МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
Учреждение образования
«Витебский государственный технологический университет»

МАТЕМАТИКА.
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Практикум

для студентов второго курса специальностей
1-36 01 01 «Технология машиностроения», 1-53 01 01-01 «Автоматизация технологических процессов и производств (машиностроение и приборостроение)», 1-43 01 07 «Техническая эксплуатация энергооборудования организаций», 1-40 05 01-01 «Информационные системы и технологии (в проектировании и производстве)», 1-28 01 01«Экономика электронного бизнеса», 1-40 05 01-10 «Информационные системы и технологии (в бизнес-менеджменте)»

Витебск
2023

УДК 517 (076.1) (075.8)

Составители:

А. В. Коваленко, А. А. Джежора, А. П. Дмитриев,
Ю. А. Завацкий, Е. Ю. Вардомацкая

Рекомендовано к изданию редакционно-издательским
советом УО «ВГТУ», протокол № 3 от 30.11.2022.

Математика. Математическая статистика : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2023. – 94 с.

Практикум содержит основные теоретические сведения, задания к практическим занятиям, примеры для самостоятельного выполнения заданий, вопросы к экзамену по двум разделам курса «Математика» для студентов специальностей 1-36 01 01, 1-43 01 07, дисциплины «Высшая математика» для студентов специальности 1-53 01 01-01 и дисциплины «Теория вероятностей и математическая статистика» для студентов специальностей 1-28 01 01, 1-40 05 01-01, 1-40 05 01-10: статистическая обработка данных и элементы ререссионного анализа. Данное издание предназначено для проведения практических занятий у студентов второго курса факультетов «Информационные технологии и робототехника» и «Экономика и бизнес-управление», а также может быть использовано в ходе изучения указанных тем студентами заочной и дистанционной форм обучения.

УДК 517 (076.1) (075.8)

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Методические указания составлены на основе практических занятий, которые авторы проводили на протяжении многих лет преподавания дисциплин «Математика», «Высшая математика» и «Теория вероятностей и математическая статистика» в Витебском государственном технологическом университете. Приведённый материал проверен на нескольких поколениях студентов и содержит необходимые сведения для будущих специалистов механико-информационных и экономико-информационных специальностей. Среди рассмотренных в практикуме типовых примеров есть задачи, имеющие практическую направленность и связанные с дисциплинами, которые будут изучать студенты в следующих семестрах.

Данные учебно-методические материалы предназначены для студентов четырёх специальностей факультета информационных технологий и робототехники и двух специальностей факультета экономики и бизнес-управления. В работе приведены теоретические вопросы для сдачи экзамена, содержание и тематика практических занятий по указанным курсам. Методические указания написаны в соответствии с учебной программой дисциплин «Математика», «Высшая математика» и «Теория вероятностей и математическая статистика» для студентов механико-технологических и экономических специальностей второго года обучения.

В практикуме рассмотрены два раздела курсов «Математика», «Высшая математика», «Теория вероятностей и математическая статистика»: статистическая обработка данных и элементы регрессионного анализа. Каждая тема практикума представляет собой методический материал для проведения практического занятия, содержит решения типовых примеров и подборку рекомендуемых к решению задач по теме занятия, а также задания для выполнения контролируемой самостоятельной работы. В начале каждого раздела приведён краткий теоретический материал, который необходимо знать студенту при подготовке к аудиторной и самостоятельной работе по заданной теме. Прежде чем приступать к решению задач практического занятия или выполнению контролируемой самостоятельной работы, студенту необходимо изучить теоретический курс лекционного материала или обратиться к академическим изданиям для более детального изучения разделов курса, которые его интересуют. Наименование тем практикума, а также их структура построены в соответствии с учебными программами дисциплин «Математика» «Высшая математика» и «Теория вероятностей и математическая статистика» для студентов специальностей: 1-36 01 01, 1-43 01 07, 1-53 01 01-01, 1-28 01 01, 1-40 05 01-01, 1-40 05 01-10. Студенты заочной формы обучения могут применять изложенный в практикуме теоретический и практический материал для самостоятельной работы по предмету и выполнению контрольных заданий. Предложенная методическая разработка поможет студентам подготовиться к прохождению теста по отдельным темам и разделам курса, так как проведение экзамена подразумевает электронный контроль знаний.

ПЕРЕЧЕНЬ ВОПРОСОВ УЧЕБНОЙ ПРОГРАММЫ ПО КУРСАМ «МАТЕМАТИКА», «ВЫСШАЯ МАТЕМАТИКА», «ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА» ДЛЯ СПЕЦИАЛЬНОСТЕЙ 1-36 01 01, 1-43 01 07, 1-53 01 01-01, 1-28 01 01, 1-40 05 01-01, 1-40 05 01-10 (ВТОРОЙ КУРС)

1. Вероятностный эксперимент. Предмет и задачи теории вероятностей.
2. Основные формулы комбинаторики.
3. Пространство элементарных событий. Алгебра событий.
4. Классификация событий.
5. Аксиоматическое, классическое, геометрическое и статистическое определение вероятностей случайных событий.
6. Свойства вероятностей случайных событий.
7. Условная вероятность. Теорема умножения вероятностей.
8. Формула полной вероятности.
9. Формула Байеса.
10. Последовательность независимых испытаний. Схема Бернулли.
11. Наивероятнейшее число.
12. Теорема Пуассона.
13. Локальная теорема Муавра – Лапласа.
14. Интегральная теорема Муавра – Лапласа.
15. Закон больших чисел в схеме Бернулли.
16. Понятие случайной величины.
17. Функция распределения, её свойства.
18. Плотность вероятности, её свойства.
19. Дискретные случайные величины.
20. Непрерывные случайные величины.
21. Характеристики, описывающие центр распределения случайной величины (математическое ожидание, мода и медиана), их свойства.
22. Характеристики, описывающие рассеивание случайной величины (дисперсия и среднее квадратическое отклонение), их свойства.
23. Одноточечное и двухточечное распределение.
24. Геометрическое и гипергеометрическое распределение.
25. Биномиальный закон распределения.
26. Равномерный закон распределения.
27. Показательный закон распределения.
28. Нормальный закон распределения.
29. Правило трех сигм для нормального закона распределения.
30. Определение многомерных случайных величин. Понятие о моделях распределения многомерных случайных величин.
31. Распределение вероятностей многомерных дискретных случайных величин.
32. Функция распределения многомерной случайной величины.

33. Непрерывные многомерные случайные величины. Плотность распределения вероятностей многомерной случайной величины.

34. Распределения составляющих многомерной случайной величины.

35. Условные распределения составляющих многомерных случайных величин.

36. Зависимые и независимые случайные величины.

37. Числовые характеристики многомерных случайных величин.

38. Функции (преобразования) двумерной случайной величины.

39. Двумерное нормальное распределение.

40. Предмет и задачи математической статистики.

41. Генеральная и выборочная совокупность.

42. Статистический ряд. Статистическое распределение случайной величины.

43. Эмпирическая функция распределения.

44. Графическое изображение статистических рядов.

45. Классификация точечных оценок.

46. Метод моментов.

47. Метод наибольшего правдоподобия.

48. Интервальные оценки параметров распределения. Точность нахождения точечных оценок. Доверительная вероятность и доверительный интервал.

49. Доверительные интервалы для математического ожидания случайной величины, имеющей нормальное распределение при известном среднем квадратическим отклонении.

50. Доверительные интервалы для математического ожидания случайной величины, имеющей нормальное распределение при неизвестном среднем квадратическим отклонении.

51. Доверительные интервалы для среднего квадратического отклонения случайной величины, имеющей нормальное распределение.

52. Статистическая проверка параметрических гипотез.

53. Статистический критерий значимости проверки нулевой гипотезы.

54. Статистическая проверка непараметрических гипотез.

55. Критерий согласия χ^2.

56. Критерий согласия λ Колмогорова.

57. Основные задачи регрессионного и корреляционного анализа.

58. Линейная регрессия.

59. Коэффициент корреляции.

60. Линейный корреляционный анализ. Модели регрессий.

ПРАКТИКУМ ПО РЕШЕНИЮ ЗАДАЧ

1 СТАТИСТИЧЕСКОЕ ОПИСАНИЕ РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

Содержание: генеральная и выборочная совокупность, статистический ряд, статистическое распределение случайной величины, эмпирическая функция распределения, графическое распределение статистических рядов.

1.1 Теоретический материал по теме практического занятия

При изучении теории вероятностей предполагалось, что вероятность наступления базовых событий известна. Исходя из этого предположения строилась математическая модель вероятностного эксперимента и на её основании вычислялись вероятности более сложных событий. Теория вероятности даёт возможность вычислять вероятности наступления событий, которые могут произойти при условии осуществления данного комплекса условий. На самом деле припроведении вероятностного эксперимента мы можем не знать, осуществился ли в действительности предполагаемый комплекс условий. В этом смысле задачи математической статистики обратны задачам теории вероятностей.

Математическая статистика оценивает структуру математических моделей, предложенных для анализа вероятностного эксперимента. Основными задачами математической статистики являются: планирование, описание, анализ и предсказание. Одним из основных методов описания и анализа данных эксперимента является выборочный метод. Рассмотрим его основные понятия.

Построим математическую модель эксперимента. Для этого продублируем эксперимент в одинаковых условиях независимым образом n раз. В результате опыта множество исходов эксперимента, которое описывается случайной величиной ξ, будет образовывать некоторое конечное множество $\{x_1, x_2, ..., x_n\}$ из множества возможных значений Ω_ξ случайной величины ξ. Наблюдаемые значения x_i случайной величины ξ называются *вариантами*. Множество Ω_ξ возможных значений случайной величины ξ называется *генеральной совокупностью*, а число элементов множества Ω_ξ – *объёмом генеральной совокупности*. Множество $\{x_1, x_2, ..., x_n\}$ называется *выборкой* или *выборочной совокупностью*, а число элементов, входящих в выборку, называется *объёмом выборки*. Выборка должна быть *репрезентативной* или представительной, то есть должна хорошо отображать свойства генеральной совокупности Ω_ξ.

Исходы эксперимента дискретной случайной величины можно представить в виде таблицы (или *простого статистического ряда*), которая состоит из двух строк, в первой из которой указываются номера эксперимента, а во второй приводятся значения исходов эксперимента. Простой статистический ряд является первоначальной формой статистического ряда.

i – номер эксперимента	1	2	...	n
x_i – исход эксперимента	x_1	x_2	...	x_n

Для оценки распределения случайной величины ξ производят группировку исходных данных. Если исходные данные случайной величины ξ, которые принимают значения $x_1, x_2, ..., x_n$, располагаются в порядке возрастания, то статистический ряд называется вариационным. Определим частоты m_i или частости $\omega_i = m_i/n$ (относительные частоты) появления одинаковых значений случайной величины ξ. В результате получаем сгруппированные статистические ряды следующего вида:

x_i – исход эксперимента	x_1	x_2	...	x_k
m_i – частоты	m_1	m_2	...	m_k
x_i – исход эксперимента	x_1	x_2	...	x_k
m_i/n – частости	m_1/n	m_2/n	...	m_k/n

При записи вариационных рядов необходимо учитывать следующие равенства: $\sum_{i=1}^{k} m_i = n$, $\sum_{i=1}^{k} m_i/n = 1$.

Предположим, что случайная величина ξ является непрерывной. Группировка данных заключается в разбиении интервала наблюдаемых значений случайной величины ξ на k частичных интервалов равной длины h: $[x_0; x_1)$, $[x_1; x_2)$, …, $[x_{k-1}; x_k]$ и нахождении частоты m_i или частости m_i/n попадания наблюдаемых значений в частичные интервалы. Количество интервалов выбирается произвольно, обычно не меньше трёх, но не более двадцати. Для определения числа интервалов разбиения может быть использована формула $k = 1 + 3{,}3222 \cdot \lg N$, где N – количество элементов совокупности. В случае равных интервалов величина интервала может быть определена по формуле

$$h = \frac{x_{\max} - x_{\min}}{k} = \frac{x_{\max} - x_{\min}}{1 + 3{,}3222 \cdot \lg N}.$$

В результате составляются интервальные статистические ряды следующего вида $\left(\sum_{i=1}^{k} m_i = n,\ \sum_{i=1}^{k} m_i/n = 1 \right)$:

Интервалы наблюдаемых значений СВ ξ	$[x_0;x_1)$	$[x_1;x_2)$	…	$[x_{k-1};x_k]$
m_i – частоты	m_1	m_2	…	m_k

Интервалы наблюдаемых значений СВ ξ	$[x_0;x_1)$	$[x_1;x_2)$	…	$[x_{k-1};x_k]$
m_i/n – частости	m_1/n	m_2/n	…	m_k/n

Определение 1.1.1 Перечень наблюдаемых значений случайной величины ξ (или интервалов наблюдаемых значений) и соответствующие им частостей $\omega_i = m_i/n$ называется статистическим законом распределения случайной величины ξ.

Частость события $B_i = [x_{i-1};x_i)$ является случайной величиной с математическим ожиданием, равным вероятности этого события, то есть $M(m_i/n) = P(x_{i-1} < \xi < x_i) = p_i$. Закон больших чисел в схеме Бернулли утверждает, что если эксперимент повторяется n раз при одинаковых условиях, то частость события сходится к вероятности этого же события: $\frac{m_i}{n} \xrightarrow{P} p_i$. Поэтому во второй строке интервального статистического ряда стоят оценки (приближённые значения) вероятностей p_i.

Определение 1.1.2 *Эмпирической функцией распределения* случайной величины ξ называется функция $F^*(x)$, определяющая для каждого значения x частость события $(\xi < x)$:

$$F^*(x) = \frac{n_x}{n}, \tag{1.1.1}$$

где n_x – число значений x_i, меньших значения x; n – объём выборки.

Из теоремы Бернулли следует, что при большом объёме выборки эмпирическая функция распределения $F^*(x)$ и функция распределения случайной величины $F(x) = P(\xi < x)$ мало отличаются друг от друга. Отличие эмпирической функции распределения от теоретической состоит в том, что теоретическая функция распределения определяет вероятность события $(\xi < x)$, а эмпирическая функция определяет относительную частоту этого же события. На рисунке 1.1.1 изображён график эмпирической функции распределения некоторой случайной величины ξ.

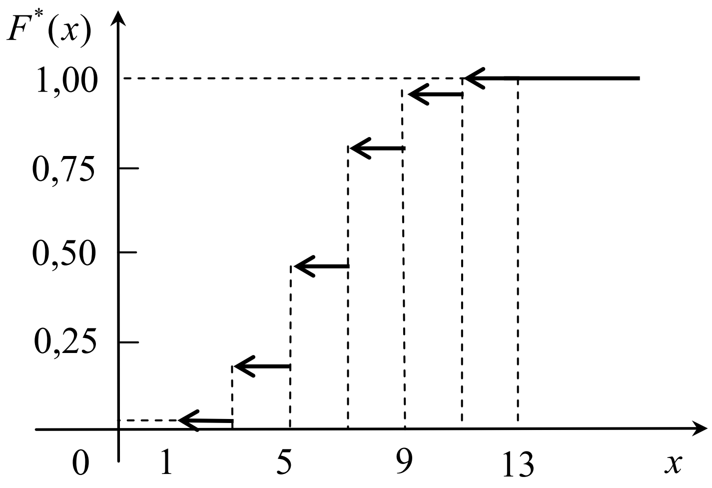

Рисунок 1.1.1 – Эмпирическая функция распределения случайной величины ξ

Для эмпирической функции распределения справедливы все свойства интегральной функции распределения. Действительно, из определения эмпирической функции распределения следует, что:

1) значения эмпирической функции распределения принадлежат отрезку $[0;1]$;

2) если x_1 – наименьшее, а x_n – наибольшее наблюдаемое значение, то $F^*(x)=0$ при значении $x \le x_1$ и $F^*(x)=1$ при значении $x > x_n$.

В отличие от интегральной функции распределения $F(x)$ эмпирическая функция распределения $F^*(x)$ является неслучайной – для разных выборок она получается различной.

Основное значение эмпирической функции распределения состоит в том, что она используется в качестве оценки интегральной функции распределения.

Для наглядности вариационные ряды представляют графиками и диаграммами. Наиболее распространёнными графиками являются полигон, гистограмма и кумулята. Полигон применяется как для дискретных, так и для интервальных спастических рядов, а гистограмма – только для изображения интервальных рядов.

Определение 1.1.3 *Полигоном частот* называют ломаную линию, отрезки которой соединяют точки $(x_1;m_1)$, $(x_2;m_2)$, …, $(x_k;m_k)$, где x_i – варианты выборки, а m_i – соответствующие им частоты.

Определение 1.1.4 *Полигоном относительных частот* или *полигоном частостей* называют ломаную линию, отрезки которой соединяют точки $(x_1;\omega_1)$, $(x_2;\omega_2)$, …, $(x_k;\omega_k)$, где x_i – варианты выборки, а ω_i – соответствующие им относительные частоты $\omega_i = m_i/n$ или частости.

При непрерывном распределении случайной величины разбиваем весь интервал, в котором заключены все наблюдаемые значения случайной величины, на интервалы длиной h и определяем m_i – сумму частот вариант, которые попали в i-й интервал.

Определение 1.1.5 *Гистограммой частот* называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы длиной h, а высоты равны отношению m_i/h.

Площадь частичного i-го прямоугольника равна $h \cdot m_i/h = m_i$ – сумме частот вариант, попавших в i-й интервал. Площадь гистограммы частот равна сумме всех частот, то есть объёму выборки n.

Определение 1.1.6 *Гистограммой относительных частот или частостей* называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиной h, а высоты равны отношению $\omega_i/h = m_i/nh$.

Площадь частичного i-го прямоугольника равна $h \cdot \omega_i/h = \omega_i$ – относительной частоте вариант, попавших в i-й интервал. Площадь гистограммы частостей равна сумме всех относительных частот, то есть единице. Если на гистограмме частостей соединить середины верхних сторон элементарных прямоугольников, то полученная замкнутая ломаная линия образует полигон распределения относительных частот. На рисунках 1.1.2 и 1.1.3 изображены полигоны и гистограммы частостей некоторых случайных величин ξ и η соответственно.

Из принципа построения полигона и гистограммы относительных частот следует, что площадь под полигоном и гистограммой частостей равна единице. В теории вероятностей гистограмме и полигону относительных частот соответствует график плотности распределения вероятностей.

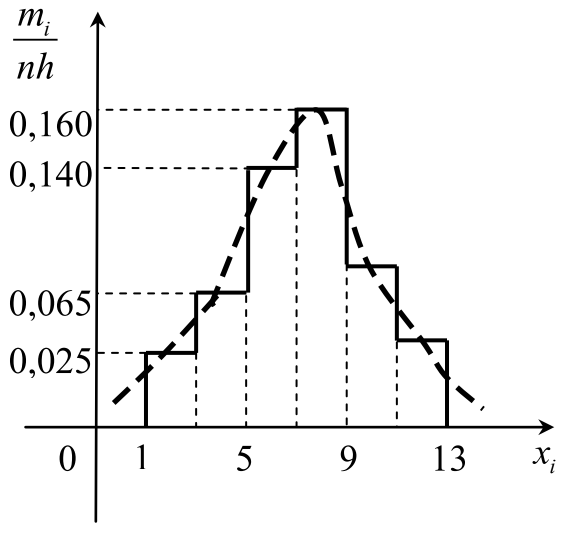

Рисунок 1.1.2 – Полигон и гистограмма частостей СВ ξ

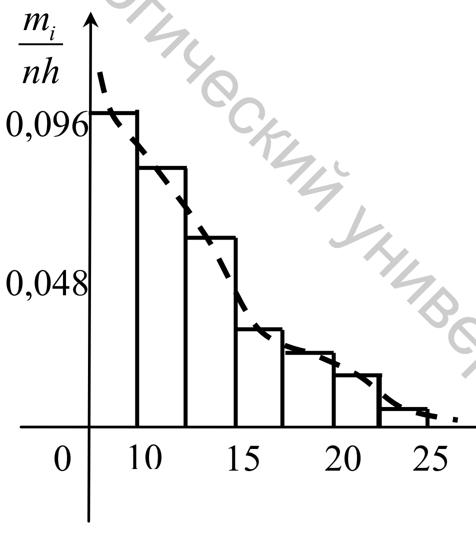

Рисунок 1.1.3 – Полигон и гистограмма частостей СВ η

Гистограмму и полигон частот интервального статистического ряда удобно использовать для визуального подбора модели распределения. Так, гистограмма, изображённая на рисунке 1.1.2, напоминает кривую Гаусса, поэтому можно выдвинуть гипотезу о нормальном распределении случайной величины ξ. Вид гистограммы на рисунке 1.1.3 по форме напоминает экспоненциальную кривую. Поэтому можно выдвинуть гипотезу о показательном распределении случайной величины η.

Для построения кумуляты на оси абсцисс откладывают наблюдаемые значения случайной величины, а на оси ординат – накопленные относительные частоты, то есть накопленные частости.

Определение 1.1.7 *Накопленной частостью* в точке x называется суммарная частость членов статистического ряда, значения которых меньше x.

Значения накопленных частостей являются значениями эмпирической функции распределения $F^*(x)$. В теории вероятностей кумуляте соответствует график функции распределения $F(x) = P(\xi < x)$. На рисунке 1.1.4 изображена кумулята интервального ряда некоторой случайной величины ξ сплошной линией. При построении используются значения эмпирической функции распределения $F^*(x)$.

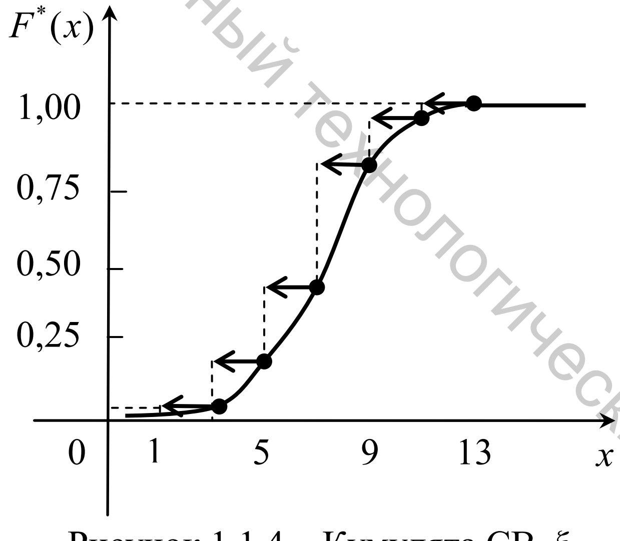

Рисунок 1.1.4 – Кумулята СВ ξ

1.2 Примеры решения типовых задач

1.2.1 На потоке производится замер времени (в минутах) на решение поставленной задачи. Производится выборка среди студентов, которые решали заданную задачу. Время на решение задач представлено в виде таблицы:

t_i – время на решение задачи	11	15	13	15	13	15	11	15	13	15

Найти распределение частот и относительных частот. Построить полигон частот и полигон относительных частот. По данному распределению выборки найти эмпирическую функцию распределения и построить её график.

Решение. В результате опыта случайная величина принимает три значения 11, 13 и 15 минут. Два студента затратили 11 минут на решение поставленной задачи, три – 13 минут, а 5 студентов – 15 минут, то есть случайная величина T – время на решение задачи студентами, равно 11, 13 и 15, а частоты этих событий – 2, 3 и 5, соответственно. Следовательно, статистический ряд распределения частот имеет вид

t_i	11	13	15
m_i	2	3	5

Найдём объём выборки: $n = 2 + 3 + 5 = 10$. Определим относительные частоты: $\omega_1 = m_1/n = 2/10 = 0,2$, $\omega_2 = 3/10 = 0,3$, $\omega_2 = 5/10 = 0,5$. Записываем распределение относительных частот:

t_i	11	13	15
m_i/n	0,2	0,3	0,5

Построим полигон частот (рис. 1.2.1) и полигон относительных частот (рис. 1.2.2).

