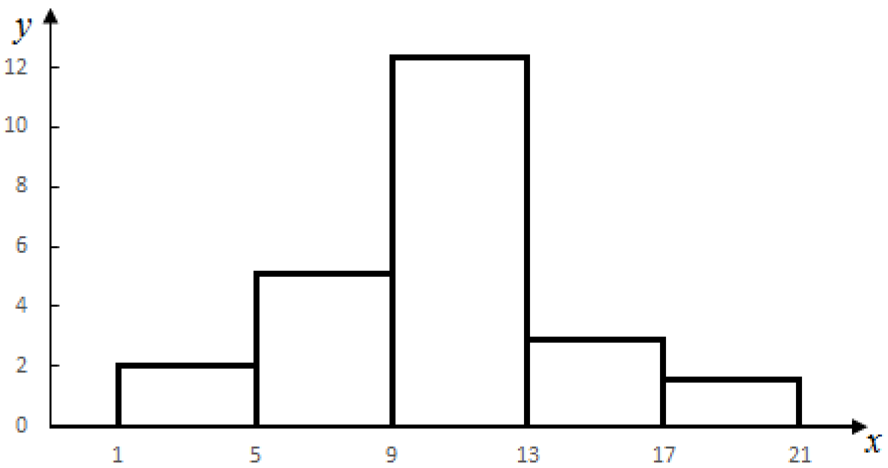

Рисунок 2 – Гистограмма

Вопросы для повторения

1 Генеральная и выборочная средние.

2 Генеральная и выборочная дисперсии.

3 Другие характеристики вариационного ряда (мода, медиана, размах варьирования, среднее абсолютное отклонение, коэффициент вариации).

Задачи для решения в аудитории

1 Генеральная совокупность задана таблицей распределения (таблица 9).

Таблица 9 − Статистическое распределение выборки

xi	1000	1200	1400
ni	1000	6000	3000

Найти генеральную среднюю хг и генеральную дисперсию.

2 На обследование каждого из десяти автомобилей было затрачено следующее время (в мин.):

25, 30, 22, 22, 54, 36 ,41, 45, 25 ,40.

Найти: а) среднее время, затраченное на обследование автомобиля,

б) дисперсию исследуемого признака.

3 Найти выборочную среднюю и выборочную дисперсию по данным таблицы10.

Таблица 10 − Число выпавших осадков в городе N

Месяц	Число выпавших осадков с сентября по февраль в городе N (в мм)	Месяц	Число выпавших осадков с сентября по февраль в городе N (в мм)
сентябрь	500	декабрь	1000
октябрь	450	январь	1200
ноябрь	200	февраль	2000

4 Выборочная совокупность задана таблицей распределения (таблица 11).

Таблица 11 – Статистическое распределение выборки.

x_i	1	2	5	8	9
n_i	3	4	6	4	3

Найти выборочную и исправленную дисперсии.

5 Выборочным путем были получены следующие данные об урожайности ржи в совхозе (таблица 12)

Таблица 12 – Урожайность ржи в совхозе

Урожайность, ц/га	Число, га
18	10
20	20
21	20

Определить выборочную среднюю *хв* и исправленное среднее квадратическое отклонение s.

6 По выборке объема n=51 найдена выборочная дисперсия Дв=5.

7 В таблице 13 приведены результаты измерения роста (в см) случайно отобранных 100 студентов.

Таблица 13 – Рост (в см) случайно отобранных 100 студентов

Рост	154-158	158-162	162-166	166-170	170-174	174-178	178-182
Число студентов	10	14	26	28	12	8	2

Найти выборочную среднюю и выборочную дисперсию роста обследованных студентов.

8 Найти выборочную среднюю, выборочную дисперсию по данному распределению выборки объема n=100 (таблица 14).

Таблица 14 – Статистическое распределение выборки

x_i	1250	1275	1280	1300
n_i	20	25	50	5

9 Уровень воды в реке по отношению к номиналу изменялся в течение 44 весенних паводков, данные измерений приведены в таблице15.

Таблица 15 – Уровень воды в реке по отношению к номиналу

Уровень (в см)	0-24	25-49	50-74	75-99	100-124	125-149	150-174	175-199	200-299	300-400
Число случаев	1	1	2	6	7	6	5	4	8	4

Найти выборочную среднюю и выборочную дисперсию уровня паводков.

10 Средняя температура июня в Ярославле измерялась в течение 40 лет. Ниже приведены данные .

10,8; 11,3; 12; 13; 10,9; 10; 11,3; 13; 10; 10; 14,7; 15; 15; 14; 13; 10,8; 10; 13; 11,3; 10; 13; 14,7; 14,7; 15; 13; 15; 15; 11,3; 10,9; 13; 15; 15; 14,7; 11,3; 11,3; 15; 15; 10,9; 13; 13.

Дать характеристику распределения признака, вычислив для этого:

1) выборочную среднюю; 2) медиану; 3) дисперсию, среднее квадратическое отклонение, размах варьирования, среднее абсолютное отклонение, коэффициент вариации; 4) моду. Построить полигон частот.

11 В ОТК были измерены диаметры 200 валиков из партии, изготовленной одним станком-автоматом. Отклонения измеренных диаметров от номинала даны в таблице 16 (в микронах).

Таблица 16 – диаметры 200 валиков

Границы уклонений	(-20)-(15)	(-15)-(-10)	(-10)-(-5)	(-5)-0	0-5	5-10	10-15	15-20	20-25	25-30
Число валиков	7	11	15	24	49	41	26	17	7	3

Произвести анализ статистических данных. Построить гистограмму частот.

Задачи для самостоятельного решения

1 Средняя температура в декабре в городе Саратове измерялась в течение 30 лет и приведены следующие данные:

-14,8; -16,9; -20; -1,1; -14,8; -20; -19,2; -14,8; -19,6; -11,1; -9,4; -16,9; -19,2; -19,2; -19,6; -11,1; -9,4; -1,1; -19,6; -19,2; -11; -11,1; -19,6; -9,4; -9,4; -1,1; -20; -20; -25; -25.

Произвести анализ статистических данных. Построить полигон частот.

2 О ходе заготовки сочных кормов в хозяйствах области поступили следующие данные, которые представлены в таблице 17.

Таблица 17 – Данные о ходе заготовки сочных кормов в хозяйствах области

Заготовлено центнеров на условную голову скота	0,75	1,25	1,75	2,25	2,75
Количество хозяйств	8	10	14	9	2

Преобразовать данный вариационный ряд в интервальный с равными интервалами так, чтобы приведенные значения вариант были центрами, дать характеристику распределения признака.

3 Имеются данные о размерах основных фондов (в млн руб.)

30 предприятий:

0,42; 0,24; 0,49; 0,67; 0,45; 0,27; 0,39; 0,21; 0,58; 0,4; 0,28; 0,73; 0,44; 0,66; 1,21; 0,62; 0,7; 0,81; 1,07; 0,68; 0,94; 0,76; 0,63; 0,88; 0,65; 1,14; 0,46; 1,38; 0,72; 0,91.

Построить интервальный вариационный ряд (h=0,2), приняв начало первого интервала равным 0,2 (млн. руб.). Вычислить $x_в$, $Д_в$ двумя способами: по исходным данным и по данным построенного ряда.

4 Из большой партии лампочек было отобрано в случайном порядке 400 штук для определения средней продолжительности горения. Выборочная средняя продолжительность горения лампочки оказалась равной 1220 ч. Найти с вероятностью 0,9973 среднюю продолжительность горения лампочки во всей партии, если среднее квадратическое отклонение продолжительности горения равно 35ч.

5 Из партии однотипных высокоомных сопротивлений отобрано для контроля 10 штук. Измерения дали следующие отклонения от номинала (таблица 18)

Таблица 18– отклонения от номинала высокоомных сопротивлений (в килоомах).

№ сопротивления	1	2	3	4	5	6	7	8	9	10
Отклонение	1	3	-2	2	4	2	5	3	-2	4

Найти выборочные средние и дисперсию отклонения фактической величины сопротивления от номинала в этой партии и определить точность оценки МО этим выборочным средним с надежностью 0,95.

Раздел 2.Статистическая оценка параметров распределения

Тема3. Точечные оценки генеральной средней и генеральной дисперсии. Интервальное оценивание

Теоретическая справка

Пусть требуется изучить количественный признак генеральной совокупности. Допустим, что из теоретических соображений удалось установить, какое именно распределение имеет признак. Возникает задача оценки параметров, которыми определяется это распределение.

Статистической оценкой неизвестного параметра теоретического распределения называют функцию от наблюдаемых случайных величин.

$$Q^* = f(x_1, x_2, ..., x_n),$$

где $x_1, x_2, ..., x_n$ – выборка.

Для того чтобы статистические оценки давали «хорошие» приближения оцениваемых параметров, они должны удовлетворять определенным требованиям: оценка должна быть несмещенной, эффективной и состоятельной.

Несмещенной называют статистическую оценку Q^*, математическое ожидание которой равно оцениваемому параметру Q при любом объеме выборки, т. е. $M(Q^*) = Q$.

Смещенной называют оценку, математическое ожидание которой не равно оцениваемому параметру.

Эффективной называют статистическую оценку, которая при заданном объеме выборки n имеет наименьшую возможную дисперсию.

При рассмотрении выборок большого объема (n велико) к статистическим оценкам предъявляется требование состоятельности.

Состоятельной называют статистическую оценку, которая при $n \to \infty$ стремится по вероятности к оцениваемому параметру. Например, если дисперсия несмещенной оценки при $n \to \infty$ стремится к нулю, то такая оценка оказывается состоятельной.

Выборочное среднее является несмещенной оценкой для генеральной средней.