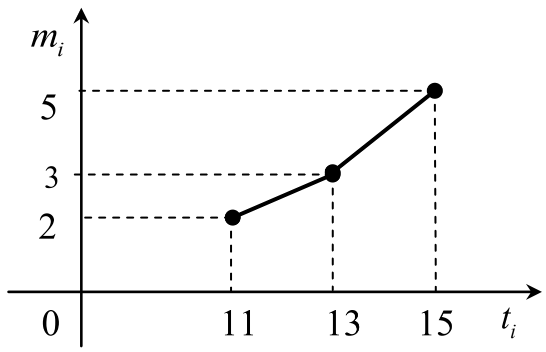

Рисунок 1.2.1 – Полигон частот случайной величины T

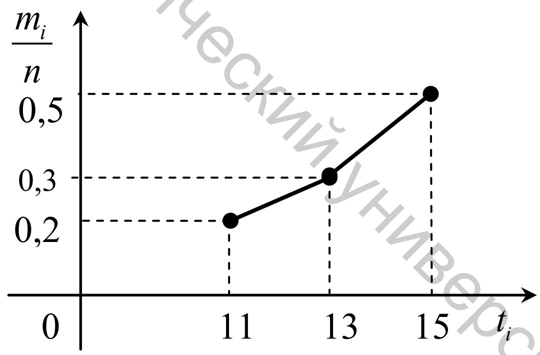

Рисунок 1.2.2 – Полигон относительных частот СВ T

Найдём эмпирическую функцию распределения по заданному распределению выборки. Наименьшая варианта равна 11, следовательно,

$$F^*(t) = 0 \text{ при } t \le 11.$$

Значение $T < 13$, а именно $t_1 = 11$, наблюдалось два раза, следовательно,

$$F^*(t) = 2/10 = 0,2 \text{ при } 11 < t \le 13.$$

Значение $T < 15$, а именно $t_1 = 11$ и $t_2 = 13$, наблюдалось $2+3=5$ раз, следовательно,

$$F^*(t) = 5/10 = 0,5 \text{ при } 13 < t \le 15.$$

Так как $t = 15$ – наибольшая варианта, то

$$F^*(t) = 1 \text{ при } t > 15.$$

Запишем искомую эмпирическую функцию распределения СВ T:

$$F^*(t) = \begin{cases} 0, & t \le 11, \\ 0,2, & 11 < t \le 13, \\ 0,5, & 13 < t \le 15, \\ 1, & t > 15. \end{cases}$$

График эмпирической функции распределения заданной случайной величины T изображён на рисунке 1.2.3.

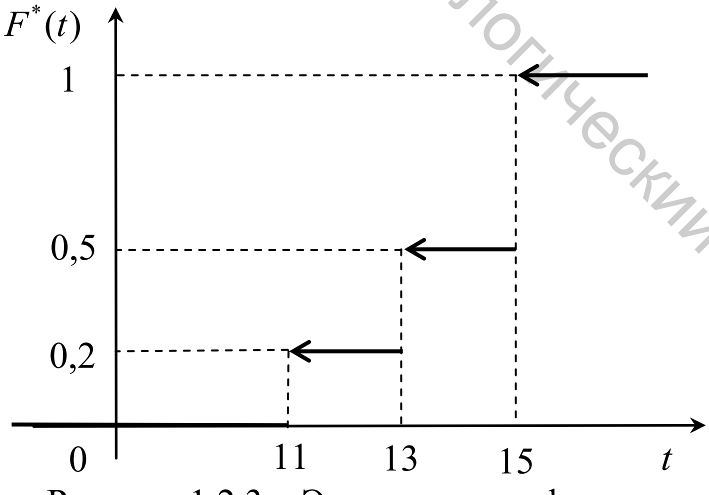

Рисунок 1.2.3 – Эмпирическая функция распределения случайной величины T

1.2.2 Пусть случайная величина ξ – температура воздуха в градусах Цельсия, измеряемая в регионе в течение суток. Из генеральной совокупности извлекается выборка объёмом $n = 200$. Для заданной выборки построить сгруппированный статистический ряд, разбивая размах варьирования на шесть интервалов. Найти эмпирическую функцию распределения и построить её график. Построить гистограмму, полигон распределения частостей или относительных частот, а так же кумуляту полученного статистического ряда. По графику гистограммы сделать прогноз о распределении случайной величины ξ. Значения выборки температуры приведены в виде таблицы:

7,1	5,0	9,6	1,2	9,9	4,0	4,5	12,8	8,1	4,8	5,2	8,3	7.6	8,2
8,2	8,8	5,4	5,4	11,6	5,6	8,8	7,8	5,8	4,8	8,1	9,1	6,1	6,0
3,0	9,4	7,3	8,3	5,1	9,8	6,8	5,2	7,0	5,0	9,5	6,9	11,0	8,9
9,0	9,5	5,4	1,6	10,0	2,8	7,8	8,8	10,2	8,4	3,4	1,9	9,0	4,6
1,7	12,0	9,7	7,4	8,9	5,7	12,6	7,9	8,0	8,1	5,3	6,3	7,1	10,0
5,0	6,0	6,9	9,7	6,4	7,4	5,4	8,9	2,9	10,9	7,7	12,7	5,1	8,0
11,8	7,0	3,5	2,5	11,4	10,4	8,8	10,7	11,2	4,0	4,1	4,3	4,2	6,3
5,2	4,7	6,0	5,6	8,2	3,7	3,6	7,5	5,3	7,3	6,5	8,4	7,9	12,2
7,4	7,2	8,7	4,2	6,6	7,3	13,0	12,4	8,7	6,0	8,5	5,1	6,5	9,4
5,0	4,8	5,2	3,7	10,0	6,0	8,6	8,6	6,9	7,6	6,7	8,5	6,9	6,0
6,8	8,8	7,7	9,8	7,8	7,8	7,2	4,0	10,4	5,4	8,6	5,2	3,9	7,1
9,2	4,9	8,1	2,6	6,0	6,0	10,6	6,9	7,5	10,5	6,1	7,5	10,9	5,4
6,1	6,2	5,9	1,3	3,5	8,0	10,8	7,9	3,8	8,3	10,3	13,0	8,5	3,1
7,5	7,5	5,7	6,7	5,5	10,2	8,0	10,5	5,3	7,4	12,9	10,3	5,1	7,6
3,3	3,2	7,6	1,0										

Решение. Составим вариационный ряд, расположив результаты эксперимента в порядке возрастания.

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

Минимальный элемент выборки равен 1,0, а максимальный элемент равен 13,0, число интервалов разбиения $k = 6$. Находим длину частичного интервала:

$$h = \frac{x_{\max} - x_{\min}}{k} = \frac{13-1}{6} = 2 .$$

Заданный отрезок $[1;13]$ разбиваем на шесть равных интервалов, длина которых равна двум градусам Цельсия. В каждом из интервалов разбиения вычисляем число вариант, которые попали в данный интервал, то есть находим частоты m_i. Составим интервальный статистический ряд, при этом укажем середины интервалов, которые необходимы для построения гистограммы, а также накопленные относительные частоты для построения эмпирической функции распределения и кумуляты.

Номер интервала i	Частичные интервалы	Середины интервалов	Частоты m_i	Относительные частоты m_i/n	Накопленные частости $\sum_{i=1}^{n} m_i/n$
1	$[1;3)$	2	10	0,05	0,05
2	$[3;5)$	4	26	0,13	0,18
3	$[5;7)$	6	56	0,28	0,46
4	$[7;9)$	8	64	0,32	0,78
5	$[9;11)$	10	30	0,15	0,93
6	$[11;13]$	12	14	0,07	1,00

Составим эмпирическую функцию распределения. По определению, эмпирическая функция распределения равна нулю, если $x \le 1$.

Если случайная величина принимает значение из интервала $(1;3]$, то $F^*(x) = 0,05$. При расчётах используем формулу (1.1.1).

Если же случайная величина попадает в интервал $(3;5]$, то эмпирическая функция распределения примет значение $F^*(x) = 0,05 + 0,13 = 0,18$.

В случае попадания случайной величины в интервал $(5;7]$ значение эмпирической функции распределения равно $F^*(x) = 0,05 + 0,13 + 0,28 = 0,46$.

Для значений случайной величины, которые лежат в интервале $(7;9]$, эмпирическая функция распределения принимает числовое значение, которое равно $F^*(x) = 0,05 + 0,13 + 0,28 + 0,32 = 0,78$.

В интервале $(9;11]$ значение эмпирической функции распределения равно $F^*(x) = 0,05 + 0,13 + 0,28 + 0,32 + 0,15 = 0,93$.

На интервале $(11;13]$ эмпирическая функция принимает значение, которое равно значению $F^*(x) = 0,05 + 0,13 + 0,28 + 0,32 + 0,15 + 0,07 = 1$.

По определению, эмпирическая функция распределения равна единице, если случайная величина принимает значение, большее 13.

Таким образом, значения эмпирической функции распределения записаны в последнем столбце интервального статистического ряда. Запишем эмпирическую функцию распределения в аналитическом виде:

$$F^*(x) = \begin{cases} 0, & \text{если} \quad x \le 1; \\ 0,05, & \text{если} \ 1 < x \le 3; \\ 0,18, & \text{если} \ 3 < x \le 5; \\ 0,46, & \text{если} \ 5 < x \le 7; \\ 0,78, & \text{если} \ 7 < x \le 9; \\ 0,93, & \text{если} \ 9 < x \le 11; \\ 1, & \text{если} \ 11 < x \le 13; \\ 1, & \text{если} \quad x > 13. \end{cases}$$

График заданной эмпирической функции распределения случайной величины ξ – температура воздуха в градусах Цельсия, измеряемая в регионе в течение суток, изображён на рисунке 1.1.1. На графике эмпирической функции распределения по оси абсцисс откладываются значения случайной величины, а по оси ординат отмечаются накопленные относительные частоты или частости.

При построении гистограммы частостей или относительных частот на оси абсцисс откладываем частичные интервалы наблюдаемых значений заданной случайной величины ξ. На каждом из полученных интервалов строим прямоугольник с высотой m_i/nh, где h является длиной указанного интервала. С геометрической точки зрения, площадь построенного прямоугольника будет равна частости заданного частичного интервала. Гистограмма частостей заданной случайной величины ξ – температура воздуха в регионе в течение суток, изображена на рисунке 1.1.2 в виде прямоугольников.

Если на гистограмме частостей соединить середины верхних сторон элементарных прямоугольников, то полученная линия образует полигон частостей. На рисунке 1.1.2 полигон частостей заданной случайной величины изображён штриховой линией.

В случае построения кумуляты, на оси абсцисс откладываем частичные интервалы наблюдаемых значений заданной случайной величины ξ, а по оси ординат – накопленные частости. Изображаем точки, абсциссы которых совпадают с правым концом частичного интервала, не входящего в него, а ординаты равны накопленным относительным частотам. Соединяя полученные точки плавной линией, строим график кумуляты. График кумуляты заданной случайной величины ξ – температура воздуха в данном регионе, изображён на рисунке 1.1.4 сплошной линией. На этом же рисунке изображён график эмпирической функции распределения (линии со стрелками).

Так как гистограмма заданной случайной величины ξ – температуры воздуха в регионе, которая изображена на рисунке 1.1.2, напоминает кривую нормального распределения, то можно выдвинуть гипотезу о нормальном распределении заданной случайной величины.

1.3 Задания для решения на практическом занятии

1.3.1 Найти эмпирическую функцию распределения по заданному распределению выборки:

x_i	2	4	7	10	12
m_i	5	26	44	18	7

1.3.2 Найти эмпирическую функцию распределения по заданному распределению выборки:

x_i	5	10	15	20	25	30
m_i	72	46	38	24	12	8

1.3.3 Результаты исследования непрерывной случайной величины ξ представлены в виде интервального статистического ряда:

Частичный интервал	$[1;4)$	$[4;7)$	$[7;10)$	$[10;13)$	$[13;16]$
Сумма частот вариант интервала	12	18	30	24	16

Найти эмпирическую функцию распределения и кумуляту. Построить их графики.

1.3.4 Результаты исследования непрерывной случайной величины η представлены в виде интервального статистического ряда:

Частичный интервал	$[2;3)$	$[3;4)$	$[4;5)$	$[5;6)$	$[6;7]$
Сумма частот вариант интервала	15	14	10	6	5

Найти эмпирическую функцию распределения и кумуляту. Построить их графики.

1.3.5 Построить полигон частот и полигон относительных частот по данному распределению выборки случайной величины ξ:

x_i	1	2	4	5	7
m_i	25	13	12	6	4

По виду графиков статистического ряда предположить закон распределения случайной величины ξ.

1.3.6 Построить полигон частот и полигон относительных частот по данному распределению выборки случайной величины η:

x_i	23	25	27	28	33	35
m_i	18	22	17	23	19	21

По виду графиков статистического ряда предположить закон распределения случайной величины η.

1.3.7 Построить гистограмму частот по данному распределению выборки случайной величины ξ (объём выборки равен $n = 100$):

Частичный интервал	$[1;4)$	$[4;7)$	$[7;10)$	$[10;13)$	$[13;16]$
Сумма частот вариант интервала	12	18	30	24	16

По виду гистограммы частот статистического распределения выдвинуть гипотезу о законе распределения случайной величины ξ.

1.3.8 Построить гистограмму относительных частот по данному распределению выборки случайной величины η (объём выборки равен $n = 100$):

Частичный интервал	$[1;4)$	$[4;7)$	$[7;10)$	$[10;13)$	$[13;16]$
Сумма частот вариант интервала	12	18	30	24	16

По виду гистограммы относительных частот статистического распределения выдвинуть гипотезу о законе распределения случайной величины η.

1.3.9 Из генеральной совокупности случайной величины ξ – среднего роста человека, произведена выборка, в которой оказалось только пять значений

160, 165, 170, 175 и 180 сантиметров. Данные эксперимента представлены в виде таблицы.

160	180	165	175	165	165	170	165	175	165
170	170	160	170	180	170	175	175	165	170
165	175	170	175	170	175	165	170	170	175
175	170	175	165	175	160	180	175	175	175
170	165	180	170	170	170	175	180	160	170

Найти объём выборки и записать статистический ряд в виде распределения частот и относительных частот. Построить эмпирическую функцию распределения и кумуляту по статистическому распределению случайной величины. Построить полигон частот и полигон относительных частот статистического распределения. На основании графиков статистического ряда выдвинуть гипотезу о распределении заданной случайной величины ξ.

1.3.10 В результате опыта произведено изучение случайной величины η. Из генеральной совокупности извлечена выборка, которая представлена в виде таблицы.

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

Для заданной выборки построить сгруппированный статистический ряд, разбивая размах варьирования на шесть интервалов. Найти эмпирическую функцию распределения и построить её график. Построить гистограмму, полигон частот, полигон распределения частостей или относительных частот, а также кумуляту полученного статистического ряда.

1.4 Задания для контролируемой самостоятельной работы

1.4.1 Из генеральной совокупности случайной величины ξ извлечена выборка, которая представлена в виде таблицы. Для заданной выборки построить сгруппированный статистический ряд, разбивая размах варьирования на шесть интервалов. Найти эмпирическую функцию распределения и построить её график. Построить гистограмму, полигон частот, полигон распределения частостей или относительных частот и кумуляту полученного статистического ряда.

1.4.1.1

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

1.4.1.2

364	330	324	341	356	347	357	361	328	344
325	358	340	372	329	366	337	370	354	360
343	348	325	350	362	336	363	345	359	324
350	333	352	327	335	363	348	356	338	372
332	368	369	334	348	347	326	331	341	339

1.4.1.3

12	1	72	40	32	58	6	33	64	11
36	50	14	3	16	44	23	54	8	68
0	26	38	15	42	5	46	21	40	19
48	13	2	30	56	17	31	7	20	52
24	10	28	66	4	34	60	45	35	9

1.4.1.4

400	140	215	180	225	110	200	320	290	150
245	240	325	230	270	220	280	210	300	295
175	255	235	130	350	170	215	160	245	200
200	310	260	265	220	340	230	285	205	265
250	190	210	330	270	275	100	240	325	375

1.4.1.5

38,0	36,5	36,0	36,7	37,2	36,5	37,6	36,1	36,0	37,5
36,8	37,7	37,4	38,2	36,2	36,8	36,6	37,1	36,4	36,5
37,2	38,1	37,0	36,3	36,8	36,3	37,3	36,7	37,8	38,2
36,4	36,9	37,8	37,5	36,4	38,3	37,7	36,1	36,6	36,3
37,6	37,3	36,6	37,1	37,9	36,2	36,3	38,4	36,8	37,9

1.4.1.6

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

1.4.1.7

98	40	135	92	68	88	86	74	98	140
43	95	64	49	106	75	55	82	115	80
85	118	46	63	90	20	72	65	76	78
60	96	30	66	73	52	82	84	35	120
100	62	94	103	80	70	130	109	82	58

1.4.1.8

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

1.4.1.9

105	260	345	340	95	368	275	35	455	60
690	124	55	80	245	25	169	460	187	570
115	10	133	205	20	160	70	178	40	196
5	391	230	0	151	50	520	65	290	45
225	110	15	142	414	320	30	575	65	310

1.4.1.10

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

1.4.1.11

225	238	235	251	248	253	241	234	256	249
250	227	223	236	240	236	221	243	233	220
232	234	228	239	221	231	254	225	242	231
226	244	247	222	230	247	235	255	223	246
238	224	252	229	237	220	244	226	245	232

1.4.1.12

36	8	0	12	22	44	4	5	6	7
8	17	54	21	30	23	15	5	34	9
27	10	20	1	41	3	50	16	17	35
9	19	29	38	13	15	24	33	26	18
18	1	11	32	2	14	3	25	8	45

1.4.1.13

175	460	490	470	505	400	310	435	150	397
455	265	465	230	420	100	420	520	389	450
400	475	341	415	357	365	205	381	445	320
325	405	125	349	245	425	373	435	535	470
250	333	410	280	450	295	430	130	440	190

1.4.1.14

65	79	51	60	90	52	78	64	53	92
72	66	88	62	63	83	55	88	69	54
86	73	67	81	76	91	84	72	56	85
57	87	80	75	54	71	93	68	79	70
86	60	74	89	82	77	55	85	86	58

1.4.1.15

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

1.4.1.16

40	130	94	48	96	115	56	84	125	88
100	92	138	76	160	152	98	128	86	38
90	10	104	142	108	80	82	115	60	98
134	72	44	106	78	52	112	70	158	120
70	102	74	64	148	110	156	190	118	68

1.4.1.17

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

1.4.1.18

200	20	150	65	160	125	80	170	280	45
25	210	15	250	70	45	190	135	40	180
100	55	110	10	120	230	130	35	10	145
195	25	175	220	5	75	30	245	300	240
50	105	60	115	165	0	260	85	95	90

1.4.1.19

7,0	2,0	7,6	15,5	9,9	11,0	14,0	10,2	9,4	11,8
10,0	13,5	12,4	10,6	5,0	8,5	11,2	17,0	11,6	15,0
5,5	10,2	18,0	7,9	10,8	1,0	8,8	11,4	10,0	9,7
12,0	7,3	4,0	12,6	19,0	12,8	17,5	9,1	4,5	16,5
14,5	12,2	10,4	13,0	8,2	16,0	10,0	6,0	10,4	3,0

1.4.1.20

196	161	151	184	193	153	166	155	204	177
157	197	185	191	183	195	158	203	167	178
160	186	198	163	172	182	202	180	179	168
169	189	162	199	164	201	173	187	154	175
187	159	170	171	200	165	181	174	191	205

1.4.1.21

10	55	35	11	41	47	39	3	66	5
22	1	12	33	15	2	50	29	30	20
1	13	8	43	26	16	28	7	19	32
14	23	0	25	9	27	17	53	5	11
7	44	24	6	60	21	4	18	35	6

1.4.1.22

75	128	154	134	104	152	143	154	166	124
125	172	120	150	158	199	155	215	149	168
170	225	174	100	137	160	179	164	112	164
116	152	131	187	155	200	162	195	162	85
150	191	95	156	183	140	108	146	175	148

1.4.1.23

51	77	56	75	15	42	23	58	48	31
29	53	41	14	47	16	65	22	69	60
45	28	13	57	32	24	17	46	21	72
70	12	27	40	25	39	37	18	33	20
11	43	55	30	59	61	44	35	19	50

1.4.1.24

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

1.4.1.25

90	54	50	35	74	50	51	78	65	38
53	59	85	63	58	48	34	59	52	52
32	70	62	56	30	80	58	51	42	66
61	55	54	72	57	55	60	36	58	80
40	69	59	45	64	53	76	57	75	56

1.4.1.26

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

1.4.1.27

120	10	240	175	180	30	90	55	330	110
50	130	5	140	15	190	170	250	105	200
60	260	55	0	10	20	210	100	270	40
25	65	220	75	150	160	35	125	45	290
360	210	70	60	80	85	280	30	230	300

1.4.1.28

25	29	15	5	16	23	14	7	22	9
10	14	26	18	12	1	19	17	14	17
20	15	11	13	27	16	24	2	24	8
18	21	15	0	19	12	28	11	16	10
4	9	16	22	23	6	13	17	3	30

1.4.1.29

810	840	930	819	882	925	852	980	958	898
960	965	900	934	822	942	890	954	833	945
870	813	970	866	938	886	950	831	915	860
905	990	844	975	910	825	904	894	990	836
839	874	816	878	848	946	828	985	856	920

1.4.1.30

5	32	18	24	12	26	7	28	9	3
40	1	34	11	1	13	38	23	2	30
16	17	10	19	48	0	14	39	46	10
7	9	31	5	20	36	22	15	29	6
8	42	6	27	4	21	36	3	8	44

2 ТОЧЕЧНЫЕ И ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

Содержание: классификация точечных оценок, метод моментов, метод-максимального правдобододобия, интервальные оценки параметров распределения, доверительные интервалы для нормального распределения.

2.1 Теоретический материал по теме практического занятия

Математическая статистика позволяет получать обоснованные выводы о параметрах или виде распределения случайных величин по совокупности наблюдений над ними. Распределение случайной величины ξ характеризуется рядом параметров. К ним относятся и числовые характеристики: математическое ожидание, дисперсия, мода, медиана и другие.Эти параметры называются *параметрами генеральной совокупности*. Они могут быть найдены из законов распределения случайных величин. Используя выборку наблюдений случайной величины, можно определить приближённые значения каждого из параметров, называемые в математической статистике *числовыми оценками параметров* или просто *оценками*. Оценки параметров распределения являются значениями некоторых функций элементов выборки – *статистик*. Рассмотрим статистики, которые используютсяпри получении оценок числовых характеристик распределения.

Предположим из генеральной совокупности с функцией распределения $F(x,\Theta)$, где Θ – неизвестный параметр, извлечена выборка $x_1, x_2, ..., x_n$. Задача оценивания неизвестного параметра Θ состоит в построении приближённых формул

$$\Theta \approx u(x_1, x_2, ..., x_n). \tag{2.1.1}$$

Функцию (2.1.1) называют *выборочной* или *статистической функцией*, а её значение в приближённом равенстве – *оценкой* и обозначают $\hat{\Theta} = u(x_1, x_2, ..., x_n)$. Любая выборка является конечной и случайной, следовательно, статистические функции $\hat{\Theta}$ также являются случайными. Таким образом, будем рассматривать оценку $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ неизвестного параметра Θ как случайную величину, а её значение, вычисленное по заданной выборке объёмом n, как одну реализацию случайной величины, то есть как одно из множества её возможных значений.

В математической статистике рассматриваются точечные и интервальные оценки параметров распределения.

Точечная оценка параметра Θ определяется одним числом $\hat{\Theta} = u(x_1, x_2, ..., x_n)$. *Интервальной оценкой* называют оценку, которая определяется двумя числами $\hat{\Theta}_1$ и $\hat{\Theta}_2$ – концами интервала, накрывающего оцениваемый параметр Θ.

Рассмотрим классификацию точечных оценок. Чтобы точечная оценка неизвестного параметра $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ была «хорошей» прежде всего с точки зрения точности и надёжности оценок, желательно, чтобы найденные на основании статистических функций $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ оценки неизвестных параметров были тесно сгруппированы около значений оцениваемых параметров распределения, то есть чтобы рассеивание случайной величины $\hat{\Theta}$ около параметра Θ было по возможности минимальным.

Определение 2.1.1 Оценка $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ называется *состоятельной*, если при увеличении числа испытаний оценка сходится по вероятности к оцениваемому параметру, то есть если $\lim\limits_{n\to\infty} P\left(\left|\Theta - \hat{\Theta}\right| < \varepsilon\right) = 1$.

Определение 2.1.2 Оценка $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ называется *несмещённой* (оценкой без систематической ошибки), если её математическое ожидание равно оцениваемому параметру, то есть $M\left(\hat{\Theta}\right) = \Theta$.

Если последнее равенство не выполняется, то оценка называется *смещённой* (содержащей систематическую ошибку). Наряду с несмещёнными оценками $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ применяются асимптотически несмещённые оценки, для которых $M\left(\hat{\Theta}\right) \to \Theta$ при увеличении объёма выборки.