Выборочная дисперсия является смещенной оценкой генеральной дисперсии, т.е. математическое ожидание выборочной дисперсии не равно оцениваемой генеральной дисперсии, а равно

$$\mathrm{M}(D_{\text{г}}) = \frac{n-1}{n} D_{\text{в}}.$$

Для исправления выборочной дисперсии достаточно умножить ее на дробь $\frac{n}{n-1}$, получим исправленную дисперсию, которую обычно обозначают через S^2. Исправленная дисперсия является несмещенной оценкой. В качестве оценки генеральной дисперсии принимают исправленную дисперсию. Для оценки среднего квадратического отклонения совокупности используют исправленное среднее квадратическое отклонение $S = \sqrt{S^2}$. Формулы для вычисления выборочной дисперсии и исправленной дисперсии отличаются только знаменателями. При достаточно больших n выборочная и исправленная дисперсии мало отличаются, поэтому на практике исправленной дисперсией пользуются, если $n<30$. Кроме выборочной средней и выборочной дисперсии, применяются и другие характеристики вариационного ряда.

Интервальной называют оценку, которая определяется двумя числами – концами интервала. Интервальные оценки позволяют установить точность и надежность оценок.

Пусть найденная по данным выборки статистическая характеристика Θ^* служит оценкой неизвестного параметра Θ. Будем считать Θ постоянным числом (Θ может быть и случайной величиной). Ясно, что Θ^* тем точнее определяет параметр Θ, чем меньше абсолютная величина разности $\left|\Theta-\Theta^*\right|$. Другими словами, если $\delta>0$ и $\left|\Theta-\Theta^*\right|<\delta$, то чем меньше δ, тем оценка точнее.

Таким образом, положительное число δ характеризует точность оценки.

Однако статистические методы не позволяют категорически утверждать, что оценка Θ^* удовлетворяет неравенству $\left|\Theta-\Theta^*\right|<\delta$, можно лишь говорить о вероятности γ, с которой это неравенство осуществляется.

Надежностью (доверительной вероятностью) оценки называют вероятность γ, с которой осуществляется неравенство $\left|\Theta-\Theta^*\right|<\delta$.

Обычно надежность оценки задается наперед, причем в качестве γ берут число, близкое к единице. Наиболее часто задают надежность, равную 0,95; 0,99 и 0,999.

Пусть вероятность того, что $\left|\Theta-\Theta^*\right|<\delta$ равна γ:

$$P\left(\left|\Theta-\Theta^*\right|<\delta\right)=\gamma.$$

Заменив неравенство равносильным ему двойным неравенством, получим:

$$P\left[\Theta^*-\delta<\Theta<\Theta^*+\delta\right]=\gamma.$$

Это соотношение следует понимать так: вероятность того, что интервал $\Theta^*-\delta<\Theta<\Theta^*+\delta$ заключает в себе (покрывает) неизвестный параметр Θ, равна γ. Интервал $\left(\Theta^*-\delta,\Theta^*+\delta\right)$ называется доверительным интервалом, который покрывает неизвестный параметр с надежностью γ.

Доверительный интервал для оценки математического ожидания при известном σ

Пусть количественный признак генеральной совокупности распределен нормально. Известно среднее квадратическое отклонение этого распределения – σ. Требуется оценить математическое ожидание *a* по выборочной средней. Найдем доверительный интервал, покрывающий *a* с надежностью γ. Выборочную среднюю $\bar{x}$ будем рассматривать как случайную величину $\bar{X}$ ($\bar{x}$ изменяется от выборки к выборке), выборочные значения признака $x_1,x_2,...,x_n$ – как одинаково распределенные независимые СВ с математическим ожиданием каждой *a* и средним квадратическим отклонением γ. Примем без доказательства, что если величина *X* распределена нормально, то и выборочная средняя тоже распределена нормально с параметрами: $M(\bar{X})=a,\ \sigma(\bar{X})=\dfrac{\sigma}{\sqrt{n}}$.

Потребуем, чтобы выполнялось равенство

$P\left(\left|\overline{X}-a\right|<\delta\right)=\gamma.$ Пользуясь формулой $P\left(\left|X-a\right|<\delta\right)=2\Phi(\delta/\sigma),$ заменив X на $\overline{X}$ и σ на $\sigma(\overline{X})=\sigma/\sqrt{n},$ получим $P\left(\left|\overline{X}-a\right|<t\delta/\sqrt{n}\right)=2\Phi(\delta\sqrt{n}/\sigma)=2\Phi(t),$ где $t=\delta\sqrt{n}/\sigma.$

Найдя из предыдущего равенства $\delta=t\sigma/\sqrt{n},$ получим окончательную формулу:

$$P\left(\overline{x}-\frac{t\sigma}{\sqrt{n}}<a<\overline{x}+\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t)=\gamma.$$

Число t определяется из равенства $\Phi(t)=\frac{\gamma}{2}$ по таблице функции Лапласа.

Доверительный интервал для оценки математического ожидания при неизвестном σ

По данным выборки можно построить случайную величину (ее возможные значения будем обозначать через *t)*:

$T=\frac{\overline{X}-a}{S/\sqrt{n}}$, которая имеет распределение Стьюдента.

Плотность распределения Стьюдента $S(t,n)=B_n\left[1+\frac{t^2}{n-1}\right]^{-n/2}$,

где $B_n=\frac{\Gamma(n/2)}{\sqrt{\pi(n-1)}\Gamma((n-1)/2)}.$

Поскольку $S(t,n)$ – четная функция, можно записать двойное неравенство:

$$P\left(\overline{X}-t_\gamma S/\sqrt{n}<a<\overline{X}-t_\gamma S/\sqrt{n}\right)=\gamma.$$

Доверительный интервал $\left(\overline{x}-t_\gamma s/\sqrt{n},\ \overline{x}-t_\gamma s/\sqrt{n}\right)=\gamma.$ $\overline{x},\ s$ вычисляются по выборке; t_γ – табличная величина.

Доверительный интервал для оценки среднего квадратического отклонения σ нормального распределения

Требуется оценить неизвестное генеральное среднее квадратическое отклонение σ по исправленной дисперсии *s*, т.е. найти доверительные интервалы, покрывающие параметр σ с заданной надежностью γ.

Потребуем выполнения соотношения $P\left(s-\delta<\sigma<s+\delta\right)=\gamma.$

Преобразуем двойное неравенство $s-\delta<\sigma<s+\delta$ в равносильное неравенство $s(1-\delta/s)<\sigma<s(1+\delta/s).$ Положив $\frac{\delta}{s}=q$, получим доверительный интервал:

$s(1-q)<\sigma<s(1+q).$

Для отыскания q пользуются таблицей значений $q=q(\gamma,n).$

Интервальная оценка вероятности (биноминального распределения) по относительной частоте

Сначала вычислим точечную оценку.

В качестве точечной оценки неизвестной вероятности p появления события A принимают относительную частоту $W = \frac{m}{n}$, где m – число появлений события A, n – число испытаний. Найдем дисперсию оценки, приняв во внимание, что $D(m) = npq : D(W) = D\left[m/n\right] = D(m)/n^2 = npq/n^2 = pq/n.$

q – вероятность не появления события А.Среднее квадратическое отклонение $\sigma_W = \sqrt{pq/n}$.

Интервальная оценка. Для биноминального распределения формулу $P\left(\left|X-a\right|<\delta\right) = 2\Phi(\delta/\sigma)$ можно записать в виде: $P\left(\left|W-p\right|<\delta\right) = 2\Phi(\delta/\sigma_W)$.

Потребуем выполнения этого соотношения с надежностью γ.

$P\left(\left|W-p\right|<\delta\right) = 2\Phi(\delta/\sigma_W) = \gamma.$

Подставляя $\sigma_W = \sqrt{pq/n}$,

получим $P\left(\left|W-p\right|<\delta\right) = 2\Phi(\delta\sqrt{n}/\sqrt{pq}) = 2\Phi(t) = \gamma,$ где $t = \sqrt{n}/\sqrt{pq}.$

Отсюда $\delta = t\sqrt{pq/n}$ и, следовательно, $P\left(\left|W-p\right|<t\sqrt{pq/n}\right) = 2\Phi(t) = \gamma.$

Случайную величину W заменим неслучайной наблюдаемой относительной частотой w и подставим $1-p$ вместо q.

$\left|w-p\right| < t\sqrt{p(1-p)/n}$.

Решим это неравенство относительно p. Возведем обе части неравенства в квадрат. Получим

$\left[\left(t^2/n\right)+1\right]p^2 - 2\left[w+\left(t^2/n\right)\right]p + w^2 < 0.$

Найдем корни трехчлена.

Меньший корень $p_1 = \frac{n}{t^2+n}\left[w + \frac{t^2}{2n} - t\sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2}\right].$

Больший корень $p_1 = \frac{n}{t^2+n}\left[w + \frac{t^2}{2n} + t\sqrt{\frac{w(1-w)}{n} + \left(\frac{t}{2n}\right)^2}\right].$

Искомый доверительный интервал $p_1 < p < p_2.$

Образцы решения задач

1 Изучается вопрос об урожайности овощей в хозяйствах области. Результаты приведены в таблице 19.

Таблица 19 – Результаты урожайности овощей в хозяйствах области

Урожайность, ц/га	25	75	125	175	225	275	325
Количество хозяйств	11	23	26	17	20	2	1

Найти граничные значения урожайности по области с надежностью 0, 97.