Дисперсия любой несмещённой оценки одного параметра Θ удовлетворяет неравенству Рао-Крамера

$$D\left(\hat{\Theta}\right) \geq -\frac{1}{N \cdot M\left(\dfrac{\partial^2 \ln f(x,\Theta)}{\partial \Theta^2}\right)}, \tag{2.1.2}$$

где $f(x,\Theta)$ – плотность распределения вероятностей случайной величины; N – число произведённых испытаний. В случае дискретной случайной величины плотность распределения вероятностей $f(x,\Theta)$ заменяется функцией распределения вероятностей.

Определение 2.1.3 Оценка $\hat{\Theta} = u(x_1, x_2, ..., x_n)$, для которой в неравенстве Рао-Крамера (2.1.2) достигается знак равенства, называется *эффективной*.

Определение 2.1.4 Оценка $\hat{\Theta} = u(x_1, x_2, ..., x_n)$ называется *достаточной*, если она использует всю информацию относительно оцениваемого параметра, содержащуюся в выборке.

Достаточные оценки построены таким образом, что никакие другие оценки не могут дать какой-либо дополнительной информации об оцениваемых параметрах.

Рассмотрим точечные оценки неизвестных параметров распределения.

Несмещённой оценкой генеральной средней (математического ожидания) является *выборочная средняя*

$$\overline{x}_B = \frac{\sum_{i=1}^{k} m_i x_i}{n}, \tag{2.1.3}$$

где x_i – варианта выборки; m_i – частота варианты x_i; $\sum_{i=1}^{k} m_i = n$.

Смещённой оценкой генеральной дисперсии служит *выборочная дисперсия*

$$D_B = \frac{\sum_{i=1}^{k} m_i \left(x_i - \overline{x}_B\right)^2}{n} = \overline{x_B^2} - \overline{x}_B{}^2. \tag{2.1.4}$$

Выборочная дисперсия является смещённой оценкой генеральной дисперсии $D_Г$, так как $M\left(D_B\right)=\frac{n-1}{n}\cdot D_Г \neq D_Г$.

Положительное значение квадратного корня из выборочной дисперсии называется *выборочным средним квадратическим отклонением* $\sigma_B=\sqrt{D_B}$.

Несмещённой оценкой генеральной дисперсии служит «*исправленная дисперсия*», которая определяется по формуле

$$s^2=\frac{n}{n-1}\cdot D_B=\frac{1}{n-1}\cdot\sum_{i=1}^{k}m_i\left(x_i-\overline{x}_B\right)^2. \qquad (2.1.5)$$

Несмещённой оценкой среднего квадратического отклонения генеральной совокупности является «исправленное» среднее квадратическое отклонение

$$s=\sqrt{s^2}=\sqrt{\frac{n}{n-1}\cdot D_B}=\sqrt{\frac{1}{n-1}\cdot\sum_{i=1}^{k}m_i\left(x_i-\overline{x}_B\right)^2}. \qquad (2.1.6)$$

Если рассматривается интервальный статистический ряд, то в качестве вариант берутся середины частичных интервалов.

Рассмотрим методы получения оценок.

1. *Метод моментов.*

Пусть ξ – непрерывная случайная величина с плотность распределения вероятностей $f(x,\Theta)$, зависящая от одномерного неизвестного параметра Θ. Тогда математическое ожидание $M(\xi)$ является функцией Θ:

$$M(\xi)=\int_{-\infty}^{+\infty}x\,f(x,\Theta)\,dx=\mu_1(\Theta).$$

Выборочное среднее $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ принимает значение, близкое к $M(\xi)$. Это позволяет записать уравнение для определения неизвестного параметра Θ:

$$\mu_1(\Theta)=\overline{X}.$$

Метод моментов аналогичным образом применяется к дискретным случайным величинам.

2. *Метод максимального правдоподобия.*

Пусть из генеральной совокупности произведена выборка заданного объёма. Составляем функцию правдоподобия $L=L(x_1,x_2,...,x_n,\Theta_1,\Theta_2,...,\Theta_m)$. В случае дискретной случайной величины функция правдоподобия представляет собой произведение вероятностей, а вслучае непрерывной случайной величины – произведение плотностей вероятности. В качестве оценки неизвестных параметров выбирается такая функция $\hat{\Theta}$, которая максимизирует функцию правдоподобия. Из системы

$$\frac{\partial L}{\partial \Theta_i} = 0, \quad i = \overline{1,m},$$

выбираем то решение, которое обращает функцию правдоподобия в максимум. Поскольку экстремум функции L и $\ln L$ достигается при одних и тех же значениях $\hat{\Theta}$, то для упрощения расчётов используется логарифмическая функция правдоподобия. В этом случае оценки максимального правдоподобия находятся из системы уравнений $\frac{\partial \ln L}{\partial \Theta_i} = 0$.

Рассмотрим интервальные оцценки параметров распределения.

Точечная оценка неизвестного параметра Θ, найденная по выборке объёма n из генеральной совокупности с функцией распределения $F(x,\Theta)$, не позволяет непосредственно ответить на вопрос, какую ошибку мы совершаем, принимая вместо точного значения неизвестного параметра Θ его приближённое значение или оценку $\hat{\Theta} = u(x_1, x_2, ..., x_n)$.

В связи с этим более выгодно пользоваться интервальной оценкой, основанной на определении некоторого интервала, внутри которого с определённой вероятностью находится неизвестное значение параметра Θ. Предположим, что по результатам выборки объёма n найдена выборочная характеристика $\hat{\Theta} = u(x_1, x_2, ..., x_n)$, которая является точечной оценкой неизвестного параметра Θ. Чем меньше разность $\left|\Theta - \hat{\Theta}\right|$, тем точнее будет оценка. Таким образом, положительное число ε характеризует точность оценки $\left|\Theta - \hat{\Theta}\right| < \varepsilon$. При этом точность ε зависит от объёма выборки. Поэтому можно только говорить о вероятности $1-\alpha$, с которой это неравенство выполняется.

Определение 2.1.5 *Доверительной вероятностью* оценки называют вероятность $P = 1-\alpha$ выполнения неравенства $\left|\Theta - \hat{\Theta}\right| < \varepsilon$.

Доверительная вероятность точечной оценки показывает, что при извлечении выборки одинакового объёма из одной и той же генеральной совокупности с функцией распределения $F(x,\Theta)$ в $(1-\alpha)$ 100 % случаях параметр Θ будет накрываться данным интервалом.

Определение 2.1.6 *Доверительным интервалом* называется интервал $\left(\hat{\Theta}-\varepsilon;\Theta+\varepsilon\right)$, накрывающий неизвестный параметр Θ с заданной доверительной вероятностью $P=1-\alpha$, с уровнем значимости α.

Рассмотрим случайную величину ξ, которая подчинена нормальному закону распределения.

Интервальной оценкой математического ожидания a случайной величины ξ, имеющей нормальное распределение *при известном среднем квадратическом отклонении* σ и выборочной средней $\overline{x}_B$, является доверительный интервал

$$\overline{x}_B-u_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}<a<\overline{x}_B+u_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}. \qquad (2.1.7)$$

Квантили нормального распределения $u_{\alpha/2}$ находят по таблицам функции Лапласа (приложение, табл. А.3) из условия $2\cdot\Phi(u_{\alpha/2})=1-\alpha$. Доверительный интервал покрывает математическое ожидание a с заданной вероятностью или надёжностью $1-\alpha$. Точность оценки математического ожидания (предельная погрешность) равна $\varepsilon=u_{\alpha/2}\sigma\big/\sqrt{n}$.

Интервальной оценкой математического ожидания a случайной величины ξ, имеющей нормальное распределение *при неизвестном среднем квадратическом отклонении* σ и известной выборочной средней $\overline{x}_B$, является доверительный интервал

$$\overline{x}_B-t_{\alpha/2;\,n-1}\cdot\frac{s}{\sqrt{n}}<a<\overline{x}_B+t_{\alpha/2;\,n-1}\cdot\frac{s}{\sqrt{n}}. \qquad (2.1.8)$$

Квантили *t*-распределения $t_{\alpha/2;\,n-1}$ находят по таблицам распределения Стьюдента (приложение, табл. А.5). Доверительный интервал покрывает математическое ожидание a с заданной вероятностью или надёжностью $1-\alpha$. Точность оценки математического ожидания (предельная погрешность) равна $\varepsilon=t_{\alpha/2;\,n-1}\,s\big/\sqrt{n}$.

Интервальной оценкой среднего квадратического отклонения σ случайной величины ξ, имеющей нормальное распределение, является доверительный интервал

$$s\cdot\gamma_1<\sigma<s\cdot\gamma_2, \qquad (2.1.9)$$

где $\gamma_1=\sqrt{\dfrac{n-1}{\chi^2_{\alpha/2;\,n-1}}}$, $\gamma_2=\sqrt{\dfrac{n-1}{\chi^2_{1-\alpha/2;\,n-1}}}$, $s=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{k}m_i\left(x_i-\overline{x}_B\right)^2}$.

Коэффициенты γ_1 и γ_2, соответствующие доверительной вероятности $P = 1 - \alpha$ и числу степеней свободы $\nu = n - 1$, находятся по таблице (приложение, табл. А.7).

Так как $n \to \infty$, распределение χ^2 приближается к нормальному распределению, то при достаточно большом объёме выборки $(n \geq 50)$ доверительный интервал для среднего квадратического отклонения можно найти по формуле

$$P\left(\frac{s}{1 + u_{\alpha/2}/\sqrt{2n}} < \sigma < \frac{s}{1 - u_{\alpha/2}/\sqrt{2n}}\right) = 1 - \alpha, \qquad (2.1.10)$$

где $u_{\alpha/2}$ – квантиль стандартизированного нормального распределения, соответствующий доверительной вероятности $1 - \alpha$.

2.2 Примеры решения типовых задач

2.2.1 Из генеральной совокупности, исследования случайной величины T – температура тела человека, измеряемая в течение определённого промежутка времени, извлечена выборка объёма $n = 100$. Данные представлены в виде статистического ряда

варианта t_i	36,2	36,4	36,6	36,8	37,0
частота m_i	6	18	52	20	4

Найти выборочную среднюю температуру, выборочную и исправленную дисперсии измерения температуры.

Решение. Находим выборочную среднюю по формуле (2.1.3):

$$\overline{t_B} = \frac{\sum_{i=1}^{5} m_i t_i}{n} = \frac{6 \cdot 36{,}2 + 18 \cdot 36{,}4 + 52 \cdot 36{,}6 + 20 \cdot 36{,}8 + 4 \cdot 37{,}0}{100} = 36{,}596.$$

Находим выборочную дисперсию по формуле (2.1.4):

$$D_B = \frac{\sum_{i=1}^{5} m_i \left(t_i - \overline{t_B}\right)^2}{n} = \frac{6 \cdot (36{,}2 - 36{,}596)^2 + 18 \cdot (36{,}4 - 36{,}596)^2}{100} +$$

$$+ \frac{52 \cdot (36{,}6 - 36{,}596)^2 + 20 \cdot (36{,}8 - 36{,}596)^2 + 4 \cdot (37{,}0 - 36{,}596)^2}{100} = 0{,}031184.$$

Находим исправленную дисперсию по формуле (2.1.5):

$$s^2 = \frac{n}{n-1} \cdot D_B = \frac{100}{100-1} \cdot 0,031184 = 0,0315.$$

2.2.2 Из генеральной совокупности исследуемой случайной величины ξ произведена выборка объёмом $n = 100$, которые записаны в виде интервального статистического ряда

Частичный интервал	$[2;4)$	$[4;6)$	$[6;8)$	$[8;10)$	$[10;12]$
Сумма частот вариант интервала	8	20	42	18	12

Найти выборочную среднюю, выборочную дисперсию и выборочное среднее квадратическое отклонение, исправленную дисперсию и исправленное среднее квадратическое отклонение наблюдаемых значений случайной величины ξ.

Решение. В начале находим середины интервалов и выбираем их в качестве вариант наблюдаемых значений случайной величины ξ: $x_1 = 3$, $x_2 = 5$, $x_3 = 7$, $x_4 = 9$, $x_5 = 11$.

Находим выборочную среднюю по формуле (2.1.3):

$$\overline{x}_B = \frac{\sum_{i=1}^{5} m_i x_i}{n} = \frac{8 \cdot 3 + 20 \cdot 5 + 42 \cdot 7 + 18 \cdot 9 + 12 \cdot 11}{100} = 7,12.$$

Находим выборочную дисперсию по формуле (2.1.4):

$$D_B = \frac{\sum_{i=1}^{5} m_i \left(x_i - \overline{x}_B\right)^2}{n} = \frac{8 \cdot (3-7,12)^2 + 20 \cdot (5-7,12)^2}{100} +$$

$$+ \frac{42 \cdot (7-7,12)^2 + 18 \cdot (9-7,12)^2 + 12 \cdot (11-7,12)^2}{100} = 4,7056.$$

Находим исправленную дисперсию по формуле (2.1.5):

$$s^2 = \frac{n}{n-1} \cdot D_B = \frac{100}{100-1} \cdot 4,7056 = 4,7531.$$

Выборочное среднее квадратическое отклонением равно $\sigma_B = \sqrt{D_B} = \sqrt{4,7056} = 2,169$, а «исправленное» среднее квадратическое отклонение равно $s = \sqrt{s^2} = \sqrt{4,7531} = 2,18$.

2.2.3 Случайная величина ξ имеет показательное распределение с плотностью вероятностей $f(x,\Theta) = \Theta \cdot e^{-\Theta x}$. По результатам наблюдаемых значений $x_1, x_2, ..., x_n$ этой величины оценить параметр Θ.

Решение. Запишем функцию правдоподобия $L = \Theta \cdot e^{-\Theta x_1} \cdot \Theta \cdot e^{-\Theta x_2} \times \times \Theta \cdot e^{-\Theta x_n} = \Theta^n \cdot e^{-\Theta \sum\limits_{i=1}^{n} x_i}$. Логарифмируя, имеем $\ln L = n \ln \Theta - \Theta \sum\limits_{i=1}^{n} x_i$. Дифференцируя по параметру Θ, находим $\dfrac{\partial L}{\partial \Theta} = \dfrac{n}{\Theta} - \sum\limits_{i=1}^{n} x_i$. Приравнивая частную производную к нулю, получаем $\dfrac{n}{\Theta} - \sum\limits_{i=1}^{n} x_i = 0$. Следовательно оценка неизвестного параметра имеет вид: $\hat{\Theta} = \dfrac{n}{\sum\limits_{i=1}^{n} n} = \dfrac{1}{\overline{x}}$.

2.2.4 Пусть случайная величина ξ имеет нормальное распределение с известным средним квадратическим отклонением $\sigma = 3$. Найти доверительный интервал для математического ожидания, если среднее выборочное наблюдаемых значений $n = 36$ прямых равноточных измерений равно $\overline{x}_B = 18$, а уровень значимости равен $\alpha = 0,05$.

Решение. Пользуясь таблицей функции Лапласа (приложение, табл. А.3), определим квантиль стандартизированного нормального распределения, учитывая, что доверительная вероятность равна $P = 1 - \alpha = 1 - 0,05 = 0,95$:

$$2 \cdot \Phi(u_{\alpha/2}) = 1 - \alpha,\ 2 \cdot \Phi(u_{\alpha/2}) = 0,95,\ \Phi(u_{\alpha/2}) = 0,475,\ u_{\alpha/2} = 1,96.$$

Для определения доверительного интервала для математического ожидания воспользуемся неравенством (2.1.7):

$$18 - 1,96 \cdot \frac{3}{\sqrt{36}} < a < 18 + 1,96 \cdot \frac{3}{\sqrt{36}} \text{ или } 17,02 < a < 18,98.$$

Следовательно, искомый доверительный интервал имеет вид $(17,02;18,98)$. Раскроем смысл полученного результата: если будет произведено достаточно большое число выборок заданного объёма, то в 95 % случаев из них доверительные интервалы накроют математическое ожидание и только в 5 %

случаев оцениваемое математическое ожидание может выйти за пределы доверительных интервалов.

2.2.5 Найти минимальный объём выборки, на основании которой можно было бы оценить математическое ожидание времени исполнения некоторой операции с ошибкой, не превышающей 15 часов, и надёжностью $P = 1 - \alpha = 0,9$, если предположить, что время выполнения указанной операции ξ является случайной величиной, имеющей нормальное распределение с средним квадратическим отклонением $\sigma = 60$.

Решение. Пользуясь таблицей функции Лапласа (приложение, табл. А.3), определим квантиль стандартизированного нормального распределения:

$$2 \cdot \Phi(u_{\alpha/2}) = 1 - \alpha,\ 2 \cdot \Phi(u_{\alpha/2}) = 0,9,\ \Phi(u_{\alpha/2}) = 0,45,\ u_{\alpha/2} = 1,64.$$

Воспользуемся формулой, связывающей предельную погрешность ε оценки математического ожидания по выборочной средней: $\varepsilon = u_{\alpha/2}\sigma \big/ \sqrt{n}$. Следовательно, $n \geq \dfrac{u_{\alpha/2}^2 \cdot \sigma^2}{\varepsilon^2} = \dfrac{1,64^2 \cdot 60^2}{15^2} = 43,0336$. Таким образом, $n = 44$.

2.2.6 В результате исследования некоторой случайной величины ξ – линейный размер некоторого объекта, подчинённого нормальному закону распределения, произведена выборка объёмом $n = 5$: $x_1 = 2, x_2 = 2,1, x_3 = 2,2, x_4 = 2,1,$ $x_5 = 2$. Оценить с помощью доверительного интервала математическое ожидание линейного размера. Доверительную вероятность $P = 1 - \alpha$ принять равной 0,99. Найти минимальное число измерений, которое необходимо выполнить, чтобы с надёжностью $1 - \alpha = 0,99$ можно было утверждать, что предельная погрешность точечной оценки исследуемой величины не превышала 0,002. Построить доверительный интервал, покрывающий среднее квадратическое отклонение σ с надёжностью $P = 1 - \alpha = 0,99$.

Решение. Найдём точечные оценки неизвестных параметров: математического ожидания a и среднего квадратического отклонения σ.

$$\hat{a} = \overline{x}_B = \frac{2 \cdot 2 + 2 \cdot 2,1 + 1 \cdot 2,2}{5} = 2,08.$$

$$\hat{\sigma} = s = \sqrt{\frac{1}{5-1}\left(2 \cdot (2 - 2,08)^2 + 2 \cdot (2,1 - 2,08)^2 + 1 \cdot (2,2 - 2,08)^2\right)} = 0,08.$$

По таблице распределения Стьюдента (приложение, табл. А.5) по доверительной вероятности $P = 1 - \alpha = 0,99$ и числу степеней свободы $\nu = n - 1 = 4$ находим квантиль *t*-распределения: $t_{\alpha/2;\, n-1} = t_{0,005;\, 4} = 4,604$. Для определения доверительного интервала воспользуемся неравенством (2.1.8).

$$\overline{x}_B - t_{\alpha/2;\,n-1} \cdot \frac{s}{\sqrt{n}} < a < \overline{x}_B + t_{\alpha/2;\,n-1} \cdot \frac{s}{\sqrt{n}}.$$

$$2{,}08 - 4{,}604 \cdot \frac{0{,}08}{\sqrt{5}} < a < 2{,}08 + 4{,}604 \cdot \frac{0{,}08}{\sqrt{5}} \text{ или } 1{,}915 < a < 2{,}245.$$

Следовательно, искомый доверительный интервал имеет вид $(1{,}915; 2{,}245)$. Таким образом, если будет произведено достаточно большое число выборок заданного объёма, то в 99 % случаев из них доверительные интервалы накроют математическое ожидание и только в 1 % случаев оцениваемое математическое ожидание может выйти за пределы доверительных интервалов.

Для расчёта истинного числа измерений, необходимых для определения истинного линейного размера объекта с погрешностью, не превышающей $\varepsilon = 0{,}02$, воспользуемся формулой предельной погрешности: $\varepsilon = t_{\alpha/2;\,n-1}\, s \big/ \sqrt{n}$.

Следовательно, $n \geq \dfrac{t^2_{\alpha/2;\,n-1} \cdot s^2}{\varepsilon^2} = \dfrac{(4{,}604)^2 \cdot (0{,}08)^2}{(0{,}002)^2} \approx 33{,}91$ измерения.

Таким образом, для того чтобы предельная погрешность ε истинного значения линейного размера исследуемого объекта не превышала $\varepsilon = 0{,}002$, следует, помимо пяти проделанных измерений, выполнить ещё $n - 5 = 34 - 5 = 29$ измерений.

По заданной доверительной вероятности $P = 1 - \alpha = 0{,}99$ и числу степеней свободы $\nu = n - 1 = 5 - 1 = 4$, используя таблицу доверительных интервалов для среднего квадратического отклонения σ (приложение 7), находим нижнюю границу $\gamma_1 = 0{,}519$ и верхнюю границу $\gamma_2 = 4{,}390$. Для определения доверительного интервала для среднего квадратического отклонения воспользуемся неравенством (2.1.9): $s \cdot \gamma_1 < \sigma < s \cdot \gamma_2$.

Следовательно, искомый доверительный интервал равен:

$$0{,}08 \cdot 0{,}519 < \sigma < 0{,}08 \cdot 4{,}390 \text{ или } 0{,}04152 < \sigma < 0{,}3512.$$

Таким образом, искомый доверительный интервал имеет вид $(0{,}04152; 0{,}3512)$. То есть, если будет произведено достаточно большое число выборок заданного объёма, то в 99 % случаев из них доверительные интервалы накроют среднее квадратическое отклонение и только в 1 % случаев оцениваемое среднее квадратическое отклонение может выйти за пределы доверительных интервалов.

2.3 Задания для решения на практическом занятии

2.3.1 Из генеральной совокупности, исследования СВ ξ, извлечена выборка объёма $n = 100$. Данные представлены в виде статистического ряда:

варианта x_i	7,2	7,6	8,4	8,0	7,4
частота m_i	8	20	45	15	12

Найти выборочную среднюю, выборочную и исправленную дисперсии и соответствующие средние квадратические отклонения.

2.3.2 В итоге пяти измерений длины стержня одним прибором (без систематических ошибок) получены следующие результаты (в *мм*): $x_1 = 70$, $x_2 = 72$, $x_3 = 81$, $x_4 = 83$, $x_5 = 81$. Найти: а) выборочную среднюю длину стержня; б) выборочную и исправленную дисперсию ошибок прибора; в) выборочное и исправленное среднее квадратическое отклонение.

2.3.3 В итоге шести измерений некоторой физической величины одним прибором (без систематических ошибок) получены следующие результаты: $x_1 = 2$, $x_2 = 5$, $x_3 = 8$, $x_4 = 7$, $x_5 = 5$, $x_6 = 3$. Найти: а) выборочную среднюю исследуемой величины; б) выборочную и исправленную дисперсию; в) выборочное и исправленное среднее квадратическое отклонение.

2.3.4 Среди студентов факультета производится измерение роста. Из общего числа студентов, которые произвели измерения роста, произведена выборка в количестве 100 студентов. В результате получен интервальный статистический ряд.

Рост	Число студентов
$[160;164)$	9
$[164;168)$	15
$[168;172)$	20
$[172;176)$	29
$[176;180)$	17
$[180;184)$	7
$[184;188]$	3

Найти: а) выборочную среднюю роста обследованных студентов; б) выборочную и исправленную дисперсию роста обследованных студентов; в) выборочное и исправленное среднее квадратическое отклонение роста обследованных студентов.

2.3.5 В результате опыта произведено изучение случайной величины η. Из генеральной совокупности извлечена выборка, которая представлена в виде таблицы, представленной в задаче № 1.3.10. Найти: а) выборочную среднюю исследуемой величины, определённую в задаче № 1.3.10; б) выборочную и исправленную дисперсию данной случайной величины; в) выборочное и исправленное среднее квадратическое отклонение этой величины.

2.3.6 Найти доверительный интервал для оценки с надёжностью 0,98 неизвестного математического ожидания нормально распределённой случайной величины ξ генеральной совокупности, если заданы генеральное среднее квадратическое отклонение $\sigma = 4$, а выборочная средняя равна $\overline{x}_B = 20$ и объём выборки равен $n = 80$.

2.3.7 Выборка из большой партии электроламп содержит 200 ламп. Средняя продолжительность горения лампы выборки оказалась равной 800 часов. Найти с надёжностью 0,90 доверительный интервал для средней продолжительности горения лампы всей партии, если известно, что среднее квадратическое отклонение, продолжительности горения лампы равно 50 часов.