Решение

Граничные значения средней урожайности по области являются границами доверительного интервала $\left(\bar{x}_{\text{В}} \cdot \frac{t \cdot \sigma}{\sqrt{n}}, \bar{x}_B + \frac{t \cdot \sigma}{\sqrt{n}}\right)$.

$$\bar{x}_B = \frac{25 \cdot 11 + 75 \cdot 23 + 125 \cdot 26 + 175 \cdot 17 + 225 \cdot 20 + 275 \cdot 2 + 325}{100}$$
$$= \frac{275 + 1725 + 3250 + 2975 + 4500 + 550 + 325}{100} = 136$$

$\sigma = \sqrt{\bar{x}^2 - (\bar{x})^2}$;

$$\overline{x^2} = \frac{25^2 \cdot 11 + 75^2 \cdot 23 + 125^2 \cdot 26 + 175^2 \cdot 17 + 225^2 \cdot 20 + 275^2 \cdot 2 + 325^2}{100}$$
$$== \frac{6875 + 129375 + 406250 + 520625 + 1012500 + 105625}{100} =$$
$$= 21812{,}5;$$

$\sigma = \sqrt{21812{,}5 - 136^2} \approx 57{,}6.$

Найдем tкак аргумент $\Phi(t) = \frac{0{,}97}{2}, \Phi(t) = 0{,}485,\ t = 2,17$(приложение А)

$$\delta = \frac{t \cdot \sigma}{\sqrt{n}}, \delta = \frac{2{,}17 \cdot 27{,}6}{\sqrt{100}} = \frac{2{,}17 \cdot 27{,}6}{10} = 12{,}4992, \qquad \delta = 12{,}5.$$

Границы доверительного интервала

$$\bar{x}_{\text{В}} - \delta = 136 - 12{,}5 = 123{,}5; \bar{x}_B + \delta = 136 + 12{,}5 = 148{,}5.$$

Округляя значения, получим $a = 136 \pm 13$, т.е. (123,149) –доверительный интервал, в котором находится средняя урожайность овощей по области с вероятностью 97%.

2 Из генеральной совокупности извлечена выборка объема 10 (таблица 20).

Таблица 20 – Выборка объема n= 10.

x_i	*-1*	*1*	*2*	*3*	*4*	*5*
n_i	*1*	*2*	*2*	*2*	*1*	*2*

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

Решение

Искомый доверительный интервал имеет вид:

$$\left(\bar{x}_{\text{в}-t_1} \cdot \frac{s}{\sqrt{n}} < a < \bar{x}_{B+t_1} \cdot \frac{s}{\sqrt{n}}\right).$$

Найдем t_1. Пользуясь таблицей (приложение А), учитывая $\gamma = 0{,}95, n = 10$, находим $t_1 = 2{,}26$.

$$\delta = \frac{t_1 \cdot s}{\sqrt{n}}, \delta = \frac{2{,}26 \cdot 2}{\sqrt{10}} = \frac{4{,}52}{\sqrt{10}} = 1{,}4125.$$

Найдем выборочную среднюю и исправленное среднее квадратическое отклонение по формулам: $\bar{x}_{\text{В}}=\frac{\sum_{i=1}^{k} n_i x_i}{n}, s = \sqrt{\frac{\sum_{i=1}^{k} n_i (x_i - \bar{x}_B)^2}{n-1}}$;

$$\bar{x}_{\text{В}} = \frac{-1 \cdot 1 + 1 \cdot 2 + 2 \cdot 2 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 2}{10} = 2{,}5;$$

$$s = \sqrt{\frac{12{,}25 + 4{,}5 + 0{,}5 + 0{,}5 + 2{,}25 + 12{,}5}{9}} = \sqrt{\frac{32{,}5}{9}} \approx 2.$$

$$\bar{x}_{B-\delta} = 2{,}5 - 1{,}4125 = 1{,}0875; \bar{x}_{B-\delta} = 2{,}5 + 1{,}4125 = 3{,}9125.$$

Округляя результаты, получим доверительный интервал $1{,}1 < a < 3{,}9$,покрывающий неизвестное математическое ожидание с надежностью 0,95.

3 По данным выборки объема $n = 15$ 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 *0,999*.

Решение

Задача сводится к отысканию доверительного интервала

$$s\,(1 - q) < \sigma < s\,(1 + q), \text{если } q < 1 \text{ или } 0 < \sigma < s(1 + q), \text{если } q > 1.$$

По данным $\gamma = 0{,}999$ и $n = 15$, по таблице (приложение А) найдем $q = 1{,}15$.

Так как $q < 1$, то подставив $s = 1\,, q = 1{,}15$ в соотношение $0 < \sigma < s(1 + q)$, получим искомый доверительный интервал $0 < \sigma < 2{,}15$.

Вопросы для повторения

1 Понятие точечной оценки, интервальной оценки.

2 Доверительная вероятность (надежность) оценки.

3 Доверительный интервал.

4 Оценка МО нормально распределенного признака при известном σ.

5 Оценка МО нормально распределенного признака при неизвестном σ. Оценка истинного значения измеряемой величины.

6 Доверительный интервал для оценки среднего квадратического отклонения нормального распределения. Оценка точности измерений.

Задачи для решения в аудитории

1 Месячная зарплата (тыс. рублей) за год каждого из пятидесяти случайно отобранных работников такова:

317, 304, 230, 285, 290, 320, 262, 274, 205, 180, 234, 221, 241, 270, 257, 290, 258, 296, 301, 150, 160, 210, 235, 308, 240, 370, 180, 244, 365, 130, 170, 250, 370, 267, 288, 231, 253, 315, 201, 256, 279, 285, 226, 367, 247, 252, 320, 160, 215, 350.

Найти: а) среднюю месячную зарплату, б) доверительный интервал для средней зарплаты с надежностью 0,99.

2 Из генеральной совокупности извлечена выборка объема n=10

(таблица 21).

Таблица 21 – Статистическое распределение выборки

Вариант xi	-2	1	2	3	4	5
Частота ni	2	1	2	2	2	1

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

3 Найти минимальный объем выборки, при котором с надежностью 0,975 точность оценки МО генеральной совокупности по выборочной средней будет равна δ=0,3, если известно среднее квадратическое отклонение σ=1,2 нормально распределенной генеральной совокупности.

4 При стократном повторении опыта событие А наступило 62 раза с надежностью 0,95. Оцените неизвестную вероятность события А.

5 По данным выборки объема n из генеральной совокупности нормально распределенного количественного признака найдено «исправленное» среднее квадратическое отклонение s. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надежностью 0,999, если: а) n=10, s=5,1: б) n=50, s=14.

6 Глубина моря измеряется прибором, систематическая ошибка которого равна 0, а случайные ошибки распределены нормально со средним квадратическим отклонением σ=15м. Сколько надо сделать независимых измерений, чтобы определить глубину с ошибкой не более 5м при доверительной вероятности 90%.

7 Произведено 5 независимых измерений для определения заряда электрона и получены следующие результаты (в электростатистических единицах):

$4{,}781*10^{-10}$; $4{,}785*10^{-10}$; $4{,}769*10^{-10}$; $4{,}792*10^{-10}$; $4{,}779*10^{-10}$.

Определить «подходящее» значение для заряда электрона и найти доверительные границы при надежности 99%

8 При анализе содержащегося SiO_2 в шлаке были получены следующие результаты: 28,6; 28,3; 28,4; 28,4. Найти оценку истинного содержимого SiO_2 в шлаке и точность оценки с доверительной вероятностью 0,95.

9 По 13 равноточным измерениям найдено исправленное среднее квадратическое отклонение s=0,12. Найти точность измерений σ с надежностью 0,99.

Задачи для самостоятельного решения

1 По данным 9 независимых равноточных измерений физической величины найдены среднее арифметическое результатов отдельных измерений xв=42,319 и исправленное среднее квадратическое отклонение s=5. Требуется оценить истинное значение измеряемой величины с надежностью 0,95. Предполагается, что результаты измерений распределены нормально.

2 По 15 равноточным измерениям найдено исправленное среднее квадратическое отклонение s=0,12. Найти точность измерений с надежностью 0,99.

3 Средняя температура июня в г. Москве измерялась в течение 20 лет:

15; 15; 15,5; 15,9; 16; 15,9; 16; 16,9;13,8; 16; 13,9; 14,7; 13; 15; 17,2; 18; 18; 19; 20; 23.

Найти выборочную среднюю июньских температур в Москве, доверительный интервал для выборочной средней с надежностью 0,95.

4 По данным выборки объема n=30 из генеральной совокупности нормально распределенного количественного признака найдено исправленное среднее квадратическое отклонение s=5,6. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надежностью 0,999.

5 Пробными испытаниями установлено, что относительная погрешность газоанализаторов данного типа равна 12%. Сколько дублирующих газоанализаторов надо поставить, чтобы обеспечить относительную погрешность результата в 10% и в 5%?

6 Сделана выборка из 10 деталей, обработанных на токарном станке. Отклонения диаметров валика этих деталей от середины поля допуска представлены в таблице 22.

Таблица 22 – Отклонения диаметров валика

Номера деталей	1	2	3	4	5	6	7	8	9	10
Отклонение размеров	2	1	-2	3	2	4	-2	5	3	4

Найти доверительную вероятность, с которой оценка математического ожидания отклонений, полученная по выборке, будет отличаться от его истинного значения по абсолютной величине не больше, чем на 21 мм.

7 По данным 16 независимых равноточных измерений физической величины найдены среднее арифметическое результатов отдельных измерений хв=23,261 и исправленное среднее квадратическое отклонение = 0,4. Требуется оценить истинное значение измеряемой величины и точность измерений с надежностью 0,95.

Примерные задания для рубежного контроля 1 (проверочная самостоятельная работа)

1 Дано распределение частот по размеру проданной мужской обуви (таблица 23).

Таблица 23 – Исходные данные

Размер обуви (x_i)	36	37	38	39	40	41	42	43
Число проданных пар (n_i)	2	1	5	8	17	21	18	8

Найти: 1) распределение относительных частот; 2) эмпирическую функцию распределения.
Построить: 1) полигон относительных частот; 2) график эмпирической функции распределения.