2.3.8 Станок-автомат штампует детали. По выборке объёма $n = 200$ вычислена выборочная средняя длины изготовленных деталей. Найти с надёжностью 0,99 точность ε, с которой выборочная средняя оценивает математическое ожидание линейного размера детали, если среднее квадратическое отклонение равно 10 *мм*.

2.3.9 Найти минимальный объём выборки, при котором с надёжностью 0,95, точность оценки математического ожидания генеральной совокупности нормально распределённой случайной величины по выборочной средней будет равна $\varepsilon = 0,4$, если среднее квадратическое отклонение равно $\sigma = 5$.

2.3.10 По данным независимых равноточных измерений некоторой физической величины найдена средняя выборочная $\overline{x}_B = 55$ и исправленное среднее квадратическое отклонение $s = 4$. Оценить истинное значение измеряемой величины при помощи доверительного интервала с надёжностью $P = 1 - \alpha = 0,95$.

2.3.11 По данным выборки объёма $n = 50$ из генеральной совокупности найдено исправленное среднее квадратическое отклонение $s = 2$ нормально распределённой случайной величины. Найти доверительный интервал, покрывающий генеральное среднее квадратическое отклонение σ с надёжностью 0,9.

2.3.12 Найти минимальный объём выборки, на основании которой можно было бы оценить математическое ожидание времени исполнения некоторой работы с ошибкой, не превышающей 10 часов, и надёжностью $P = 1 - \alpha = 0,99$, если предположить, что время выполнения указанной работы ξ является случайной величиной, имеющей нормальное распределение со средним квадратическим отклонением $\sigma = 80$.

2.3.13 В результате исследования некоторой случайной величины ξ, подчинённой нормальному закону распределения, произведена выборка: $x_1 = 3,5,\ x_2 = 4,1,\ x_3 = 5,7,\ x_4 = 4,1,\ x_5 = 3,5$. Оценить с помощью доверительного интервала математическое ожидание измерений. Доверительную вероятность $P = 1-\alpha$ принять равной 0,95. Найти минимальное число измерений, которое необходимо выполнить, чтобы с надёжностью 0,95 можно было утверждать, что предельная погрешность точечной оценки исследуемой величины не превышала 0,005. Построить доверительный интервал, покрывающий среднее квадратическое отклонение σ с надёжностью $P = 1-\alpha = 0{,}95$.

2.3.14 Определить численность выборки при обследовании остатков на расчётных счетах у клиентов банка, чтобы с вероятностью 0,9 предельная ошибка равнялась 15 денежным единицам, а среднее квадратическое отклонение – 350 денежным единицам.

2.3.15 При подсчёте объёмов ежедневных продаж джемперов предпринимателем в течение месяца получены следующие результаты: 1; 8; 7; 10; 13; 4; 9; 9; 10; 14; 8; 7; 9; 6; 8; 11; 3; 15; 5; 5; 11; 6; 5; 12; 6; 7; 8; 14; 11; 9. Найти: а) выборочную среднюю ежедневных продаж джемперов предпринимателем; б) выборочную и исправленную дисперсию; в) выборочное и исправленное среднее квадратическое отклонение объёмов продаж джемперов.

2.3.16 Случайная величина ξ имеет распределение Пуассона $P(\xi = m) = \dfrac{\lambda^m}{m!}e^{-\lambda}$ с неизвестным параметром λ. Используя метод моментов получения точечных оценок, найти по реализации выборки $(x_1, x_2, ..., x_n)$ значение оценки $\hat{\lambda}$ неизвестного параметра λ. Исходные данные: $x_1 = 16$, $x_2 = 33$, $x_3 = 18$, $x_4 = 36$, $x_5 = 12$, $x_6 = 34$, $x_7 = 25$, $x_8 = 17$, $n = 50$.

2.3.17 Случайная величина ξ имеет биномиальное распределение $P(\xi = m) = C_n^m p^m q^{n-m}$ с неизвестным параметром λ. Используя метод моментов получения точечных оценок, найти по реализации выборки $(x_1, x_2, ..., x_n)$ значение оценки $\hat{\lambda}$ неизвестного параметра λ. Исходные данные: $x_1 = 14$, $x_2 = 9$, $x_3 = 12$, $x_4 = 15$, $x_5 = 8$, $x_6 = 7$, $x_7 = 8$, $x_8 = 11$, $n = 30$.

2.4 Задания для контролируемой самостоятельной работы

2.4.1 Из генеральной совокупности случайной величины ξ извлечена выборка, которая представлена в виде таблицы в задаче 1.4.1. Найти выборочную среднюю, выборочную дисперсию и выборочное среднее квадратическое отклонение, исправленную дисперсию и исправленное среднее квадратическое отклонение наблюдаемых значений случайной величины ξ.

2.4.2 Из генеральной совокупности нормально распределённой случайной величины ξ – месячной заработной платы рабочих предприятия, рассчитанной в условных денежных единицах, извлечена выборка: x_1, x_2, x_3, x_4, x_5, x_6, x_7.

Требуется:

1) с помощью доверительного интервала оценить математическое ожидание случайной величины ξ – месячной заработной платы рабочих предприятия, при известном среднем квадратическом отклонении $\sigma = \sigma_0$ и уровне значимости $\alpha = \alpha_1$;

2) найти минимальный объём выборки, на основании которой можно было бы оценить математическое ожидание месячной заработной платы рабочих с ошибкой, не превышающей ε_1 условных денежных единиц месячной заработной оплаты при известном среднем квадратическом отклонении $\sigma = \sigma_0$ и уровне значимости $\alpha = \alpha_2$;

3) с помощью доверительного интервала оценить математическое ожидание нормально распределённой случайной величины ξ при неизвестном среднем квадратическом отклонении и уровне значимости $\alpha = \alpha_3$;

4) найти минимальный объём выборки, на основании которой можно было бы оценить математическое ожидание случайной величины ξ – месячной заработной платы рабочих с ошибкой, не превышающей ε_2 условных денежных единиц месячной заработной оплаты при неизвестном среднем квадратическом отклонении и уровне значимости $\alpha = \alpha_1$;

5) с помощью доверительного интервала оценить при уровне значимости $\alpha = \alpha_2$ среднее квадратическое отклонение нормально распределённой случайной величины ξ.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	σ_0	α_1	α_2	α_3	ε_1	ε_1
2.4.2.1	3,2	3,5	3,7	4,0	3,7	3,5	3,2	0,5	0,05	0,01	0,40	0,164	0,023
2.4.2.2	3,8	4,2	4,6	5,0	4,6	4,2	3,8	0,6	0,10	0,02	0,01	0,154	0,038
2.4.2.3	4,5	5,0	5,5	6,0	5,5	5,0	4,5	1,0	0,20	0,05	0,02	0,125	0,047
2.4.2.4	5,2	5,8	6,4	7,0	6,4	5,8	5,2	1,2	0,01	0,10	0,05	0,105	0,015
2.4.2.5	5,9	6,6	7,3	8,0	7,3	6,6	5,9	1,4	0,02	0,01	0,20	0,172	0,059
2.4.2.6	6,6	7,4	8,2	9,0	8,2	7,4	6,6	1,6	0,40	0,02	0,10	0,147	0,065
2.4.2.7	7,3	8,2	9,1	10,0	9,1	8,2	7,3	1,8	0,20	0,05	0,40	0,115	0,077
2.4.2.8	7,7	8,8	9,9	11,0	9,9	8,8	7,7	2,0	0,02	0,10	0,01	0,136	0,083
2.4.2.9	5,4	7,6	9,8	12,0	9,8	7,6	5,4	4,0	0,05	0,01	0,02	0,183	0,009
2.4.2.10	3,1	6,4	9,7	13,0	9,7	6,4	3,1	5,0	0,01	0,02	0,05	0,199	0,092
2.4.2.11	0,7	0,8	0,9	1,0	0,9	0,8	0,7	0,2	0,10	0,05	0,20	0,167	0,024
2.4.2.12	1,4	1,6	1,8	2,0	1,8	1,6	1,4	0,5	0,40	0,10	0,10	0,152	0,039
2.4.2.13	2,1	2,4	2,7	3,0	2,7	2,4	2,1	0,7	0,05	0,01	0,40	0,124	0,043
2.4.2.14	1,3	1,7	2,1	2,5	2,1	1,7	1,3	0,9	0,10	0,02	0,01	0,106	0,012
2.4.2.15	2,0	2,5	3,0	3,5	3,0	2,5	2,0	1,2	0,20	0,05	0,02	0,173	0,051
2.4.2.16	3,3	3,7	4,1	4,5	4,1	3,7	3,3	0,8	0,01	0,10	0,05	0,148	0,068
2.4.2.17	4,0	4,5	5,0	5,5	5,0	4,5	4,0	1,1	0,02	0,01	0,20	0,113	0,072
2.4.2.18	4,7	5,3	5,9	6,5	5,9	5,3	4,7	1,2	0,40	0,02	0,10	0,138	0,084

2.4.2.19	5,4	6,1	6,8	7,5	6,8	6,1	5,4	1,4	0,10	0,05	0,40	0,185	0,007
2.4.2.20	6,1	6,9	7,7	8,5	7,7	6,9	6,1	1,6	0,05	0,10	0,01	0,196	0,098
2.4.2.21	6,8	7,7	8,6	9,5	8,6	7,7	6,8	1,6	0,20	0,01	0,02	0,163	0,028
2.4.2.22	7,5	8,5	9,5	10,5	9,5	8,5	7,5	2,0	0,01	0,02	0,05	0,155	0,035
2.4.2.23	6,7	8,3	9,9	11,5	9,9	8,3	6,7	3,2	0,02	0,05	0,10	0,127	0,044
2.4.2.24	4,4	7,1	9,8	12,5	9,8	7,1	4,4	3,0	0,40	0,10	0,20	0,107	0,017
2.4.2.25	2,1	5,9	9,7	13,5	9,7	5,9	2,1	4,1	0,05	0,01	0,40	0,179	0,056
2.4.2.26	0,9	1,1	1,3	1,5	1,3	1,1	0,9	0,4	0,10	0,02	0,01	0,143	0,062
2.4.2.27	1,7	2,0	2,3	2,6	2,3	2,0	1,7	0,6	0,20	0,05	0,02	0,117	0,075
2.4.2.28	2,4	2,8	3,2	3,6	3,2	2,8	2,4	0,9	0,01	0,10	0,05	0,132	0,081
2.4.2.29	1,0	1,6	2,2	2,8	2,2	1,6	1,0	0,8	0,02	0,01	0,20	0,187	0,005
2.4.2.30	1,6	2,3	3,0	3,7	3,0	2,3	1,6	1,4	0,40	0,02	0,10	0,193	0,096

3 ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

Содержание: статистическая гипотеза, статистический критерий значимости проверки нулевой гипотезы, проверка гипотезы о нормальном распределении, статистическая проверка непараметрических гипотез по критерию согласия χ^2 Пирсона.

3.1 Теоретический материал по теме практического занятия

В предыущих разделах были рассмотрены методы получения оценок неизвестных параметров распределения. Нахождение точечных и интервальных оценок является предварительной стадией вероятностного эксперимента. Конечная цель исследования заключается в сравнительной оценке различных технологических процессов по их производительности, тчности и экономичности. Задачи такого рода носят название *задач сравнения.*

При решении экономических и инженерных задач приходится делать предположения о виде законов распределения случайных величин или о соотношении между их числовыми характеристиками. Такие предположения в математической статистике называют *гипотезами.*

Определение 3.1.1 Статистическая гипотеза называется *непараметрической*, если в ней сформулированы предположения относительно функции распределения.

Определение 3.1.2 Статистическая гипотеза называется *параметрической*, если в ней сформулированы предположения относительно значений параметров функции распределения исследуемой случайной величины.

Одновременно с выдвигаемой гипотезой законов распределения случайной величины можно выдвинуть и другие альтернативные гипотезы. Если выдвинутая гипотеза отвергается, то принимается другая альтернативная гипотеза. Исходя из этого предположения, гипотезы подразделяются на нулевые и альтернативные.

Определение 3.1.3 *Нулевой* гипотезой H_0 называют основную или выдвинутую первоначально гипотезу.

При выдвинутой нулевой гипотезе различие между сравниваемыми величинами отсутствуют, а наблюдаемые отклонения объясняются лишь случайными изменениями выборки.

Определение 3.1.4 *Альтернативной* называется гипотеза H_a, которая составляет конкуренцию нулевой гипотезе H_0 в том смысле, что если нулевая гипотеза отвергается, то за основу принимается альтернативная гипотеза.

Определение 3.1.5 Параметрическая гипотеза называется *простой*, если она содержит только одно предположение относительно параметров распределения.

Например, если a – математическое ожидание случайной величины ξ, то гипотеза $H_0 : a = 0$ является простой гипотезой.

Определение 3.1.6 Параметрическая гипотеза называется *сложной*, если она состоит из конечного или бесконечного числа простых гипотез.

Например, если a – математическое ожидание нормально распределённой случайной величины, то альтернативная гипотеза $H_a : a > 1$ является сложной, так как она состоит из бесконечного множества простых гипотез $H_0 : a = c$, где c – любое число большее единицы.

Определение 3.1.7 *Статистическим критерием* называют случайную величину K, с помощью которой принимаются решения о принятии или отвержении нулевой гипотезы.

Для проверки нулевых гипотез по выборочным данным рассчитываются частные значения входящих в критерий величин и таким образом получают частное или наблюдаемое значение критерия

Определение 3.1.8 *Ошибкой первого рода* называется ошибка отклонения верной нулевой гипотезы H_0.

Определение 3.1.9 *Уровнем значимости* статистического критерия называется вероятность α совершения ошибки первого рода.

Определение 3.1.10 *Ошибкой второго рода* называется ошибка принятия ложной нулевой гипотезы H_0.

Определение 3.1.11 Нулевой непараметрической гипотезой называется гипотеза относительного общего вида функции распределения случайной величины ξ, то есть гипотеза вида $H_0 : F(x) = F_0(x)$.

Проверка гипотезы о предполагаемом распределении производится на основании вычисления некоторой выборочной статистики (критерия), распределение которой получено из предположения истинности нулевой гипотезы и сравнения наблюдаемого значения этой выборочной статистики с критическим значением. Приняв ту или иную гипотезу, из неё выводят определённое следствие и рассматривают, насколько оно оправдывается на опыте, то есть проверяется согласие принятой гипотезы с опытом.

Для проверки нулевой гипотезы о виде закона распределения случайной величины рассмотрим один из наиболее распространённых критериев согласия, а именно критерий χ^2 Пирсона. Данный критерий позволяет проверить гипотезу о согласии данных выборки с конкретным распределением для любой случайной величины, как непрерывной, так и дискретной.

Критерий χ^2 – один из наиболее используемых критериев для проверки гипотезы о виде закона распределения случайных величин. Указанный критерий позволяет проводить проверку гипотезы соответствия опытного закона распределения теоретическому закону распределения не только в случаях, когда последний известен полностью, но и тогда, когда параметры теоретического закона распределения определяются на основе опытных данных. Таким образом, χ^2 Пирсона даёт возможность произвести проверку согласия эмпирической функции распределения с гипотетической функцией $F(x)$, принадлежащей к некоторому множеству функций определённого вида (нормальных, показательных, биномиальных и так далее).

Предположим, что проведено n независимых опытов, в каждом из которых случайная величина ξ приняла определённое значение. Результаты опытов оформляются в виде статистического интервального ряда, который содержит k интервалов, с указанием частот m_i, попадания случайной величины ξ в заданный интервал, где $i = \overline{1;k}$, а $\sum_{i=1}^{k} m_i = n$. При этом в каждом из интервалов желательно иметь 5–10 наблюдаемых значений случайной величины ξ.

Рассмотрим схему применения критерия χ^2 Пирсона.

1. Исходя из теоретического или предполагаемого закона распределения, находят вероятности p_i попадания случайной величины ξ в каждый из k полученных интервалов.

2. Вычисляем значение критерия χ^2 (выборочную статистику), которое соответствует опытным данным по формуле

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(m_i - np_i\right)^2}{np_i}. \tag{3.1.1}$$

3. Определяем число ν степеней свободы распределения по формуле

$$\nu = k - r - 1. \tag{3.1.2}$$

где r – число параметров предполагаемого закона распределения, вычисленных опытным путём, а k – число частичных интервалов.

4. Критерий χ^2 построен таким образом, что чем ближе к нулю наблюдаемое значение χ^2, тем вероятнее, что нулевая гипотеза справедлива. Поэтому для проверки нулевой гипотезы применяется критерий χ^2 с правосторонней критической областью. Для того чтобы проверить нулевую гипотезу, необходимо найти по таблицам квантилей χ^2-распределения (приложение, табл. А.4) по заданному уровню значимости α и числу степеней свободы ν критическое значение $\chi^2_{\alpha;\nu}$, удовлетворяющее условию $P\left(\chi^2 \geq \chi^2_{\alpha;\nu}\right) = \alpha$. Сравнивая наблюдаемое значение выборочной статистики χ^2, вычисленное по формуле (3.1.1), с критическим значением $\chi^2_{\alpha;\nu}$, принимаем одно из двух решений:

а) если $\chi^2_{набл} \geq \chi^2_{\alpha;\nu}$, то нулевая гипотеза H_0 отвергается в пользу альтернативной H_a, то есть считается, что гипотетическая функция $F(x)$ не согласуется с результатами эксперимента;

б) если $\chi^2_{набл} < \chi^2_{\alpha;\nu}$, то считается, что нет оснований для отклонения нулевой гипотезы H_0, то есть гипотетическая функция $F(x)$ хорошо согласуется с результатами эксперимента.

3.2 Примеры решения типовых задач

3.2.1 Результаты исследования 200 предприятий по стоимости основных производственных фондов (случайная величина ξ) в условных денежных единицах представлены в виде сгруппированного интервального статистического ряда:

Интервалы стоимости производственных фондов, у. е.	$[19;20)$	$[20;21)$	$[21;22)$	$[22;23)$	$[23;24)$	$[24;25]$
Частоты m_i	10	26	56	64	30	14

Проверить нулевую гипотезу о нормальном распределении стоимости производственных фондов предприятий при уровне значимости $\alpha = 0,05$.

Решение. Из условия следует, что точные параметры предполагаемого нормального распределения неизвестны, поэтому нулевая гипотеза может быть записана в виде: $H_0 : F(x)$ является функцией нормального распределения с двумя параметрами, математическим ожиданием $M(x)=\hat{a}=\overline{x}_B$ и дисперсией $D(x)=\hat{\sigma}^2=s^2$.

Для проверки предложенной нулевой гипотезы определяем значения x_i' середин интервалов и находим точечные оценки математического ожидания и среднего квадратического отклонения нормально распределённой случайной величины ξ.

$$\hat{a}=\overline{x}_B=\frac{\sum_{i=1}^{6} m_i x_i'}{n}=$$

$$=\frac{10\cdot 19,5+26\cdot 20,5+56\cdot 21,5+64\cdot 22,5+30\cdot 23,5+14\cdot 24,5}{200}=22,1 \text{ (у. е.)};$$

$$\hat{\sigma}^2=s^2=\frac{\sum_{i=1}^{6} m_i (x_i'-\overline{x}_B)^2}{n-1}=$$

$$=\frac{10\cdot(-2,6)^2+26\cdot(-1,6)^2+56\cdot(-0,6)^2+64\cdot 0,4^2+30\cdot 1,4^2+14\cdot 2,4^2}{199}=1,52 \text{ (у. е.)};$$

$$\hat{\sigma}=s=\sqrt{1,52} \text{ (у. е.)}.$$

Для проверки гипотезы о нормальном распределении случайной величины ξ с помощью критерия χ^2 в начале интервалы наблюдаемых значений нормируют, то есть выражают их в единицах среднего квадратического отклонения s: $u_i=(x_i-\overline{x}_B)/s$, причём наименьшее значение u_i полагают равным $(-\infty)$, наибольшее – $(+\infty)$.

Вычислим теоретические вероятности p_i попадания случайной величины ξ с математическим ожиданием 22,1 и средним квадратическим отклонением 1,52 условных денежных единиц в частичные интервалы $\left[x_{i-1};x_i\right)$ по формуле

$$p_i=P\left(x_{i-1}\le\xi<x_i\right)=\Phi(u_i)-\Phi(u_{i-1});\ \ (i=\overline{1;6}),$$

где $\Phi(u_i)=\frac{1}{\sqrt{2\pi}}\int_0^{u_i} e^{-\frac{t^2}{2}}dt$, значения которой находят по таблице значений функции Лапласа (приложение,табл. А.3).

Например, вероятность того, что случайная величина попадёт в первый частичный интервал $\left(-\infty;20\right)$, равна:

$$p_1 = P(-\infty < \xi < 20) = \Phi\left(\frac{20-20,1}{1,233}\right) - \Phi\left(\frac{-\infty - 22,1}{1,233}\right) = \Phi(-1,70) - \Phi(-\infty) =$$

$$= \Phi(+\infty) - \Phi(1,70) = 0,5 - 0,45543 = 0,04457.$$

$$p_2 = P(20 \le \xi < 21) = \Phi\left(\frac{21-20,1}{1,233}\right) - \Phi\left(\frac{20-22,1}{1,233}\right) = \Phi(-0,89) - \Phi(-1,70) =$$

$$= \Phi(1,70) - \Phi(0,89) = 0,45543 - 0,31327 = 0,14216$$

и так далее. После этого вычисляются теоретические частоты нормального распределения $n_{теор} = np_i$ и наблюдаемое значение критерия $\chi^2_{набл} = \sum_{i=1}^{k} \frac{(m_i - np_i)^2}{np_i}$. Вычисления, необходимые для определения наблюдаемого значения выборочной статистики χ^2, приведём в таблице:

Интервалы изменения наблюдаемых значений СВ ξ	Частоты m_i	Нормированные интервалы $[u_{i-1};u_i)$	p_i	np_i	$(m_i - np_i)^2$	$\frac{(m_i - np_i)^2}{np_i}$
$[19;20)$	10	$(-\infty;-1,70)$	0,04457	9	1	0,11
$[20;21)$	26	$[-1,70;-0,89)$	0,14216	28,4	5,76	0,20
$[21;22)$	56	$[-0,89;-0,08)$	0,28918	56,2	0,04	0,0007
$[22;23)$	64	$[-0,08;+0,73)$	0,29918	59,8	17,64	0,29
$[23;24)$	30	$[+0,73;+1,54)$	0,17092	34,2	17,64	0,52
$[24;25]$	14	$[+1,54;+\infty)$	0,06178	12,4	2,56	0,23
$\sum$	200		1,00000	200,0		$\chi^2_{набл} = 1,35$

В результате вычислений получили наблюдаемое значение критерия $\chi^2_{набл} = 1,35$. Найдём по таблице квантилей χ^2-распределения по заданному уровню значимости $\alpha = 0,05$ и числу степеней свободы $\nu = k - r - 1 = 6 - 2 - 1 = 3$ критическое значение $\chi^2_{0,05;3} = 7,815$. Так как $\chi^2_{набл} = 1,35 < \chi^2_{0,05;3} = 7,815$, то нет оснований для отклонения нулевой гипотезы о нормальном распределении стоимости производственных фондов предприятий с параметрами $a = 22{,}1$ и $\sigma = 1{,}233$.

3.2.2 В цехе с 5 станками ежедневно регистрируется число вышедших из строя станков (случайная величина ξ). Всего проведено 100 наблюдений. Получены следующие числовые данные $\left(n = \sum_{i=1}^{6} m_i = 100 \right)$:

Число вышедших из строя станков	0	1	2	3	4	5
Частоты m_i	8	28	31	18	9	6

Проверить гипотезу о том, что распределение числа вышедших из строя станков (случайная величина ξ) имеет распределение Пуассона при уровне значимости $\alpha = 0{,}05$.