2 Дать характеристику распределения признака по данным таблицы 24.

Таблица 24 – Данные о стаже 9 рабочих в производственной бригаде

Рабочие	1	2	3	4	5	6	7	8	9
Производственный стаж (лет)	2	5,5	9	10	10,5	12	15	17	21

3 Из генеральной совокупности извлечена выборка объема n = 12 (таблица 25).

Таблица 25 – Исходные данные

x_i	– 0,5	– 0,4	– 0,2	0	0,2	0,6	0,8	1	1,2	1,5
n_i	1	2	1	1	1	1	1	1	2	1

Определить с надежностью 0,95 математическое ожидание a нормально распределенного признака генеральной совокупности при помощи доверительного интервала.

Список литературы

1 Баврин, И. И. Высшая математика [Текст] / И. И. Баврин. – Москва : Просвещение, 2004.

2 Владимирский, Б. М. Математика. Общий курс [Текст] : учебник для бакалавров / Б. М. Владимирский. – Москва : Просвещение, 2008.

3 Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистики : учебное пособие для студентов [Текст]/ В. Е. Гмурман. – Москва : Высшая школа, 2006.

4 Ильин, В. А. Высшая математика [Текст]/ В. А. Ильин. – Москва : Проспект, 2005.

5 Карелина, И. Г. Математика [Текст] / И. Г. Карелина. – Воронеж : Воронежский государственный университет, 2004.

6 Колемаев, В. А., Калинин В. Н. Теория вероятностей и математическая статистка [Текст] / В. А. Колемаев, В. Н. Калинин. – Москва : Инфра, 1997.

7 Лобоцкая, Н. Л. Основы высшей математики [Текст] / Н. Л. Лобоцкая. – Минск : Высшая школа, 2000.

8 Маркович, Э. С. Курс высшей математики с элементами теории вероятностей и математической статистики [Текст] / Э. С. Маркович. – Москва : Высшая школа, 1982.

9 Минорский, В. П. Сборник задач по высшей математике [Текст] / В. П. Минорский. – Москва : Изд-во физ.-мат., 2001.

10 Крамер, Г. Математические методы статистики [Текст] / Г. Крамер – Москва : Наука, 1975.

11 Кремер, Н. М. Теория вероятностей и математическая статистика [Текст] / Н. М. Кремер. – Москва : Наука, 2000.

Приложения

Приложение А

Статистические таблицы

Таблица А 1 – Значения функции $\varphi(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{x^2}{2}}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблица А 2 – Значения функции $\Phi(x)=\frac{1}{\sqrt{2\pi}}\cdot\int\limits_0^x e^{-\frac{z^2}{2}}dz$

	Ф(*x*)	***x***	**Ф(*x*)**	***x***	**Ф(*x*)**	***x***	**Ф(*x*)**
0,00							
0,01	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,02	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,03	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,04	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,05	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,06	0,0199	0,37	0,1443	0,69	0,2549	1,01	0,3438
0,07	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,08	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,09	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,10	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,11	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,12	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,13	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,14	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,15	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,16	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,17	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,18	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,19	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,20	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,21	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,22	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,23	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,24	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,25	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,26	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,27	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,28	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,29	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,30	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,31	0,1179	0,62	0,2324	0,94	0,3264		
	0,1217	0,63	0,2357	0,95	0,3289		

Продолжение таблицы А2

1,26	0,3962	1,59	*0,4441*	1,92	0,4726	*2,50*	0,4938
1,27	0,3980	1,60	*0,4452*	1,93	0,4732	*2,52*	0,4941
1,28	0,3997	1,61	*0,4463*	1,94	0,4738	*2,54*	0,4945
1,29	0,4015	1,62	*0,4474*	1,95	0,4744	*2,56*	0,4948
1,30	0,4032	1,63	*0,4484*	1,96	0,4750	*2,58*	0,4951
1,31	0,4049	1,64	*0,4495*	1,97	0,4756	*2,60*	0,4953
1,32	0,4066	1,65	*0,4505*	1,98	0,4761	*2,62*	0,4956
1,33	0,4082	1,66	*0,4515*	1,99	0,4767	*2,64*	0,4959
1,34	0,4099	1,67	*0,4525*	2,00	0,4772	*2,66*	0,4961
1,35	0,4115	1,68	*0,4535*	2,02	0,4783	*2,68*	0,4963
1,36	0,4131	1,69	*0,4545*	2,04	0,47,93	*2,70*	0,4965
1,37	0,4147	1,70	*0,4554*	2,06	0,4803	*2,72*	0,4967
1,38	0,4162	1,71	*0,4564*	2,08	0,4812	*2,74*	0,4969
1,39	0,4177	1,72	*0,4573*	2,10	0,4821	*2,76*	0,4971
1,40	0,4192	1,73	*0,4582*	2,12	0,4830	*2,78*	0,4973
1,41	0,4207	1,74	*0,4591*	2,14	0,4838	*2,80*	0,4974
1,42	0,4222	1,75	*0,4599*	2,16	0,4846	*2,82*	0,4976
1,43	0,4236	1,76	*0,4608*	2,18	0,4854	*2,84*	0,4977
1,44	0,4251	1,77	*0,4616*	2,20	0,4861	*2,86*	0,4979
1,45	0,4265	1,78	*0,4625*	2,22	0,4868	*2,88*	0,4980
1,46	0,4279	1,79	*0,4633*	2,24	0,4875	*2,90*	0,4981
1,47	0,4292	1,80	*0,4641*	2,26	0,4881	*2,92*	0,4982
1,48	0,4306	1,81	*0,4649*	2,28	0,4887	*2,94*	0,4984
1,49	0,4319	1,82	*0,4656*	2,30	0,4893	*2,96*	0,4985
1,50	0,4332	1,83	*0,4664*	2,32	0,4898	*2,98*	0,4986
1,51	0,4345	1,84	*0,4671*	2,34	0,4904	*3,00*	0,49865
1,52	0,4357	1,85	*0,4678*	2,36	0,4909	*3,20*	0,49931
1,53	0,4370	1,86	*0,4686*	2,38	0,4913	*3,40*	0,49966
1,54	0,4382	1,87	*0,4693*	2,40	0,4918	*3,60*	0,499841
1,55	0,4394	1,88	*0,4699*	2,42	0,4922	*3,80*	0,499928
1,56	0,4406	1,89	*0,4706*	2,44	0,4927	*4,00*	0,499968
1,57	0,4418	1,90	*0,4713*	2,46	0,4931	*4,50*	0,499997
1,58	0,4429	1,91	*0,4719*	2,48	0,4934	*5,00*	0,499997

Таблица А 3 – Значения $q = q(\gamma, n)$

n \ γ	**0,95**	**0,99**	**0,999**
5	1,37	2,67	5,64
6	1,09	2,01	3,88
7	0,92	1,62	2,98
8	0,80	1,38	2,42
9	0,71	1,20	2,06
10	0,65	1,08	1,80
11	0,59	0,98	1,60
12	0,55	0,90	1,45
13	0,52	0,83	1,33
14	0,48	0,78	1,23
15	0,46	0,73	1,15
16	0,44	0,70	1,07
17	0,42	0,66	1,01
18	0,40	0,63	0,96
19	0,39	0,60	0,92

n \ γ	**0,95**	**0,99**	**0,999**
20	0,37	0,58	0,88
25	0,32	0,49	0,73
30	0,28	0,43	0,63
35	0,26	0,38	0,56
40	0,24	0,35	0,50
45	0,22	0,32	0,46
50	0,21	0,30	0,43
60	0,188	0,269	0,38
70	0,174	0,245	0,34
80	0,161	0,226	0,31
90	0,151	0,211	0,29
100	0,143	0,198	0,27
150	0,115	0,160	0,211
200	0,099	0,136	0,185
250	0,089	0,120	0,162

Таблица А 4 – Критические точки распределения χ^2

Число степеней свободы *k*	Уровень значимости α					
	0,01	0,025	0,05	0,95	0,975	0,89
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Таблица А 5 – Значения $t_\gamma = t(\gamma, n)$

γ / ***n***	**0,95**	**0,99**	**0,999**
5	2,78	4,60	8,61
6	2,57	4,03	6,86
7	2,45	3,71	5,96
8	2,37	3,50	5,41
9	2,31	2,36	5,04
10	2,26	3,25	4,78
11	2,23	3,17	4,59
12	2,20	3,11	4,44
13	2,18	3,06	4,32
14	2,16	3,01	4,22
15	2,15	2,98	4,14
16	2,13	2,95	4,07
17	2,12	2,92	4,02
18	2,11	2,90	3,97
19	2,10	2,88	3,92

γ / ***n***	**0,95**	**0,99**	**0,999**
20	2,093	2,861	3,883
25	2,064	2,797	3,745
30	2,045	2,756	3,659
35	2,032	2,720	3,600
40	2,023	2,708	3,558
45	2,016	2,692	3,527
50	2,009	2,679	3,502
60	2,001	2,662	3,464
70	1,996	2,649	3,439
80	1,001	2,640	3,418
90	1,987	2,633	3,403
100	1,984	2,627	3,392
120	1,980	2,617	3,374
∞	1,960	2,576	3,291

Лукерьянова Елена Александровна

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

(Часть 1)

Материалы для практических занятий и самостоятельной работы
для студентов очной, очно-заочной и заочной форм обучения
37.03.01 «Психология», 37.05.02 « Психология служебной деятельности»

Редактор Л. П. Чукомина

Подписано в печать 28.01.19	Формат 60 х 84 1/16	Бумага 65 г/м^2
Печать цифровая	Усл. печ. л 2,5	Уч.-изд.2,5
Заказ №31	Тираж 25	Не для продажи

БИЦ Курганского государственного университета.
640020, г. Курган, ул. Советская, 63/4.
Курганский государственный университет.