Решение. Согласно условию задачи, необходимо проверить нулевую гипотезу H_0: $F(x)$ – функция распределения числа вышедших из строя станков имеет вид $F(x,\mu) = \sum_{i=1}^{x} \frac{\mu^i \cdot e^{-\mu}}{i!}$ с параметром $\hat{\mu} = \overline{x}_B = \frac{8 \cdot 0 + 28 \cdot 1 + 31 \cdot 2 + 18 \cdot 3}{100} +$ $+ \frac{9 \cdot 4 + 6 \cdot 5}{100} = 2{,}1$. Вычислим теоретические вероятности p_i выхода из строя x_i станков в течение n наблюдений по формуле Пуассона (приложение А):

$$p_i = P_n(x_i) = \frac{\hat{\mu}^{x_i} \cdot e^{-\hat{\mu}}}{x_i!} = \frac{(2{,}1)^{x_i} \cdot e^{-2{,}1}}{x_i!}; \; x_i = \overline{0;5}.$$

Вычисления, необходимые для определения наблюдаемого значения выборочной статистики χ^2, приведём в таблице:

Число вышедших из строя станков	Частоты m_i	p_i	np_i	$(m_i - np_i)^2$	$\frac{(m_i - np_i)^2}{np_i}$
0	8	0,122	12,2	17.64	1,45
1	28	0,257	25,7	5,29	0,21
2	31	0,270	27,0	16,00	0,59
3	18	0,189	18,9	0,81	0,04
4	9	0,099	9,9	0,81	0,08
5	6	0,063	6,3	0,03	0,01
$\sum$	100	1,000	100,0		$\chi^2_{набл} = 2{,}38$

В результате вычислений получили наблюдаемое значение критерия $\chi^2_{набл} = 2{,}38$. Найдём по таблице квантилей χ^2-распределения по заданному уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $\nu = k - r - 1 = 6 - 1 - 1 = 4$ критическое значение $\chi^2_{0,05;4} = 9{,}488$. Так как $\chi^2_{набл} = 2{,}38 < \chi^2_{0,05;3} = 9{,}488$, то нет оснований для отклонения нулевой гипотезы о том, что распределение числа вышедших из строя станков является распределением Пуассона с параметром $\mu = 2{,}1$.

Проверить гипотезу о том, что распределение числа вышедших из строя станков (случайная величина ξ) имеет распределение Пуассона при уровне значимости $\alpha = 0{,}05$.

Решение. Согласно условию задачи, необходимо проверить нулевую гипотезу H_0: $F(x)$ – функция распределения числа вышедших из строя станков имеет вид $F(x,\mu) = \sum\limits_{i=1}^{x} \frac{\mu^i \cdot e^{-\mu}}{i!}$ с параметром $\hat{\mu} = \overline{x}_B = \frac{8\cdot 0 + 28\cdot 1 + 31\cdot 2 + 18\cdot 3}{100} + \frac{9\cdot 4 + 6\cdot 5}{100} = 2{,}1$. Вычислим теоретические вероятности p_i выхода из строя x_i станков в течение n наблюдений по формуле Пуассона (приложение, табл. А.1):

$$p_i = P_n(x_i) = \frac{\hat{\mu}^{x_i} \cdot e^{-\hat{\mu}}}{x_i!} = \frac{(2{,}1)^{x_i} \cdot e^{-2{,}1}}{x_i!};\ x_i = \overline{0;5}.$$

Вычисления, необходимые для определения наблюдаемого значения выборочной статистики χ^2, приведём в таблице:

Число вышедших из строя станков	Частоты m_i	p_i	np_i	$(m_i - np_i)^2$	$\frac{(m_i - np_i)^2}{np_i}$
0	8	0,122	12,2	17.64	1,45
1	28	0,257	25,7	5,29	0,21
2	31	0,270	27,0	16,00	0,59
3	18	0,189	18,9	0,81	0,04
4	9	0,099	9,9	0,81	0,08
5	6	0,063	6,3	0,03	0,01
$\sum$	100	1,000	100,0		$\chi^2_{набл} = 2{,}38$

В результате вычислений получили наблюдаемое значение критерия $\chi^2_{набл} = 2{,}38$. Найдём по таблице квантилей χ^2-распределения по заданному уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $\nu = k - r - 1 =$

$= 6 - 1 - 1 = 4$ критическое значение $\chi^2_{0,05;4} = 9,488$. Так как $\chi^2_{набл} = 2,38 < \chi^2_{0,05;3} = 9,488$, то нет оснований для отклонения нулевой гипотезы о том, что распределение числа вышедших из строя станков является распределением Пуассона с параметром $\mu = 2,1$.

3.3 Задания для решения на практическом занятии

3.3.1 Используя критерий χ^2 Пирсона, при уровне значимости 0,05 установить, согласуется ли гипотеза о нормальном распределении генеральной совокупности с данными выборки объёма $n = 100$.

Частичный интервал	$[4;9)$	$[9;14)$	$[14;19)$	$[19;24)$	$[24;29)$	$[29;34)$	$[34;39]$
Частоты m_i	5	9	14	41	15	10	6

3.3.2 Имеются данные распределения товарооборота 50 магазинов города за год. Данные товарооборота приведены в таблице, где в первом столбце указан товарооборот в условных денежных единицах, а во втором – число магазинов, товарооборот которых оказался в данном интервале.

Частичный интервал	Частоты m_i	Частичный интервал	Частоты m_i	Частичный интервал	Частоты m_i
$[70;100)$	1	$[190;220)$	8	$[310;340)$	2
$[100;130)$	2	$[220;250)$	8	$[340;370)$	2
$[130;160)$	5	$[250;280)$	6	$[370;400)$	1
$[160;190)$	8	$[280;310)$	5	$[400;430)$	2

Используя критерий χ^2 Пирсона, при уровне значимости 0,10 установить, согласуется ли гипотеза о нормальном распределении генеральной совокупности с данными выборки объёма $n = 50$.

3.3.3 Из большой партии деталей было отобрано 500 штук с целью исследования закона распределения деталей по линейной длине. Результаты опытов приведены в таблице:

Частичный интервал	Частоты m_i	Частичный интервал	Частоты m_i
$[19,85;19,95)$	160	$[20,35;20,45)$	15
$[19,95;20,05)$	125	$[20,45;20,55)$	13
$[20,05;20,15)$	80	$[20,55;20,65)$	12
$[20,15;20,25)$	50	$[20,65;20,75)$	8
$[20,25;20,35)$	35	$[20,75;20,85]$	2

Используя критерий χ^2 Пирсона, при уровне значимости 0,02, установить, согласуется ли гипотеза о показательном распределении генеральной совокупности с данными выборки объёма $n = 500$.

3.3.4 На телефонной станции производились наблюдения за числом неправильных соединений в минуту. Всего проведено 200 наблюдений. Получены следующие числовые данные $\left(n = \sum_{i=1}^{6} m_i = 100 \right)$:

Число неверных соединений (случайная величина ξ)	0	1	2	3	4	5	6	7	8	9
Частоты m_i	8	12	20	29	31	33	27	24	10	6

Проверить гипотезу о том, что распределение числа вышедших из строя станков (случайная величина ξ) имеет распределение Пуассона при уровне значимости $\alpha = 0,1$.

3.3.5 Данные об урожайности ржи на различных участках сельскохозяйственного предприятия приведены в следующей таблице:

Урожайность, ц/га	$[8;11)$	$[11;14)$	$[14;17)$	$[17;20)$	$[20;23)$	$[23;26)$	$[26;29]$
Доля участка, % общей площади	3	10	24	30	20	8	5

Составить интервальный статистический ряд распределения относительных частот наблюдаемых значений непрерывной случайной величины ξ – урожайности ржи на различных участках сельхозугодий. Построить гистограмму и полигон относительных частот урожайности ржи. По видам гистограммы и полигона относительных частот случайной величины ξ сделать предварительный выбор вида закона распределения урожайности ржи, в центнерах с гектара, на различных участках сельхозугодий. Найти эмпирическую функцию распределения и построить её график. Вычислить числовые характеристики выборки: среднюю выборочную, выборочную дисперсию, выборочное среднеквадратическое отклонение.

Найти интервальные оценки параметров нормального распределения (математического ожидания и среднего квадратического отклонения) с доверительной вероятностью $P = 1 - \alpha = 0,98$.

Построить доверительный интервал для математического ожидания урожайности ржи, если известно среднее квадратическое отклонение $\sigma = 6$ ц/га.

Проверить согласие эмпирической функции распределения с модельной нормальной функцией распределения при помощи критерия χ^2 Пирсона.

Составить интервальный статистический ряд распределения относительных частот наблюдаемых значений непрерывной случайной величины ξ – урожайности ржи на различных участках сельхозугодий. Построить гистограмму и полигон относительных частот урожайности ржи. По видам гистограммы и полигона относительных частот случайной величины ξ сделать предварительный выбор вида закона распределения урожайности ржи, в центнерах с гектара, на различных участках сельхозугодий. Найти эмпирическую функцию распределения и построить её график. Вычислить числовые характеристики выборки: среднюю выборочную, выборочную дисперсию, выборочное среднеквадратическое отклонение.

Найти интервальные оценки параметров нормального распределения (математического ожидания и среднего квадратического отклонения) с доверительной вероятностью $P = 1 - \alpha = 0,98$.

Построить доверительный интервал для математического ожидания урожайности ржи, если известно среднее квадратическое отклонение $\sigma = 6$ ц/га. Проверить согласие эмпирической функции распределения с модельной нормальной функцией распределения при помощи критерия χ^2 Пирсона.

3.4 Задания для контролируемой самостоятельной работы

3.4.1 При исследовании непрерывной случайной величины ξ, эксперимент продублировали n раз при одних и тех же условиях. В результате эксперимента все полученные значения случайной величины ξ попали в промежуток $\left[x_0; x_k\right]$. После обработки данных интервал наблюдаемых значений случайной величины разбит на интервалы равной длины $\left[x_0; x_1\right)$, $\left[x_1; x_2\right)$, …, $\left[x_{k-1}; x_k\right]$. Результаты исследования случайной величины ξ представлены в виде интервального статистического ряда. В первой строке интервального ряда указываются интервалы наблюдаемых значений случайной величины ξ, а во второй – частоты m_i появления значений этой СВ ξ в указанном интервале $\left(\sum_i m_i = n\right)$.

1. Описать случайную величину ξ, которая могла бы соответствовать интервальному статистическому ряду вашего варианта, указав при этом единицы измерения значений выбранной случайной величины. При решении остальных пунктов задания необходимо указывать наименование вами выбранной случайной величины.

2. Составить интервальный статистический ряд распределения относительных частот наблюдаемых значений непрерывной случайной величины ξ.

Построить гистограмму и полигон относительных частот СВ ξ. По видам гистограммы и полигона относительных частот случайной величины ξ сделать предварительный выбор вида закона распределения этой случайной величины.

3. Найти эмпирическую функцию распределения и построить её график.

4. Вычислить числовые характеристики выборки: среднюю выборочную, выборочную дисперсию (смещённую и несмещённые оценки), выборочное среднеквадратическое отклонение.

5. В зависимости от того какое предположение Вы сделали насчёт вида закона для случайной величины ξ, записать её плотность распределения вероятностей и функцию распределения.

6. Исходя из предположения, которое было выдвинуто относительно закона распределения случайной величины ξ, найти интервальные оценки параметров распределения (математическое ожидание и среднее квадратическое отклонение) с доверительной вероятностью $P = 1 - \alpha$, где $\alpha = 0,01$, если номер варианта равен $4k$; $\alpha = 0,02$, если номер варианта равен $4k + 1$; $\alpha = 0,05$, если номер варианта равен $4k + 2$; $\alpha = 0,1$, если номер варианта равен $4k + 3$. По условию задачи среднее квадратическое отклонение неизвестно, поэтому для записи доверительного интервала для математического ожидания при известном среднем квадратическом отклонении принять его равным $\sigma = \dfrac{x_k - x_0}{2}$.

7. Проверить согласие эмпирической функции распределения с модельной нормальной функцией распределения при помощи критерия согласия χ^2.

3.4.1.1

18 – 20	20 – 22	22 – 24	24 – 26	26 – 28	28 – 30	30 – 32	32 – 34	34 – 36	36 – 38
14	36	130	210	102	82	61	45	15	5

3.4.1.2

31 – 34	34 – 37	37 – 40	40 – 43	43 – 46	46 – 49	49 – 52	52 – 55	55 – 58	58 – 61
3	16	32	45	86	175	135	56	32	20

3.4.1.3

1 – 5	5 – 9	9 – 13	13 – 17	17 – 21	21 – 25	25 – 29	29 – 33	33 – 37	37 – 41
9	13	27	78	167	86	54	47	12	7

3.4.1.4

37 – 42	42 – 47	47 – 52	52 – 57	57 – 62	62 – 67	67 – 72	72 – 77	77 – 82
12	16	46	76	133	87	45	21	14

3.4.1.5

12 – 18	18 – 24	24 – 30	30 – 36	36 – 42	42 – 48	48 – 54	54 – 60	60 – 66
12	24	43	65	97	64	32	13	10

3.4.1.6

1,3 – 2,0	2,0 – 2,7	2,7 – 3,4	3,4 – 4,1	4,1 – 4,8	4,8 – 5,5	5,5 – 6,2	6,2 – 6,9	6,9 – 7,6
14	30	54	80	42	21	17	8	4

3.4.1.7

4,6 – 5,4	5,4 – 6,2	6,2 – 7,0	7,0 – 7,8	7,8 – 8,6	8,6 – 9,4	9,4 – 10,2	10,2 – 11,0	11,0 – 11,8
8	10	14	28	64	25	16	9	6

3.4.1.8

36,5 – 36,7	36,7 – 36,9	36,9 – 37,1	37,1 – 37,3	37,3 – 37,5	37,5 – 37,7	37,7 – 37,9	37,9 – 37,9	37,9 – 38,1
4	8	16	21	61	27	16	7	2

3.4.1.9

168 – 171	171 – 174	174 – 177	177 – 180	180 – 183	183 – 186	186 – 189	189 – 192
2	5	11	22	48	18	8	6

3.4.1.10

3,4 – 3,8	3,8 – 4,2	4,2 – 4,6	4,6 – 5,0	5,0 – 5,4	5,4 – 5,8	5,8 – 6,2	6,2 – 6,6
3	5	8	12	16	36	20	4

3.4.1.11

65 – 66	66 – 67	67 – 68	68 – 69	69 – 70	70 – 71	71 – 72	72 – 73	73 – 74	75 – 76
10	24	63	158	121	78	56	23	12	5

3.4.1.12

47,1 – 47,3	47,3 – 47,5	47,5 – 47,7	47,7 – 47,9	47,9 – 48,1	48,1 – 48,3	48,3 – 48,5	48,5 – 48,7	48,7 – 48,9
6	18	38	87	126	96	26	12	5

3.4.1.13

12,3 – 12,7	12,7 – 13,1	13,1 – 13,5	13,5 – 13,9	13,9 – 14,3	14,3 – 14,7	14,7 – 15,1	15,1 – 15,5	15,5 – 15,9
4	17	35	87	129	91	27	12	3

3.4.1.14

189 – 194	194 – 199	199 – 204	204 – 209	209 – 214	214 – 219	219 – 224	224 – 229	229 – 234
4	6	13	34	75	98	70	12	3

3.4.1.15

10,1 – 11,1	11,1 – 12,1	12,1 – 13,1	13,1 – 14,1	14,1 – 15,1	15,1 – 16,1	16,1 – 17,1	17,1 – 18,1	18,1 – 19,1
5	14	44	73	41	18	12	6	2

3.4.1.16

162 – 166	166 – 170	170 – 174	174 – 178	178 – 182	182 – 186	186 – 190	190 – 194
6	11	15	32	56	29	14	5

3.4.1.17

2,0 – 2,3	2,3 – 2,6	2,6 – 2,9	2,9 – 3,2	3,2 – 3,5	3,5 – 3,8	3,8 – 4,1	4,1 – 4,4
4	6	18	26	64	21	9	4

3.4.1.18

8,15 – 8,18	8,18 – 8,21	8,21 – 8,24	8,24 – 8,27	8,27 – 8,30	8,30 – 8,33	8,33 – 8,36
4	8	18	38	21	10	6

3.4.1.19

2304 – 2314	2314 – 2324	2324 – 2334	2334 – 2344	2344 – 2354	2354 – 2364	2364 – 2374
3	6	20	37	18	9	5

3.4.1.20

158,2 – 158,4	158,4 – 158,6	158,6 – 158,8	158,8 – 160,0	160,0 – 160,2	160,2 – 160,4	160,4 – 160,6
3	10	24	68	29	15	5

3.4.1.21

0,12 – 0,125	0,125 – 0,13	0,13 – 0,135	0,135 – 0,14	0,14 – 0,145	0,145 – 0,15	0,15 – 0,155
2	8	12	38	16	10	4

3.4.1.22

350 – 351	351 – 352	352 – 353	353 – 354	354 – 355	355 – 356
4	10	14	32	19	11

3.4.1.23

48 – 52	52 – 56	56 – 60	60 – 64	64 – 68	68 – 72
3	12	32	17	12	4

3.4.1.24

15,1 – 15,3	15,3 – 15,5	15,5 – 15,7	15,7 – 15,9	15,9 – 16,1	16,1 – 16,3
2	5	16	35	18	4

3.4.1.25

17,0 – 17,6	17,6 – 18,2	18,2 – 18,8	19,4 – 20,0	20,6 – 21,2	21,8 – 22,4	23,0 – 23,6	23,6 – 24,2
3	5	16	41	56	39	18	6

3.4.1.26

0,45 – 0,85	0,85 – 1,25	1,25 – 1,65	1,65 – 2,05	2,05 – 2,45	2,45 – 2,85	2,85 – 3,25	3,25 – 3,65
3	8	18	39	17	13	9	5

3.4.1.27

20 – 25	25 – 30	30 – 35	35 – 40	40 – 45	45 – 50	50 – 55	55 – 60	60 – 65
3	9	36	93	158	86	39	5	3

3.4.1.28

14 – 17	17 – 20	20 – 23	23 – 26	26 – 29	29 – 32	32 – 35	35 – 38	38 – 41	41 – 44
3	5	37	127	413	148	43	13	7	4

3.4.1.29

5,15 – 5,35	5,35 – 5,55	5,55 – 5,75	5,75 – 5,95	5,95 – 6,15	6,15 – 6,35
2	9	32	47	38	10

3.4.1.30

1700 – 1720	1720 – 1740	1740 – 1760	1760 – 1780	1780 – 1800	1800 – 1820
34	45	98	142	86	39

4 МАТЕМАТИЧЕСКИЕ МОДЕЛИ ФУНКЦИЙ В РЕГРЕССИОННОМ АНАЛИЗЕ

Содержание: аппроксимация функций,эмпирические формулы, метод наименьших квадратов, параметры линейной аппроксимирующей функции, параметры квадратичной аппроксимирующей функции, параметры гиперболической аппроксимирующей функции, параметры показательной аппроксимирующей функции, параметры логарифмической аппроксимирующей функции.

4.1 Теоретический материал по теме практического занятия

Одной из важных задач математической статистики является нахождение связи между случайными величинами ξ и η. Во многих случаях одна из переменных может быть неслучайной, в то время как другая переменная имеет случайные отклонения, обусловленныеошибками измерений или другими причинами.

4.1.1 Эмпирические формулы
4.1.1.1 Постановка задачи аппроксимации функций

При эмпирическом изучении зависимости переменной y от переменной x производятся измерения переменной y при различных значениях величины x. Представим результаты эксперимента зависимости y от переменной x в виде таблицы 4.1.1 или точечной диаграммы (рис. 4.1.1).

Таблица 4.1.1 – Результаты эксперимента

x_i	x_1	x_2	x_3	…	x_i	…	x_n
y_i	y_1	y_2	y_3	…	y_i	…	y_n

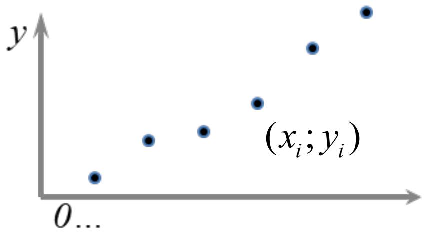

Рисунок 4.1.1 – Точечная диаграмма

Важное практическое применение имеет следующая задача: найти аналитическое выражение зависимости функции y от переменной x, то есть необходимо подобрать непрерывную функцию $y = \varphi(x)$, такую, чтобы $y_i = \varphi(x_i)$. Функции $y = \varphi(x)$, полученные в результате решения такого рода задач, называются *эмпирическими* или *аппроксимирующими*.

Подбор аппроксимирующей функции по экспериментальным данным (x_i, y_i), $i = \overline{1,n}$, можно осуществить различными методами исходя из разных моделей эксперимента. Рассмотрим две наиболее распространённые модели.

1. *Интерполяционная модель.*

Предположим, что некоторая функция $y = f(x)$ задана таблицей значений (x_i, y_i), $i = \overline{1,n}$, (таблица 4.1.1). Пусть значения функции y при выбранных значениях x_i точно известны, а её аналитическое выражение $y = f(x)$ в общем случае считается неизвестной функцией. Необходимо по таблице значений функции $y = f(x)$ подобрать эмпирическую функцию $y = \varphi(x)$, удовлетворяющую условиям $\varphi(x_i) = y_i$, $i = \overline{1,n}$.

Геометрически задача подбора функции $\varphi(x)$ по заданным частным значениям функции $y = f(x)$ означает построение кривой $y = \varphi(x)$, проходящей через точки плоскости с координатами (x_i, y_i), $i = \overline{1,n}$. Через заданные точки (x_i, y_i) можно провести бесконечное множество различных кривых. Таким образом, задача подбора эмпирической функции $y = \varphi(x)$ по точным значениям или узлам (x_i, y_i), $i = \overline{1,n}$, неизвестной функции $y = f(x)$ не определена.

Предположим теперь, что функция $y = \varphi(x)$ не произвольная, а является многочленом степени $n-1$ при n узлах. Тогда задача подбора эмпирической функции $\varphi(x)$ приобретает более определённый характер: по таблице значений $(x_i, y_i),\ i=\overline{1,n}$, неизвестной функции $y = f(x)$ построить аппроксимирующий многочлен $P_{n-1}(x)$, удовлетворяющий условиям $P_{n-1}(x_i),\ i=\overline{1,n}$. Построенный многочлен называется *интерполяционным многочленом*. Существуют различные виды интерполяционных многочленов: Лагранжа, Ньютона и другие.

В практических расчётах часто используется *кусочно-линейная интерполяция* (функцию на отрезке $[x_k; x_{k+1}]$, $k=\overline{1,n}$, заменяют многочленом Лагранжа первой степени), *кусочно-квадратичная интерполяция* (по любым трём узлам x_{k-1}, x_k, x_{k+1} строят многочлен Лагранжа второй степени) или *сплайн-интерполяция*.

2. *Аппроксимация функций по методу наименьших квадратов.*

Пусть результаты измерений $(x_i, y_i), i=\overline{1,n}$ приведены в виде таблицы или точечной диаграммы. Аналитическое выражение для функции $y = f(x)$ нам неизвестно. Значения функции $y = f(x)$ при значениях $x = x_i$ определены приближённо с некоторой случайной погрешностью.

Наличие случайных погрешностей делает нецелесообразным подбор такой эмпирической функции $y=\varphi(x)$, которая бы точно описывала все экспериментальные данные, то есть график её проходил бы через точки $(x_i, y_i),\ i=\overline{1,n}$. В этом случае предпочтительно подобрать такую аппроксимирующую функцию $\varphi(x)$, которая бы «сглаживала» случайные погрешности измерений.

Задача построения аппроксимирующей функции $y=\varphi(x)$ по экспериментальным данным в модели метода наименьших квадратов состоит из двух этапов:

1) определение вида аппроксимирующей функции $y=\varphi(x)$, то есть выбор класса функций, к которому принадлежит аппроксимирующая функция (линейная, квадратичная, степенная, линейная и другая);

2) определение параметров $a_0, a_1, a_2, ..., a_k$ аппроксимирующей функции $y=\varphi(x, a_0, a_1, a_2, ..., a_k)$ выбранного вида.

4.1.1.2 Определение вида аппроксимирующей функции

В общем случае, не существует стандартных методов, которые образовывали бы строгую теоретическую базу для выбора аппроксимирующей функции $y=\varphi(x)$. Однако выбрать аппроксимирующую функцию, если придерживаться следующих приведённых пунктов.

1. При подборе аппроксимирующей функции $y = \varphi(x)$ следует учитывать характер расположения экспериментальных точек (x_i, y_i), $i = \overline{1, n}$, на точечной диаграмме.

Предположим, что на точечной диаграмме точки (x_i, y_i), $i = \overline{1, n}$, располагаются вдоль некоторой линии, например, вдоль прямой. Тогда в качестве аппроксимирующей функции рекомендуется выбирать линейную функцию $\varphi(x) = a \cdot x + b$, зависящую от двух параметров a и b.

Если расположение точек (x_i, y_i), $i = \overline{1, n}$, на точечной диаграмме напоминает параболу, то в качестве аппроксимирующей функции лучше выбрать квадратичную функцию $\varphi(x) = ax^2 + bx + c$, зависящую от трёх параметров a, b и c.

Обычно при «зрительном» анализе экспериментальных точек, на точечной диаграмме используют графики известных функций. Если характер расположения точек на точечной диаграмме напоминает одну из кривых некоторого семейства (степенных, показательных, логарифмических и других функций), то аппроксимирующая функция $\varphi(x)$ выбирается из этого семейства функций.

Таким образом, на основании точечной диаграммы, можно подобрать достаточно много кривых, проходящих вблизи экспериментальных точек $(x_i, y_i), i = \overline{1, n}$, поэтому вопрос, какую именно из них следует выбрать в качестве аппроксимирующей, решается неоднозначно.

2. При подборе аппроксимирующей функции $y = \varphi(x)$ рекомендуется, если это возможно, провести её «линеаризацию», то есть сделать такую замену переменных, которая сводила бы нелинейную функцию $y = \varphi(x)$ к линейной функции.

Предположим, что замена переменных $t = t(x)$, $z = z(y)$ приводит нелинейную функцию $y = \varphi(x)$ к линейной функции $z = at + b$. Если теперь в системе координат Otz построить точечную диаграмму (t_i, z_i), $i = \overline{1, n}$, то точки должны располагаться вдоль прямой линии. В противном случае не рекомендуется выбирать функцию $\varphi(x)$ из рассматриваемого класса функций.

Как правило, «линеаризация» достигается логарифмированием исходной функции.

3. При подборе аппроксимирующей функции $y = \varphi(x)$ следует по возможности сочетать характер расположения точек на точечной диаграмме с логически-профессиональным анализом условия задачи.

4. Для простоты дальнейших этапов построения аппроксимирующей функции $\varphi(x) = \varphi(x, a_0, a_1, a_2, ..., a_k)$ желательно, чтобы они были линейными относительно параметров a_k.

5. При подборе аппроксимирующей функции для описания «криволинейных» зависимостей предпочтительно использовать параболы не выше третьей степени.

Предположим, что используя вышеприведённые пункты 1–5, нами определён класс степенных, показательных, логарифмических и других функций, из которых необходимо выбрать аппроксимирующую функцию. После выбора вида аппроксимирующей функции найдём параметры $a_0, a_1, a_2, ..., a_k$ для предполагаемой функции $\varphi(x, a_0, a_1, a_2, ..., a_k)$.

4.1.1.3 Определение параметров аппроксимирующей функции по методу наименьших квадратов

Пусть при эмпирическом изучении зависимости переменной y от переменной x производятся измерения переменной y при различных значениях величины x. Результаты эксперимента зависимости y от переменной x представлены в виде таблицы 4.1.1.

Предположим, что известна аппроксимирующая функция $\varphi(x) = \varphi(x, a_0, a_1, a_2, ..., a_k)$. Составляем разности $e_i = \varphi(x_i) - y_i$, где $i = \overline{1, n}$, которые называются *отклонениями* или *погрешностями*. Числа y_i это числа стоящие во второй строке таблицы 4.1.1, $\varphi(x_i)$ – значения аппроксимирующей функции $\varphi(x)$ при соответствующих значениях аргумента x_i, чисел из первой строки таблицы 4.1.1. Требуется подобрать параметры $a_0, a_1, a_2, ..., a_k$ аппроксимирующей функции $\varphi(x)$ так, чтобы отклонения e_i оказались наименьшими.

Наиболее распространённым критерием малости отклонений является критерий, лежащий в основе метода наименьших квадратов: параметры функции $\varphi(x)$ выбрать так, чтобы оказалась минимальной сумма квадратов отклонений: $S = \sum_{i=1}^{n} e_i^2 \geq 0$. Подбор параметров $a_0, a_1, a_2, ..., a_k$ в аппроксимирующей функции $\varphi(x) = \varphi(x, a_0, a_1, a_2, ..., a_k)$ по методу наименьших квадратов производится таким образом, что сумма квадратов отклонений ординат $\varphi(x_i)$ аппроксимирующей функции от наблюдаемых значений y_i была минимальной, то есть

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(\varphi(x_i, a_0, a_1, ..., a_k) - y_i\right)^2 \to \min. \qquad (4.1.1)$$

Задача определения тех значений параметров $a_0, a_1, a_2, ..., a_k$, при которых функция $S = S(a_0, a_1, a_2, ..., a_k)$, имеющая непрерывные частные производные, достигает минимума, то, на основании необходимого условия локального экстремума функции нескольких переменных, сводится к решению системы уравнений:

$$\frac{\partial S}{\partial a_0}=0,\ \frac{\partial S}{\partial a_1}=0,\ \frac{\partial S}{\partial a_2}=0,\ \ldots,\ \frac{\partial S}{\partial a_k}=0. \qquad (4.1.2)$$

Для проверки адекватности предполагаемой эмпирической формулы экспериментальным данным удобно использовать разделённые разности.

Так, например, эмпирическая формула $y=ax+b$ оказывается пригодной, то есть неплохо будет отражать вид зависимости между переменными x и y, если первые разделённые разности $f(x_{i+1},x_i)=\frac{y_{i+1}-y_i}{x_{i+1}-x_i}$ мало отличаются друг от друга.

В случаях, когда таблица результатов измерений имеет постоянный шаг , то есть каждый из последующих значений x отличается от предыдущего на постоянное число, то достаточно сравнить неразделённые разности $\Delta y_i=y_{i+1}-y_i$, и по их приближённому равенству можно сделать вывод об экспериментальной линейной зависимости.

Эмпирическая функция распределения $y=ax^2+bx+c$ оказывается пригодной в случае, когда мало отличаются друг от друга разделённые разности второго порядка $f(x_{i+2};x_{i+1};x_i)=\frac{f(x_{i+2};x_{i+1})-f(x_{i+1};x_i)}{x_{i+2}-x_{i+1}}$, где в числителе стоят разделённые разности первого порядка. Если каждое последующее значение x отличается от предыдущего на одно и то же число, то сравнивают неразделённые разности второго порядка $\Delta^2 y_i=\Delta y_{i+1}-\Delta y_i$, где $\Delta y_i=y_{i+1}-y_i$ $\left(i=\overline{1,n}\right)$ и, в случае приближённого равенства, делаем вывод о квадратичной зависимости экспериментальных данных.

В дальнейшем рассмотрим методику определения параметров некоторых аппроксимирующих функций по методу наименьших квадратов.

4.1.2 Определение параметров некоторых эмпирических функций, вычисляемых по методу наименьших квадратов

4.1.2.1 Определение параметров линейной аппроксимирующей функции, вычисляемых по методу наименьших квадратов

Предположим необходимо установить зависимость между двумя величинами x и y, результаты измерений которых занесены в таблицу 4.1.1.

Рассмотрим в декартовой системе координат Oxy точки $M_1(x_1;y_1)$, $M_2(x_2;y_2)$, …, $M_n(x_n;y_n)$, которые изображены на рисунке 4.1.2.

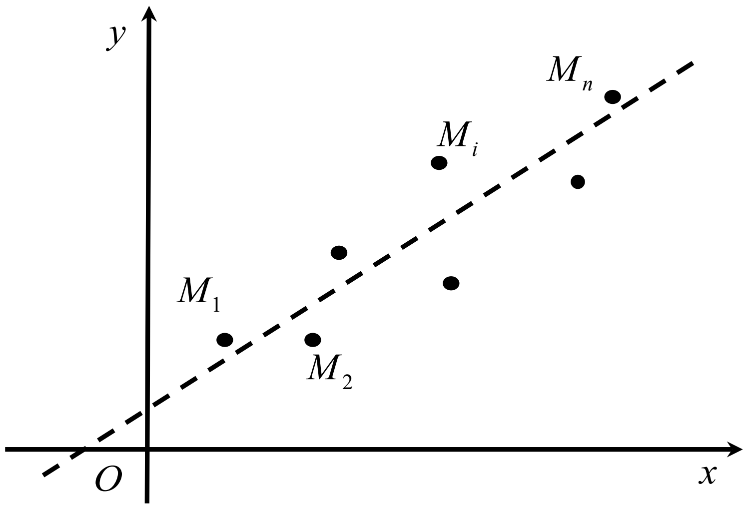

Рисунок 4.1.2 – Экспериментальные точки и линия эмпирической функции, определённой в пункте 4.1.2.1

Из точечной диаграммы видим, что экспериментальные точки располагаются вдоль некоторой прямой. Но если эти точки почти лежат на одной прямой, то естественно предположить, что между двумя величинами x и y существует линейная зависимость, которая выражается формулой

$$y = ax + b, \tag{4.1.3}$$

где a и b – параметры, которые необходимо определить.

Представим равенство (4.1.3) в виде

$$ax + b - y = 0. \tag{4.1.4}$$

Точки $M_i(x_i, y_i)$ принадлежат искомой прямой (4.1.3) с определённой погрешностью. Подставляя экспериментальные значения x_i и y_i $\left(i = \overline{1,n}\right)$ в формулу (4.1.4), получаем систему отклонений

$$\begin{cases} e_1 = ax_1 + b - y_1, \\ e_2 = ax_2 + b - y_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ e_n = ax_n + b - y_n. \end{cases} \tag{4.1.5}$$

По методу наименьших квадратов подберём неизвестные параметры a и b таким образом, чтобы полученная величина $S = \sum_{i=1}^{n} e_i^2 \geq 0$ была наименьшей.

Учитывая систему (4.1.5) преобразуем функцию S:

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(ax_i + b - y_i \right)^2 . \quad (4.1.6)$$

Функция двух переменных $S = S(a,b)$ является функцией двух переменных относительно переменных a и b. Чтобы эта функция принимала минимальное значение необходимо частные производные функции по данным переменным приравнять к нулю:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0, \\ \dfrac{\partial S}{\partial b} = 0, \end{cases} \quad \text{или} \quad \begin{cases} \sum\limits_{i=1}^{n} 2(ax_i + b - y_i)x_i = 0, \\ \sum\limits_{i=1}^{n} 2(ax_i + b - y_i) = 0. \end{cases}$$

Преобразуем эту систему уравнений

$$\begin{cases} a \cdot \sum\limits_{i=1}^{n} x_i^2 + b \cdot \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i, \\ a \cdot \sum\limits_{i=1}^{n} x_i + b \cdot n = \sum\limits_{i=1}^{n} y_i. \end{cases} \quad (4.1.7)$$

Решая систему (4.1.7), определяем оптимальные значения неизвестных параметров a и b в эмпирической формуле (4.1.3). На рисунке 4.1.2 штриховой линией изображён график аппроксимированной функции $y = ax + b$, коэффициенты которой определены из системы (4.1.7).

4.1.2.2 Определение параметров квадратичной аппроксимирующей функции по методу наименьших квадратов

Предположим необходимо установить зависимость между двумя величинами x и y, результаты измерений которых занесены в таблицу 4.1.1.

Рассмотрим в декартовой системе координат Oxy точки $M_1(x_1; y_1)$, $M_2(x_2; y_2)$, …, $M_n(x_n; y_n)$, которые изображены на рисунке 4.1.3.

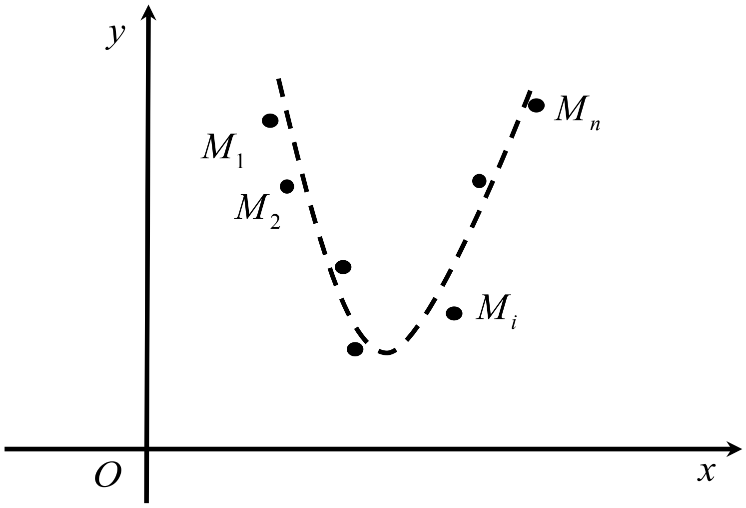

Рисунок 4.1.3 – Экспериментальные точки и линия эмпирической функции, определённой в пункте 4.1.2.2

Из точечной диаграммы видим, что экспериментальные точки располагаются вдоль некоторой линии, которая напоминает участок параболы. Следовательно, в качестве аппроксимирующей функции можно выбрать параболу вида $y = ax^2 + bx + c$. Подберём параметры a, b и c таким образом, чтобы сумма квадратов отклонений аппроксимирующей функции $y = ax^2 + bx + c$ от экспериментальных значений был наименьшим.

Представим равенство $y = ax^2 + bx + c$ в виде

$$ax^2 + bx + c - y = 0. \tag{4.1.8}$$

Точки $M_i(x_i, y_i)$ принадлежат искомой параболе с определённой погрешностью. Подставляя экспериментальные значения x_i и y_i $\left(i = \overline{1,n}\right)$ в формулу (4.1.8), получаем систему отклонений

$$\begin{cases} e_1 = ax_1^2 + bx_1 + c - y_1, \\ e_2 = ax_2^2 + bx_2 + c - y_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ e_n = ax_n^2 + bx_n + c - y_n. \end{cases} \tag{4.1.9}$$

По методу наименьших квадратов подберём неизвестные параметры a, b и c таким образом, чтобы полученная величина $S = \sum\limits_{i=1}^{n} e_i^2 \geq 0$ была наименьшей. Учитывая систему (4.1.9), преобразуем функцию S:

$$S=\sum_{i=1}^{n}e_i^2=\sum_{i=1}^{n}\left(ax_i^2+bx_i+c-y_i\right)^2 . \tag{4.1.10}$$

Функция трёх переменных $S=S(a,b,c)$ является функцией трёх переменных относительно переменных a, b и c. Чтобы эта функция принимала минимальное значение необходимо частные производные функции по данным переменным приравнять к нулю:

$$\begin{cases}\dfrac{\partial S}{\partial a}=0,\\ \dfrac{\partial S}{\partial b}=0,\\ \dfrac{\partial S}{\partial c}\end{cases} \quad \text{или} \quad \begin{cases}\sum\limits_{i=1}^{n}2(ax_i^2+bx_i+c-y_i)\cdot x_i^2=0,\\ \sum\limits_{i=1}^{n}2(ax_i^2+bx_i+c-y_i)\cdot x_i=0,\\ \sum\limits_{i=1}^{n}2(ax_i^2+bx_i+c-y_i)\cdot 1=0,\end{cases}$$

Преобразуем эту систему уравнений

$$\begin{cases}a\cdot\sum\limits_{i=1}^{n}x_i^4+b\cdot\sum\limits_{i=1}^{n}x_i^3+c\cdot\sum\limits_{i=1}^{n}x_i^2=\sum\limits_{i=1}^{n}x_i^2y_i,\\ a\cdot\sum\limits_{i=1}^{n}x_i^3+b\cdot\sum\limits_{i=1}^{n}x_i^2+c\cdot\sum\limits_{i=1}^{n}x_i=\sum\limits_{i=1}^{n}x_i y_i,\\ a\cdot\sum\limits_{i=1}^{n}x_i^2+b\cdot\sum\limits_{i=1}^{n}x_i+c\cdot n=\sum\limits_{i=1}^{n}y_i.\end{cases} \tag{4.1.11}$$

Решая систему (4.1.11) определяем оптимальные значения неизвестных параметров a, b и c в эмпирической формуле $y=ax^2+bx+c$. На рисунке 4.1.3 штриховой линией изображён график аппроксимированной функции $y=ax^2+bx+c$, коэффициенты которой определены из системы (4.1.11).

4.1.2.3 Определение параметров гиперболической аппроксимирующей функции по методу наименьших квадратов

Предположим необходимо установить зависимость между двумя величинами x и y, результаты измерений которых занесены в таблицу 4.1.1.

Рассмотрим в декартовой системе координат Oxy точки $M_1(x_1;y_1)$, $M_2(x_2;y_2)$, …, $M_n(x_n;y_n)$, которые изображены на рисунке 4.1.4.

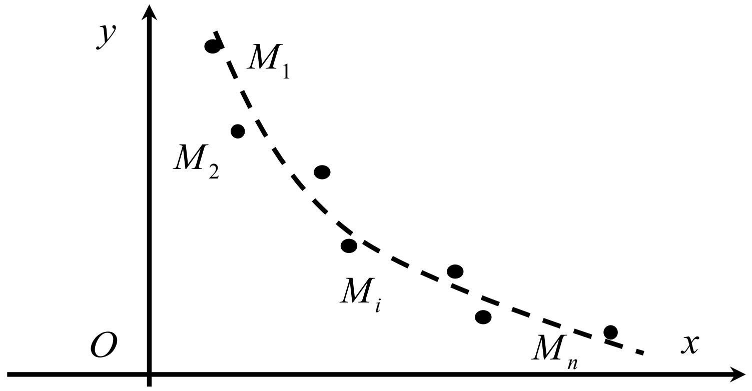

Рисунок 4.1.4 – Экспериментальные точки и линия эмпирической функции, определённой в пункте 4.1.2.3

Из точечной диаграммы видим, что экспериментальные точки располагаются вдоль некоторой линии, которая напоминает участок гиперболы. Следовательно, в качестве аппроксимирующей функции можно выбрать гиперболу вида $y=\dfrac{a}{x}+b$. Подберём параметры a, b и c таким образом, чтобы сумма квадратов отклонений аппроксимирующей функции $y=\dfrac{a}{x}+b$ от экспериментальных значений был наименьшим.

Введём новую переменную $z=\dfrac{1}{x}$, тогда уравнение $y=\dfrac{a}{x}+b$ принимает вид $y=a\cdot z+b$, то есть переменные y и z будут связаны линейной зависимостью с параметрами a и b. В этом случае по методу наименьших квадратов, неизвестные параметры определяются из системы (4.1.7), где вместо экспериментальных точек x_i выбираем точки $z_i=1/x_i$, а $i=\overline{1,n}$. В результате получаем систему линейных алгебраических уравнений

$$\begin{cases} a\cdot\sum\limits_{i=1}^{n}\dfrac{1}{x_i^2}+b\cdot\sum\limits_{i=1}^{n}\dfrac{1}{x_i}=\sum\limits_{i=1}^{n}\dfrac{y_i}{x_i}, \\ a\cdot\sum\limits_{i=1}^{n}\dfrac{1}{x_i}+b\cdot n=\sum\limits_{i=1}^{n}y_i, \end{cases} \tag{4.1.12}$$

решение которой, определяет неизвестные параметры a и b эмпирической гиперболической зависимости.

Выбор гиперболической функции между переменными x и y, может быть объяснён геометрически: если изображение экспериментальных точек $M_i(x_i,y_i)$ группируются вдоль некоторой гиперболы (рис. 4.1.4), то выбирают гиперболическую зависимость.

4.1.2.4 Определение параметров экспоненциальной аппроксимирующей функции по методу наименьших квадратов

Предположим необходимо установить зависимость между двумя величинами x и y, результаты измерений которых занесены в таблицу 4.1.1.

Рассмотрим в декартовой системе координат Oxy точки $M_1(x_1;y_1)$, $M_2(x_2;y_2)$, …, $M_n(x_n;y_n)$, которые изображены на рисунке 4.1.5.

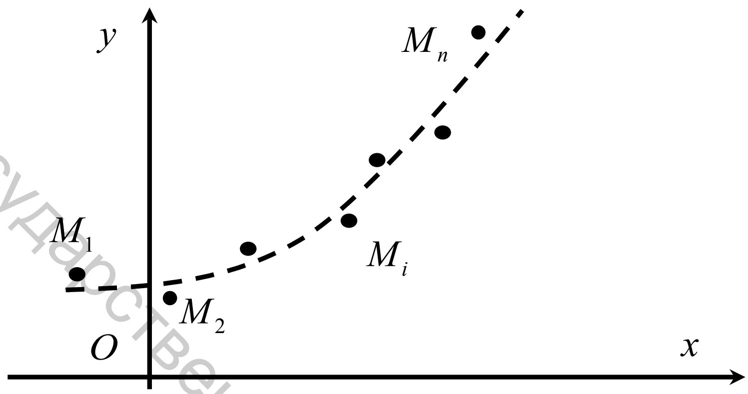

Рисунок 4.1.5 – Экспериментальные точки и линия эмпирической функции, определённой в пункте 4.1.2.4

Из точечной диаграммы видим, что экспериментальные точки располагаются вдоль некоторой линии, которая напоминает участок показательной функции. Следовательно, в качестве аппроксимирующей функции можно выбрать экспоненциальную функцию вида $y = b \cdot e^{a \cdot x}$. Подберём параметры a и b таким образом, чтобы сумма квадратов отклонений аппроксимирующей функции $y = b \cdot e^{a \cdot x}$ от экспериментальных значений был наименьшим.

Чтобы свести исследование предполагаемой функции к линейному случаю, прологарифмируем выражение $y = b \cdot e^{a \cdot x}$ по основанию e и воспользуемся свойствами логарифма:

$$\ln y = \left(\ln b \cdot e^{a \cdot x}\right) = \ln b + \ln e^{a \cdot x} = \ln b + a \cdot x \cdot \ln e = a \cdot x + \ln b,$$

или

$$\ln y = a \cdot x + \ln b. \qquad (4.1.13)$$

Сравнив формулу (4.1.13) с уравнением линейной зависимостью (4.1.3), делаем вывод о том, что переменные $\ln y$ и x связаны линейной зависимостью с параметрами a и $\ln b$. Для определения неизвестных параметров воспользуемся системой (4.1.7), заменив в ней y_i на величину $\ln y_i$, а параметр b на параметр $\ln b$. В результате получаем систему линейных алгебраических уравнений относительно неизвестных параметров a и $\ln b$:

$$\begin{cases} a \cdot \sum\limits_{i=1}^{n} x_i^2 + \ln b \cdot \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i \cdot \ln y_i, \\ a \cdot \sum\limits_{i=1}^{n} x_i + n \cdot \ln b = \sum\limits_{i=1}^{n} \ln y_i, \end{cases} \qquad (4.1.14)$$

решение которой определяет неизвестные параметры a и $\ln b$ или b эмпирической экспоненциальной зависимости $y = b \cdot e^{a \cdot x}$.

Выбор экспоненциальной функции между переменными x и y, может быть объяснён геометрически: если изображение экспериментальных точек $M_i(x_i, y_i)$ группируются вдоль некоторой экспоненты (рисунок 4.1.5), то выбирают экспоненциальную зависимость.

4.1.2.5 Определение параметров логарифмической аппроксимирующей функции по методу наименьших квадратов

Предположим необходимо установить зависимость между двумя величинами x и y, результаты измерений которых занесены в таблицу 4.1.1.

Рассмотрим в декартовой системе координат Oxy точки $M_1(x_1; y_1)$, $M_2(x_2; y_2)$, …, $M_n(x_n; y_n)$, которые изображены на рисунке 4.1.6.

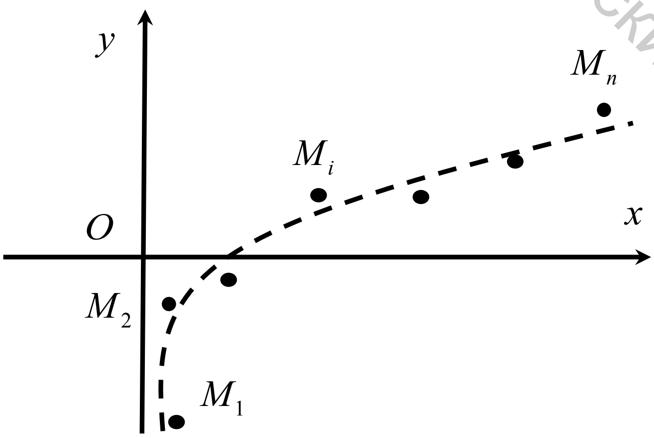

Рисунок 4.1.6 – Экспериментальные точки и линия эмпирической функции, определённой в пункте 4.1.2.5

Из точечной диаграммы видим, что экспериментальные точки располагаются вдоль некоторой линии, которая напоминает участок логарифмической функции. Следовательно, в качестве аппроксимирующей функции можно выбрать экспоненциальную функцию вида $y = a \cdot \ln x + b$. Подберём параметры a и b таким образом, чтобы сумма квадратов отклонений $S(a;b) = \sum_{i=1}^{n}\left(a \cdot \ln x_i + b - y_i\right)^2$ аппроксимирующей функции $y = a \cdot \ln x + b$ от экспериментальных значений был наименьшим.

Сравнив формулу $y = a \cdot \ln x + b$ с уравнением линейной зависимостью (4.1.3), делаем вывод о том, что переменные y и $\ln x$ связаны линейной зависимостью с параметрами a и b. Для определения неизвестных параметров воспользуемся системой (4.1.7), заменив в ней x_i на величину $\ln x_i$. В результате получаем систему линейных алгебраических уравнений относительно неизвестных параметров a и $\ln b$:

$$\begin{cases} a \cdot \sum_{i=1}^{n} \ln^2 x_i + b \cdot \sum_{i=1}^{n} \ln x_i = \sum_{i=1}^{n} (y_i \cdot \ln x_i), \\ a \cdot \sum_{i=1}^{n} \ln x_i + b \cdot n = \sum_{i=1}^{n} y_i, \end{cases} \tag{4.1.15}$$

решение которой, определяет неизвестные параметры a и b эмпирической логарифмической зависимости $y = a \cdot \ln x + b$.

Выбор логарифмической функции между переменными x и y, может быть объяснён геометрически: если изображение экспериментальных точек $M_i(x_i, y_i)$ группируются вдоль графика некоторой логарифмической функции (рис. 4.1.6), то выбирают логарифмическую зависимость.

4.2 Примеры решения типовых задач

4.2.1 В результате работы фирмы в течение 5 лет получены следующие данные

x_i	1	2	3	4	5
y_i	1	4	3	6	6

где x_i – инвестиционные вложения; y_i – прибыль, которую она получает за год. Значения инвестиционных вложений и прибыли в таблице задаются в условных единицах. Индекс $i = \overline{1;5}$ указывает год исследования работы фирмы. Предполагая, что между прибылью и инвестициями в производство существует некоторая зависимость найти аппроксимирующую функцию, используя метод наименьших квадратов. Определить наиболее правдоподобную величину прибыли на 6 год деятельности фирмы, если она в этот год инвестирует в производство 7 у. е. Считать условия работы фирмы неизменными и независимыми от внешних факторов.

Решение. Построим в декартовой системе координат точки $M_i(x_i; y_i)$, где $i=\overline{1,5}$, то есть точки $M_1(1;1)$, $M_2(2;4)$, $M_3(3;3)$, $M_4(4;6)$, $M_5(5;6)$. Экспериментальные точки M_i изображены на рисунке 4.2.1.

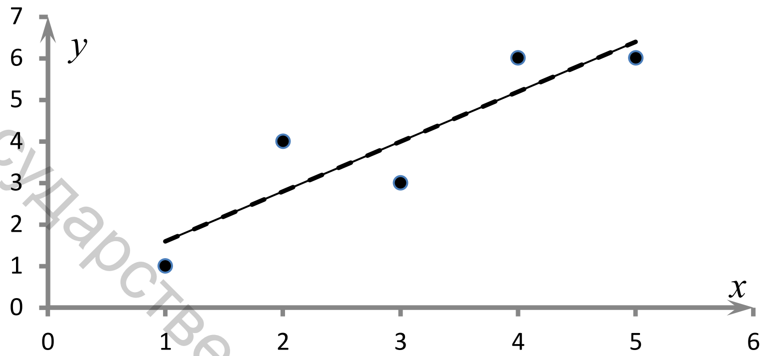

Рисунок 4.2.1 – Экспериментальные точки и линия эмпирической функции, пример 4.2.1

Зависимость получения прибыли (y) фирмы от инвестиционных вложений (x), выражается функцией $y = f(x)$. Заданные точки приблизительно располагаются вдоль некоторой прямой, поэтому можно предположить линейную зависимость прибыли фирмы от ее инвестиционных вложений.

Используя метод наименьших квадратов, определим линейную функцию $y = ax + b$, которая аппроксимирует исходную $y = f(x)$.

Найдем параметры k и b эмпирической формулы $y = kx + b$ из системы линейных уравнений (4.1.7).

В условии данной задачи зависимость получения прибыли фирмы от инвестиционных вложений проводилась в течение пяти лет, а, следовательно, $n = 5$. Построим вспомогательную таблицу.

i	x_i	y_i	x_i^2	$x_i \cdot y_i$
1	1	1	1	1
2	2	4	4	8
3	3	3	9	9
4	4	6	16	24
5	5	6	25	30
$\sum$	15	20	55	72

В результате получаем систему линейных уравнений:

$$\begin{cases} 55 \cdot a + 15 \cdot b = 72, \\ 15 \cdot a + 5 \cdot b = 20. \end{cases}$$

Решением данной системы являются числа $a = 1,2$ и $b = 0,4$.

Следовательно, зависимость полученной прибыли фирмой от инвестиционных вложений, может быть выражена функцией: $y = 1,2 \cdot x + 0,4$. На рисунке 4.2.1 эмпирическая функция изображена штриховой линией. Используя полученную зависимость, можно предположить, что на следующий год, при вложении в производство 7 у. е., фирма получит наиболее правдоподобную прибыль равную 8,8 у. е.

4.2.2 В результате измерений зависимых величин x и y получены следующие данные, которые представлены в виде таблицы

x	–1	0	1	2	3	4	5
y	11,0	1,6	–1,4	0,5	4,8	16,0	28,0

Определить вид зависимости между указанными величинами и найти параметры эмпирической формулы.

Решение. Построим в прямоугольной декартовой системе координат точки $M_i(x_i; y_i)$, где $i = \overline{1,7}$, то есть точки $M_1(-1; 11)$, $M_2(0; 1,6)$, $M_3(1; -1,4)$, $M_4(2; 0,5)$, $M_5(3; 4,8)$, $M_6(4; 16)$, $M_7(5; 28)$. Экспериментальные точки M_i изображены на рисунке 4.2.2.

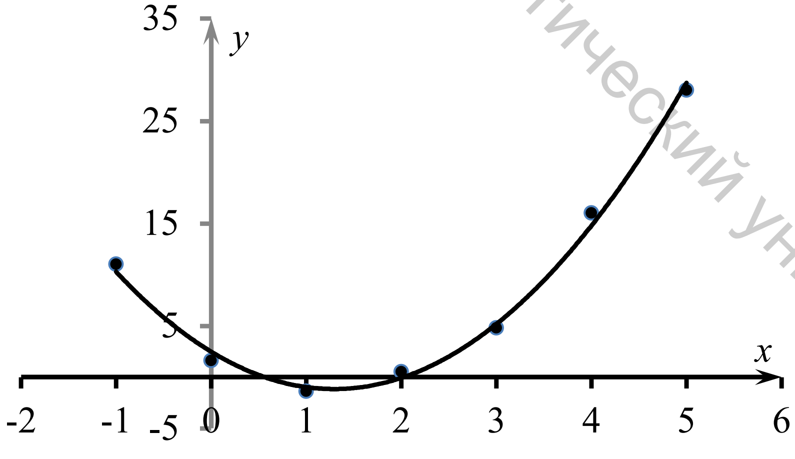

Рисунок 4.2.2 – Экспериментальные точки и линия эмпирической функции, пример 4.2.2

Из рисунка видим, что экспериментальные точки незначительно отклоняются от некоторой параболы, которая на рисунке 4.2.2 изображена штриховой линией. Поэтому, мы можем предположить, что зависимость между переменными x и y будет выражаться формулой

$$y = ax^2 + bx + c.$$

Данный вывод подтверждается и приблизительным равенством неразделённых разностей второго порядка, при постоянном шаге: $\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i$. Найдём неразделённые разности первого порядка: $\Delta y_1 = y_2 - y_1 = -9{,}4$, $\Delta y_2 = y_3 - y_2 = -3$, $\Delta y_3 = 1{,}9$, $\Delta y_4 = 4{,}3$, $\Delta y_5 = 11{,}2$, $\Delta y_6 = 12$. Вычисляем неразделённые разности второго порядка: $\Delta^2 y_1 = 6{,}4$, $\Delta^2 y_2 = 4{,}9$, $\Delta^2 y_3 = 2{,}4$, $\Delta^2 y_4 = 6{,}9$, $\Delta^2 y_5 = 0{,}8$. Первые неразделённые разности монотонно возрастают, а вторые колеблются, в принципе, в сравнительно небольшом промежутке.

Построим вспомогательную таблицу.

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i \cdot y_i$	$x_i^2 \cdot y_i$
1	-1	11	1	-1	1	-11	11
2	0	1,6	0	0	0	0	0
3	1	-1,4	1	1	1	-1,4	-1,4
4	2	0,5	4	8	16	1	2
5	3	4,8	9	27	81	14,4	43,2
6	4	16	16	64	256	64	256
7	5	28	25	125	625	140	700
$\sum$	14	60,5	56	224	980	207	1010,8

Подставляя полученные значения в систему (4.1.11), приходим к системе линейных алгебраических уравнений

$$\begin{cases} 980 \cdot a + \cdot 224 \cdot b + 56 \cdot c = 1010{,}8, \\ 224 \cdot a + 56 \cdot b + 14 \cdot c = 207, \\ 56 \cdot a + 14 \cdot b + 7 \cdot c = 60{,}5. \end{cases}$$

Решаем систему линейных алгебраических уравнений любым методом или с использованием прикладных программ. Решением данной системы являются числа $a = 2{,}1762$, $b = -5{,}6333$ и $c = 2{,}5$.

Таким образом, эмпирическая функция имеет вид

$$y = 2{,}1762 \cdot x^2 - 5{,}6333 \cdot x + 2{,}5.$$

График этой функции изображён на рисунке 4.2.2 штриховой линией.

4.2.3 В результате измерений зависимых величин x и y получены следующие данные, которые представлены в виде таблицы

x	1	2	3	6	12
y	9,5	4,5	3,8	1,6	1,4

Определить вид зависимости между указанными величинами и найти параметры эмпирической формулы.

Решение. Построим в прямоугольной декартовой системе координат точки $M_i(x_i; y_i)$, где $i=\overline{1,5}$, то есть точки $M_1(1; 9{,}5)$, $M_2(2; 4{,}5)$, $M_3(3; 3{,}8)$, $M_4(6; 1{,}6)$, $M_5(12; 1{,}4)$. Экспериментальные точки M_i изображены на рисунке 4.2.3.

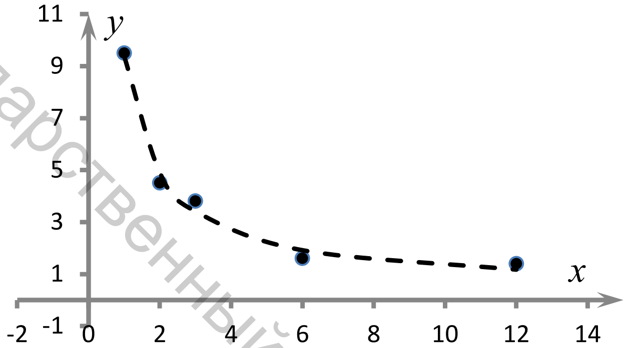

Рисунок 4.2.3 – Экспериментальные точки и линия эмпирической функции, пример 4.2.3

Из рисунка видим, что экспериментальные точки незначительно отклоняются от некоторой гиперболы, которая на рисунке 4.2.3 изображена штриховой линией. Поэтому, мы можем предположить, что зависимость между переменными x и y будет выражаться формулой

$$y = a/x + b.$$

Построим вспомогательную таблицу.

i	x_i	y_i	$1/x_i$	$1/x_i^2$	y_i/x_i
1	1	9,5	1	1	9,5
2	2	4,5	0,5	0,25	2,25
3	3	3,8	0,3333	0,1111	1,26667
4	6	1,6	0,1667	0,0278	0,26667
5	12	1,4	0,0833	0,0069	0,11667
$\sum$	24	20,8	2,0833	1,3958	13,4

Подставляя полученные значения в систему (4.1.12), приходим к системе линейных алгебраических уравнений

$$\begin{cases} 1,3958 \cdot a + \cdot 2,0833 \cdot b = 13,4, \\ 2,0833 \cdot a + 5 \cdot b = 20,8. \end{cases}$$

Решаем систему линейных алгебраических уравнений любым методом или с использованием прикладных программ. Решением данной системы являются числа $a = 8,9684$, $b = 0,4232$ и $c = 2,5$.

Таким образом, эмпирическая функция имеет вид

$$y = \frac{8,9684}{x} + 0,4232.$$

График этой функции изображён на рисунке 4.2.3 штриховой линией.

4.2.4 В результате измерений зависимых величин x и y получены следующие данные, которые представлены в виде таблицы

x	-1	-0,5	0	0,5	1
y	0,6	2	2,2	4,3	9,8

Определить вид зависимости между указанными величинами и найти параметры эмпирической формулы.

Решение. Построим в прямоугольной декартовой системе координат точки $M_i(x_i; y_i)$, где $i = \overline{1,5}$, то есть точки $M_1(-1; 0,6)$, $M_2(-0,5; 2,0)$, $M_3(0; 2,2)$, $M_4(0,5; 4,3)$, $M_5(1; 9,8)$. Экспериментальные точки M_i изображены на рисунке 4.2.4.

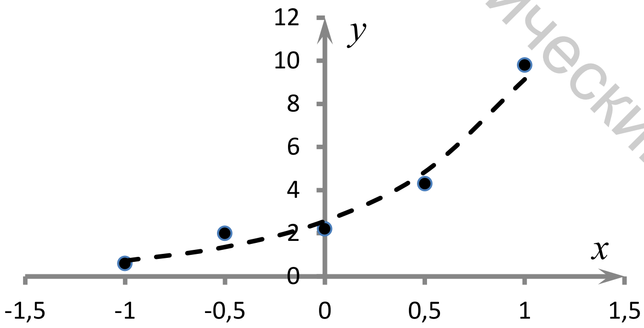

Рисунок 4.2.4 – Экспериментальные точки и линия эмпирической функции, пример 4.2.4

Из рисунка видим, что экспериментальные точки незначительно отклоняются от некоторой экспоненциальной линии, которая на рисунке 4.2.4 изображена штриховой линией. Поэтому, мы можем предположить, что зависимость между переменными x и y будет выражаться формулой

$$y = b \cdot e^{a \cdot x}.$$

Для определения неизвестных параметров a и b построим вспомогательную таблицу.

i	x_i	y_i	x_i^2	$\ln y_i$	$x_i \cdot \ln y_i$
1	-1	0,6	1	-0,511	0,51083
2	-0,5	2	0,25	0,6931	-0,3466
3	0	2,2	0	0,7885	0
4	0,5	4,3	0,25	1,4586	0,72931
5	1	9,8	1	2,2824	2,28238
$\sum$	0	18,9	2,5	4,7118	3,17594

Подставляя полученные значения в систему (4.1.14), приходим к системе линейных алгебраических уравнений

$$\begin{cases} 2{,}5 \cdot a + 0 \cdot \ln b = 3{,}17594, \\ a \cdot 0 + 5 \cdot \ln b = 4{,}7118, \end{cases}$$

Решаем систему линейных алгебраических уравнений любым методом или с использованием прикладных программ. Решением данной системы являются числа $a = 1{,}2704$, $\ln b = 0{,}9424$ или $b = 2{,}566$.

Таким образом, эмпирическая функция имеет вид

$$y = 2{,}566 \cdot e^{1{,}2704 \cdot x}.$$

График этой функции изображён на рисунке 4.2.4 штриховой линией.

4.2.5 В результате измерений зависимых величин x и y получены следующие данные, которые представлены в виде таблицы

x	1	2	3	4	5
y	-3,6	1,1	1,9	4,4	4,6

Определить вид зависимости между указанными величинами и найти параметры эмпирической формулы.

Решение. Построим в прямоугольной декартовой системе координат точки $M_i(x_i; y_i)$, где $i = \overline{1,5}$, то есть точки $M_1(1; -3{,}6)$, $M_2(2; 1{,}1)$, $M_3(3; 1{,}9)$, $M_4(4; 4{,}4)$, $M_5(5; 4{,}6)$. Экспериментальные точки M_i изображены на рисунке 4.2.5.

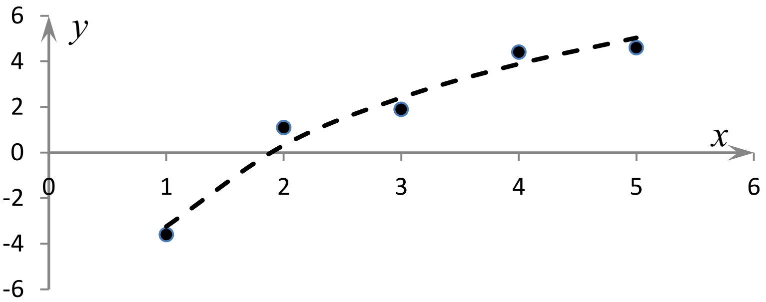

Рисунок 4.2.5 – Экспериментальные точки и линия эмпирической функции, пример 4.2.5

Из рисунка видим, что экспериментальные точки незначительно отклоняются от некоторой логарифмической кривой, которая на рисунке 4.2.5 изображена штриховой линией. Поэтому, мы можем предположить, что зависимость между переменными x и y будет выражаться формулой

$$y = a \cdot \ln x + b\,.$$

Для определения неизвестных параметров a и b построим вспомогательную таблицу.

i	x_i	y_i	$\ln x_i$	$\ln^2 x_i$	$y_i \cdot \ln x_i$
1	1	-3,6	0	0	0
2	2	1,1	0,6931	0,4805	0,76246
3	3	1,9	1,0986	1,2069	2,08736
4	4	4,4	1,3863	1,9218	6,0997
5	5	4,6	1,6094	2,5903	7,40341
$\sum$	15	8,4	4,7875	6,1995	16,3529

Подставляя полученные значения в систему (2.2.13), приходим к системе линейных алгебраических уравнений

$$\begin{cases} 6{,}1995 \cdot a + \cdot 4{,}7875 \cdot b = 16{,}3529, \\ 4{,}7875 \cdot a + 5 \cdot b = 8{,}4. \end{cases}$$

Решаем систему линейных алгебраических уравнений любым методом или с использованием прикладных программ. Решением данной системы являются числа $a = 5{,}1439$ и $b = -3{,}2453$.

Таким образом, эмпирическая функция имеет вид

$$y = y = 5{,}1439 \cdot \ln x - 3{,}2453\,.$$

График этой функции изображён на рисунке 4.2.5 штриховой линией.

4.3 Задания для решения на практическом занятии

4.3.1 При исследовании зависимости длины s (в *метрах*) тормозного пути от скорости движения v (км/ч) автомобиля получены следующие экспериментальные данные:

s	10	15	20	25	30	35	40
v	7,5	8,8	9,8	12,5	15	20	27

Определить вид аппроксимирующей функции.

4.3.2 Равноточные измерения некоторой величины S – торговая площадь магазина, отвечающие ряду значений аргумента T – годовой товарооборот в миллионах рублей. Экспериментальные данные зависимости товарооборота от площади представлены в виде таблицы:

S_i	1	2	3	4	5
T_i	2,2	3,1	4,5	5,5	6,2

Определить наиболее правдоподобный вид зависимости между переменными и найти эмпирическую функцию зависимости для этих переменных.

4.3.3 Измерение температуры корпуса работающего агрегата, производимого с интервалом 5 минут, дало следующие результаты:

t	5	10	15	20	25
T	59,3	59,8	60,1	64,9	70,2

Считая, что зависимость между временем измерения температуры и самой температурой корпуса работающего агрегата имеет вид $T = at^2 + bt + c$, найти оценки параметров a, b и c по методу наименьших квадратов.

4.3.4 Себестоимость C (у. е.) одного экземпляра книги в зависимости от тиража n (тыс. экз.) характеризуется данными, приведёнными в таблице:

n_i	1	2	3	4	5	10	20	30	50	100
C_i	1,25	1,15	1,00	0,80	0,65	0,41	0,36	0,20	0,15	0,1

Применяя способ наименьших квадратов, определить коэффициенты для гиперболической зависимости вида $y = a/x + b$.

4.3.5 Конденсатор заряжен до напряжения U_0, отвечающего моменту начала отсчёта времени, после чего он разряжается через некоторое сопротивление. Напряжение U измеряется с округлением до пяти вольт в различные моменты времени. Результаты измерений приведены в таблице:

t_i, (c)	0	1	2	3	4	5	6	7	8	9	10
U_i, (В)	100	75	55	40	30	20	15	10	10	5	5

Известно, что зависимость напряжения U от времени t имеет вид: $U = U_0 \cdot e^{-\omega \cdot t}$. Применяя метод наименьших квадратов, определить коэффициенты U_0 и ω.

4.3.6 Станок-автомат выпускает стержни различной длины l. Отклонения Δl от стандартной длины стержня приведены в таблице:

l	1	2	3	4	5
Δl	-0,5	0,1	0,9	1,6	1,8

Известно, что зависимость отклонения Δl от длины стержня l имеет вид: $\Delta l = a \cdot \ln l + b$. Применяя метод наименьших квадратов, определить коэффициенты a и b.

4.3.7 Равноточные измерения некоторой величины y, отвечающие ряду значений аргумента x, привели к результатам, которые приведены в таблице:

x	0	1	2	3	4
y	-2,4	-5,8	-10,8	-14,3	-19

Установить вид зависимости между величинами x и y (геометрически и с помощью разделённых разностей). Найти параметры эмпирической функции зависимости для этих переменных x и y.

4.3.8 Равноточные измерения некоторой величины y, отвечающие ряду значений аргумента x, привели к результатам, которые приведены в таблице:

x	-3	-2	-1	0	1	2	3
y	-7	1,8	6,9	10	8,1	1,1	-5,4

Установить вид зависимости между величинами x и y (геометрически и с помощью разделённых разностей). Найти параметры эмпирической функции зависимости для этих переменных x и y.

4.3.9 Пусть зависимость признака y от признака x характеризуется ниже приведённой таблицей

x	2	4	6	8	10
y	5,9	2	1,2	0,9	0,2

Изобразить экспериментальные точки на координатной плоскости. Предположить вид аппроксимирующей функции и, исходя из предположения, определить её параметры.

4.3.10 Пусть зависимость признака y от признака x характеризуется ниже приведённой таблицей:

x	-2	-1	0	1	2
y	0,5	0,3	1,3	2,3	9,9

Изобразить экспериментальные точки на координатной плоскости. Предположить вид аппроксимирующей функции и, исходя из предположения, определить её параметры.

4.3.11 Пусть зависимость признака y от признака x характеризуется ниже приведённой таблицей:

x	1	2	3	4	5
y	-0,9	1	1,7	2,1	2,4

Изобразить экспериментальные точки на координатной плоскости. Предположить вид аппроксимирующей функции и, исходя из предположения, определить её параметры.

4.3.12 Пусть зависимость признака y от признака x характеризуется ниже приведённой таблицей:

x	-1	0	1	2	3
y	-9,8	-3,1	0,5	-1,4	-6,5

Изобразить экспериментальные точки на координатной плоскости. Предположить вид аппроксимирующей функции и, исходя из предположения, определить её параметры.

4.3.13 Пусть зависимость признака y от признака x характеризуется ниже приведённой таблицей:

x	1	3	5	7	9
y	0,9	-1,1	-3,1	-4,8	-6,9

Изобразить экспериментальные точки на координатной плоскости. Предположить вид аппроксимирующей функции и, исходя из предположения, определить её параметры.

4.4 Задания для контролируемой самостоятельной работы

4.4.1 При исследовании зависимости признака y от признака x получены следующие экспериментальные данные:

x	x_1	x_2	x_3	x_4	x_5	x_6
y	y_1	y_2	y_3	y_4	y_5	y_6

Определить вид зависимости между переменными и найти эмпирическую функцию зависимости для этих переменных.

3.4.1.1

x	-1	0	1	2	3	4
y	1,2	2,8	5,1	6,9	8,8	11,1

3.4.1.2

x	-2	-1	0	1	2	3
y	19,6	8,8	1,6	-0,6	-0,1	5,2

3.4.1.3

x	1	3	5	7	9	11
y	10,8	5,2	3,6	3,1	2,8	2,6

3.4.1.4

x	-1,5	-1	-0,5	0	0,5	1
y	0,1	0,5	1,1	3,2	8,5	22

3.4.1.5

x	1	2	3	4	5	6
y	3,2	4,4	5,1	5,6	6,2	6,8

3.4.1.6

x	0	1	2	3	4	5
y	2,1	4,7	8,3	10,8	14,1	16,8

3.4.1.7

x	-3	-2	-1	0	1	2
y	5,1	0,2	-0,8	1,9	8,8	19,6

3.4.1.8

x	2	3	4	5	6	7
y	5,2	3,5	3,1	2,5	2,3	2

3.4.1.9

x	-2	-1,5	-1	-0,5	0	0,5
y	0,1	0,2	0,5	2,1	8,8	40,1

3.4.1.10

x	2	3	4	5	6	7
y	-1,6	-,8	0,2	0,3	0,6	0,9

3.4.1.11

x	-2	-1	0	1	2	3
y	-1,9	0,3	1,9	4,2	5,8	7,9

3.4.1.12

x	0	1	2	3	4	5
y	3,3	-4,2	-5,1	0,3	10,6	27,3

3.4.1.13

x	1	2	3	4	5	6
y	5,1	3,4	2,9	2,8	2,6	2,4

3.4.1.14

x	-0,5	0	0,5	1	1,5	2
y	0,9	1,8	3,2	5,4	9	14,5

3.4.1.15

x	1	3	5	7	9	11
y	2,1	5,2	6,8	7,8	8,5	9,2

3.4.1.16

x	1	2	3	4	5	6
y	-1,1	2,2	4,9	8,1	10,8	14,1

3.4.1.17

x	-2	-1	0	1	2	3
y	0,3	-6,8	-8,2	-2,9	7,8	24,5

3.4.1.18

x	2	4	6	8	10	12
y	0,1	-1,1	-1,4	-1,5	-1,7	-1,8

3.4.1.19

x	1	2	3	4	5	6
y	2,4	3	3,8	4,6	5,5	6,8

3.4.1.20

x	2	4	6	8	10	12
y	0,1	2,2	3,4	4,2	4,9	5,5

3.4.1.21

x	-3	-2	-1	0	1	2
y	-10,8	-8,1	-4,9	-1,4	0,7	4,2

3.4.1.22

x	0	1	2	3	4	5
y	-3,1	-7,9	-9,1	-5,8	1,2	11,6

3.4.1.23

x	1	3	5	7	9	11
y	4	-0,8	-1,4	-1,9	-2,3	-2,5

3.4.1.24

x	-2	-1	0	1	2	3
y	1	1,9	2,9	5	8,2	13,2

3.4.1.25

x	3	4	5	6	7	8
y	-0,6	0,5	1,4	2,1	2,7	3,3

3.4.1.26

x	2	3	4	5	6	7
y	-6,2	-2,8	0,1	3,4	5,8	9,2

3.4.1.27

x	-1	0	1	2	3	4
y	9,7	-4,9	-13,6	-16,8	-13,8	-5,1

3.4.1.28

x	2	5	8	11	14	17
y	4,2	0,5	-0,5	-1	-1,2	-1,3

3.4.1.29

x	-1	-0,5	0	0,5	1	1,5
y	1,6	2,7	3,9	6	8,6	13,1

3.4.1.30

x	0,5	1	1,5	2	2,5	3
y	2,2	4,8	6,6	7,8	8,6	9,4

ЛИТЕРАТУРА

1. Бирюкова, Л. Г. Теория вероятностей и математическая статистика : учеб. пособие / Л. Г. Бирюкова, Г. И. Бобрик, В. И. Матвеев. – Москва : Инфра-М, 2019. – 160 с.

2. Белько, И. В. Теория вероятностей и математическая статистика. Примеры и задачи / И. В. Белько, Г. П. Свирид. – Минск : Новое знание, 2002. – 250 с.

3. Белько, И. В. Теория вероятностей и математическая статистика / И. В. Белько, Г. П. Свирид. – Минск : ООО «Новое знание», 2004. – 251 с.

4. Булдык, Г. М. Теория вероятностей и математическая статистика : учеб. пособие для студентов экономических спец. вузов / Г. М. Булдык. – Минск : Вышэйшая школа, 1989. – 285 с.

5. Венецкий, И. Г. Теория вероятностей и математическая статистика : учеб. пособие для студентов экономических спец. вузов / И. Г. Венецкий, Г. С. Кильдишев. – 3-е изд., перераб. и доп. – Москва : Статистика, 1975. – 264 с.

6. Высшая математика : методические указания и контрольные задания для студентов заочной формы обучения. В 4 ч. / В. С. Денисов [и др.]. – Витебск : УО «ВГТУ», 2006. – Ч. 3–4.

7. Высшая математика. Случайные величины в теории вероятностей : методические указания к практическим занятиям для студентов второго курса экономических специальностей / А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2015. – 116 с.

8. Высшая математика. Теория вероятностей и математическая статистика : методические указания к практическим занятиям для студентов дневной и заочной форм обучения / Д. Е. Дунина [и др.]. – Витебск : УО «ВГТУ», 2012. – 128 с.

9. Высшая математика. Теория вероятностей : методические указания к практическим занятиям для студентов дневной и заочной форм обучения / Д. Е. Дунина [и др.]. – Витебск : УО «ВГТУ», 2009. – 102 с.

10. Высшая математика. Числовые и функциональные ряды. Случайные события в теории вероятностей : методические указания к практическим занятиям для студентов второго курса экономических специальностей / А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2014. – 101 с.

11. Высшая математика. Элементы математической статистики и линейного программирования : методические указания к практическим занятиям для студентов второго курса экономических специальностей / А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2016. – 109 с.

12. Герасимович, А. И. Математическая статистика / А. И. Герасимович. – Минск : Выш. шк, 1983. – 280 с.

13. Гмурман, В. Е. Теория вероятностей и математическая статистика : учеб. пособие для втузов / В. Е. Гмурман. – 5-е изд., перераб. и доп. – Москва : Высшая школа, 1977. – 479 с.

14. Гринберг, А. С. Теория вероятностей и математическая статистика : курс лекций / А. С. Гринберг, О. Б. Плющ, Б. В. Новыш; Академия управления при Пре-

зиденте Республики Беларусь. – 3-е изд., доп. – Минск : Академия управления при Президенте Республики Беларусь, 2005. – 186 с.

15. Гуринович, С. Л. Математика. Задачи с экономическим содержанием / С. Л. Гуринович. – Минск : Новое знание, 2008. – 263 с.

16. Гурский, Е. И. Сборник задач по теории вероятностей и математической статистике / Е. И. Гурский. – Минск : Выш. шк, 1984. – 223 с.

17. Гусак, А. А. Теория вероятностей : справочное пособие к решению задач / А. А. Гусак, Е. А. Бричикова. – Минск : ТетраСистемс, 2007. – 288 с.

18. Данко, П. Е. Высшая математика в упражнениях и задачах. В 2 ч. Ч. 2 / П. Е. Данко, А. Г. Попов. – Москва : Высш. школа, 1967. – 350 с.

19. Ефимов, А. В. Сборник задач по математике для втузов. Специальные курсы / А. В. Ефимов [и др]. – Москва : Наука, 1984. – 606 с.

20. Жевняк, Р. М. Высшая математика. В 5 ч. Ч. 5 / Р. М. Жевняк, А. А. Карпук. – Минск : Выш. шк., 1988. – 254 с.

21. Жевняк, Р. М. Теория вероятностей и математическая статистика : учеб. пособие для вузов / Р. М. Жевняк, А. А. Карпук, В. Т. Унукович. – Минск : Харвест, 2000. – 384 с.

22. Индивидуальные задания по высшей математике: Операционное исчисление. Элементы теории устойчивости. Теория вероятностей. Математическая статистика / А. П. Рябушко [и др.]. – Минск : Выш. шк., 2006. – 336 с.

23. Индивидуальные задания по теории вероятностей и математической статистике : типовые расчёты / Ю. В. Муранов [и др.]. – Витебск : УО «ВГТУ», 2000. – 66 с.

24. Карасёв, А. И. Курс высшей математики для экономических вузов. В 2 ч. Ч. 2 / А. И. Карасёв, З. М. Аксютина, Т. И. Савельева. – Москва : Высш. школа, 1990. – 272 с.

25. Карасёв, А. И. Теория вероятностей и математическая статистика / А. И. Карасёв. – Москва : Статистика, 1979. – 279 с.

26. Колемаев, В. А. Теория вероятностей и математическая статистика / В. А. Колемаев [и др.]. – Москва : Высш. школа, 1991. – 400 с.

27. Колемаев, В. А. Теория вероятностей и математическая статистика : учебник / В. А. Колемаев, В. Н. Калинина; под ред. В. А. Колемаева. – Москва : ИНФРА-М, 1997. – 302 с.

28. Корячко, В. П. Интеллектуальные системы и нечеткая логика : учебник для магистров высших учебных заведений / В. П. Корячко, М. А. Бакулева, В. И. Орешков. – Москва : КУРС, 2017. – 347 с.

29. Кудрявцев, В. А. Краткий курс высшей математики / В. А. Кудрявцев, Б. П. Демидович. – Москва : Наука, 1989. – 655 с.

30. Кузнецов, А. В. Руководство к решению задач по математическому программированию / А. В. Кузнецов, Н. И. Холод, Л. С. Костевич. – Минск : Выш. школа, 1978. – 256 с.

31. Кузнецов, Ю. Н. Математическое программирование / Ю. Н. Кузнецов, В. И. Кузубов, А. Б. Волощенко. – Москва : Высш. школа, 1980. – 300 с.

32. Маталыцкий, М. А. Теория вероятностей и математическая статистика / М. А. Маталыцкий, Г. А. Хацкевич ; Министерство образования Республики Беларусь. – Минск : Вышэйшая школа, 2017. – 591 с.

33. Математика. Аналитическая геометрия. Линейная и векторная алгебра : методические указания к практическим занятиям для студентов первого курса / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2017. – 103 с.

34. Математика. Дифференциальные уравнения и системы дифференциальных уравнений. Операционное исчисление : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2019. – 101 с.

35. Математика. Интегральное исчисление функции одной переменной. Дифференциальное исчисление функции нескольких переменных : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2018. – 98 с.

36. Математика. Кратные интегралы. Элементы теории поля. Ряды : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2021. – 108 с.

37. Математика. Линейные операторы векторных пространств. Введение в математический анализ. Дифференциальное исчисление функции одной переменной : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2017. – 120 с.

38. Математика. Случайные события в вероятностном пространстве : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2021. – 106 с.

39. Математика. Случайные величины в вероятностном пространстве : практикум / сост. А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2022. – 106 с.

40. Прохоров, Ю. В. Теория вероятностей / Ю. В. Прохоров, Ю. А. Розанов. – Москва : Наука, 1973. – 496 с.

41. Руководство к решению задач по высшей математике. В 2 ч. Ч. 2 / Е. И. Гурский [и др.]. – Минск : Выш. шк., 1989. – 400 с.

42. Теория вероятностей и математическая статистика : для спец. 1-й ступени высшего образования, закрепленных за УМО, в качестве учебно-методического пособия / А. И. Волковец [и др.]. – Минск : БГУИР, 2017. – 71 с.

43. Теория вероятностей и математическая статистика : задания для выполнения типовых расчётов для студентов второго курса механико-технологических специальностей / А. В. Коваленко [и др.]. – Витебск : УО «ВГТУ», 2012. – 77 с.

44. Теория вероятностей и математическая статистика [Электронный ресурс]: электронный учебно-методический комплекс / науч. Рук. Ю. С. Харин. – Минск : УО «БГУ», 2010.

45. Чистяков, В. П. Курс теории вероятностей / В. П. Чистяков. – Москва : Наука, 1982. – 243 с.

46. Чудесенко, В. Ф. Сборник заданий по специальным курсам высшей математики (Типовые расчёты) / В. Ф. Чудесенко. – Москва : Высш. школа, 1983. – 111 с.

ПРИЛОЖЕНИЕ А

Таблица А.1 – Значения функции распределения Пуассона

$$P\left(X=m\right)=\frac{\mu^{m}}{m!}\cdot e^{-\mu}$$

При μ, равном 0,1; 0,2; …; 1,0

m \ μ	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
0	0,90484	0,81873	0,74082	0,67032	0,60653	0,54881	0,49659	0,44933	0,40657	0,36788
1	0,09048	0,16375	0,22225	0,26813	0,30327	0,32929	0,34761	0,35946	0,36591	0,36788
2	0,00452	0,01637	0,03334	0,05363	0,07582	0,09879	0,12166	0,14379	0,16466	0,18394
3	0,00015	0,00109	0,00333	0,00715	0,01264	0,01976	0,02839	0,03834	0,04940	0,06131
4	0,00000	0,00005	0,00025	0,00072	0,00158	0,00296	0,00497	0,00767	0,01111	0,01533
5		0,00000	0,00002	0,00006	0,00016	0,00036	0,00070	0,00123	0,00200	0,00307
6			0,00000	0,00000	0,00001	0,00004	0,00008	0,00016	0,00030	0,00051
7					0,00000	0,00000	0,00001	0,00002	0,00004	0,00007
8							0,00000	0,00000	0,00000	0,00001
9										0,00000

При μ, равном 2; 3; 4; …; 11

m \ μ	2	3	4	5	6	7	8	9	10	11
0	0,13534	0,04979	0,01832	0,00674	0,00248	0,00091	0,00034	0,00012	0,00005	0,00002
1	0,27067	0,14936	0,07326	0,03369	0,01487	0,00638	0,00268	0,00111	0,00045	0,00018
2	0,27067	0,22404	0,14653	0,08422	0,04462	0,02234	0,01073	0,00500	0,00227	0,00101
3	0,18045	0,22404	0,19537	0,14037	0,08924	0,05213	0,02863	0,01499	0,00757	0,00370
4	0,09022	0,16803	0,19537	0,17547	0,13385	0,09123	0,05725	0,03374	0,01892	0,01019
5	0,03609	0,10082	0,15629	0,17547	0,16062	0,12772	0,09160	0,06073	0,03783	0,02242
6	0,01203	0,05041	0,10420	0,14622	0,16062	0,14900	0,12214	0,09109	0,06306	0,04109
7	0,00344	0,02160	0,05954	0,10444	0,13768	0,14900	0,13959	0,11712	0,09008	0,06458
8	0,00086	0,00810	0,02977	0,06528	0,10326	0,13038	0,13959	0,13176	0,11260	0,08879
9	0,00019	0,00270	0,01323	0,03627	0,06884	0,10140	0,12408	0,13176	0,12511	0,10853
10	0,00004	0,00081	0,00529	0,01813	0,04130	0,07098	0,09926	0,11858	0,12511	0,11938
11	0,00001	0,00022	0,00192	0,00824	0,02253	0,04517	0,07219	0,09702	0,11374	0,11938
12	0,00000	0,00006	0,00064	0,00343	0,01126	0,02635	0,04813	0,07277	0,09478	0,10943
13		0,00001	0,00020	0,00132	0,00520	0,01419	0,02962	0,05038	0,07291	0,09259
14		0,00000	0,00006	0,00047	0,00223	0,00709	0,01692	0,03238	0,05208	0,07275
15			0,00002	0,00016	0,00089	0,00331	0,00903	0,01943	0,03472	0,05335
16			0,00000	0,00005	0,00033	0,00145	0,00451	0,01093	0,02170	0,03668
17				0,00001	0,00012	0,00060	0,00212	0,00579	0,01276	0,02373

m \ μ	2	3	4	5	6	7	8	9	10	11
18				0,00000	0,00004	0,00023	0,00094	0,00289	0,00709	0,01450
19					0,00001	0,00009	0,00040	0,00137	0,00373	0,00840
20					0,00000	0,00003	0,00016	0,00062	0,00187	0,00462
21						0,00001	0,00006	0,00026	0,00089	0,00242
22						0,00000	0,00002	0,00011	0,00040	0,00121
23							0,00001	0,00004	0,00018	0,00058
24							0,00000	0,00002	0,00007	0,00027
25								0,00001	0,00003	0,00012
26								0,00000	0,00001	0,00005
27									0,00000	0,00002

Таблица А.2 – Значения функции

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

	0	1	2	3	4	5	6	7	8	9
0,0	0,39894	0,39892	0,39886	0,39876	0,39862	0,39844	0,39822	0,39797	0,39767	0,39733
0,1	0,39695	0,39654	0,39608	0,39559	0,39505	0,39448	0,39387	0,39322	0,39253	0,39181
0,2	0,39104	0,39024	0,38940	0,38853	0,38762	0,38667	0,38568	0,38466	0,38361	0,38251
0,3	0,38139	0,38023	0,37903	0,37780	0,37654	0,37524	0,37391	0,37255	0,37115	0,36973
0,4	0,36827	0,36678	0,36526	0,36371	0,36213	0,36053	0,35889	0,35723	0,35553	0,35381
0,5	0,35207	0,35029	0,34849	0,34667	0,34482	0,34294	0,34105	0,33912	0,33718	0,33521
0,6	0,33322	0,33121	0,32918	0,32713	0,32506	0,32297	0,32086	0,31874	0,31659	0,31443
0,7	0,31225	0,31006	0,30785	0,30563	0,30339	0,30114	0,29887	0,29659	0,29431	0,29200
0,8	0,28969	0,28737	0,28504	0,28269	0,28034	0,27798	0,27562	0,27324	0,27086	0,26848
0,9	0,26609	0,26369	0,26129	0,25888	0,25647	0,25406	0,25164	0,24923	0,24681	0,24439
1,0	0,24197	0,23955	0,23713	0,23471	0,23230	0,22988	0,22747	0,22506	0,22265	0,22025
1,1	0,21785	0,21546	0,21307	0,21069	0,20831	0,20594	0,20357	0,20121	0,19886	0,19652
1,2	0,19419	0,19186	0,18954	0,18724	0,18494	0,18265	0,18037	0,17810	0,17585	0,17360
1,3	0,17137	0,16915	0,16694	0,16474	0,16256	0,16038	0,15822	0,15608	0,15395	0,15183
1,4	0,14973	0,14764	0,14556	0,14350	0,14146	0,13943	0,13742	0,13542	0,13344	0,13147
1,5	0,12952	0,12758	0,12566	0,12376	0,12188	0,12001	0,11816	0,11632	0,11450	0,11270
1,6	0,11092	0,10915	0,10741	0,10567	0,10396	0,10226	0,10059	0,09893	0,09728	0,09566
1,7	0,09405	0,09246	0,09089	0,08933	0,08780	0,08628	0,08478	0,08329	0,08183	0,08038
1,8	0,07895	0,07754	0,07614	0,07477	0,07341	0,07206	0,07074	0,06943	0,06814	0,06687
1,9	0,06562	0,06438	0,06316	0,06195	0,06077	0,05959	0,05844	0,05730	0,05618	0,05508

	0	1	2	3	4	5	6	7	8	9
2,0	0,05399	0,05292	0,05186	0,05082	0,04980	0,04879	0,04780	0,04682	0,04586	0,04491
2,1	0,04398	0,04307	0,04217	0,04128	0,04041	0,03955	0,03871	0,03788	0,03706	0,03626
2,2	0,03547	0,03470	0,03394	0,03319	0,03246	0,03174	0,03103	0,03034	0,02965	0,02898
2,3	0,02833	0,02768	0,02705	0,02643	0,02582	0,02522	0,02463	0,02406	0,02349	0,02294
2,4	0,02239	0,02186	0,02134	0,02083	0,02033	0,01984	0,01936	0,01888	0,01842	0,01797
2,5	0,01753	0,01709	0,01667	0,01625	0,01585	0,01545	0,01506	0,01468	0,01431	0,01394
2,6	0,01358	0,01323	0,01289	0,01256	0,01223	0,01191	0,01160	0,01130	0,01100	0,01071
2,7	0,01042	0,01014	0,00987	0,00961	0,00935	0,00909	0,00885	0,00861	0,00837	0,00814
2,8	0,00792	0,00770	0,00748	0,00727	0,00707	0,00687	0,00668	0,00649	0,00631	0,00613
2,9	0,00595	0,00578	0,00562	0,00545	0,00530	0,00514	0,00499	0,00485	0,00470	0,00457
3,0	0,00443	0,00430	0,00417	0,00405	0,00393	0,00381	0,00370	0,00358	0,00348	0,00337
3,1	0,00327	0,00317	0,00307	0,00298	0,00288	0,00279	0,00271	0,00262	0,00254	0,00246
3,2	0,00238	0,00231	0,00224	0,00216	0,00210	0,00203	0,00196	0,00190	0,00184	0,00178
3,3	0,00172	0,00167	0,00161	0,00156	0,00151	0,00146	0,00141	0,00136	0,00132	0,00127
3,4	0,00123	0,00119	0,00115	0,00111	0,00107	0,00104	0,00100	0,00097	0,00094	0,00090
3,5	0,00087	0,00084	0,00081	0,00079	0,00076	0,00073	0,00071	0,00068	0,00066	0,00063
3,6	0,00061	0,00059	0,00057	0,00055	0,00053	0,00051	0,00049	0,00047	0,00046	0,00044
3,7	0,00042	0,00041	0,00039	0,00038	0,00037	0,00035	0,00034	0,00033	0,00031	0,00030
3,8	0,00029	0,00028	0,00027	0,00026	0,00025	0,00024	0,00023	0,00022	0,00021	0,00021
3,9	0,00020	0,00019	0,00018	0,00018	0,00017	0,00016	0,00016	0,00015	0,00014	0,00014
4,0	0,00013	0,00013	0,00012	0,00012	0,00011	0,00011	0,00011	0,00010	0,00010	0,00009
4,1	0,00009	0,00009	0,00008	0,00008	0,00008	0,00007	0,00007	0,00007	0,00006	0,00006
4,2	0,00006	0,00006	0,00005	0,00005	0,00005	0,00005	0,00005	0,00004	0,00004	0,00004
4,3	0,00004	0,00004	0,00004	0,00003	0,00003	0,00003	0,00003	0,00003	0,00003	0,00003
4,4	0,00002	0,00002	0,00002	0,00002	0,00002	0,00002	0,00002	0,00002	0,00002	0,00002
4,5	0,00002	0,00002	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001
4,6	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001
4,7	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00000	0,00000	0,00000	0,00000

Таблица А.3 – Значения функции Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,00000	0,37	0,14431	0,73	0,26730	1,10	0,36433
0,01	0,00399	0,38	0,14803	0,74	0,27035	1,11	0,36650
0,02	0,00798	0,39	0,15173	0,75	0,27337	1,12	0,36864
0,03	0,01197	0,40	0,15542	0,76	0,27637	1,13	0,37076
0,04	0,01595	0,41	0,15910	0,77	0,27935	1,14	0,37286
0,05	0,01994	0,42	0,16276	0,78	0,28230	1,15	0,37493
0,06	0,02392	0,43	0,16640	0,79	0,28524	1,16	0,37698
0,07	0,02790	0,44	0,17003	0,80	0,28814	1,17	0,37900
0,08	0,03188	0,45	0,17364	0,81	0,29103	1,18	0,38100
0,09	0,03586	0,46	0,17724	0,82	0,29389	1,19	0,38298
0,10	0,03983	0,47	0,18082	0,83	0,29673	1,20	0,38493
0,11	0,04380	0,48	0,18439	0,84	0,29955	1,21	0,38686
0,12	0,04776	0,49	0,18793	0,85	0,30234	1,22	0,38877
0,13	0,05172	0,50	0,19146	0,86	0,30511	1,23	0,39065
0,14	0,05567	0,51	0,19497	0,87	0,30785	1,24	0,39251
0,15	0,05962	0,52	0,19847	0,88	0,31057	1,25	0,39435
0,16	0,06356	0,53	0,20194	0,89	0,31327	1,26	0,39617
0,17	0,06749	0,54	0,20540	0,90	0,31594	1,27	0,39796
0,18	0,07142	0,55	0,20884	0,91	0,31859	1,28	0,39973
0,19	0,07535	0,56	0,21226	0,92	0,32121	1,29	0,40147
0,20	0,07926	0,57	0,21566	0,93	0,32381	1,30	0,40320
0,21	0,08317	0,58	0,21904	0,94	0,32639	1,31	0,40490
0,22	0,08706	0,59	0,22240	0,95	0,32894	1,32	0,40658
0,23	0,09095	0,60	0,22575	0,96	0,33147	1,33	0,40824
0,24	0,09483	0,61	0,22907	0,97	0,33398	1,34	0,40988
0,25	0,09871	0,62	0,23237	0,98	0,33646	1,35	0,41149
0,26	0,10257	0,63	0,23565	0,99	0,33891	1,36	0,41308
0,27	0,10642	0,64	0,23891	1,00	0,34134	1,37	0,41466
0,28	0,11026	0,65	0,24215	1,01	0,34375	1,38	0,41621
0,29	0,11409	0,66	0,24537	1,02	0,34614	1,39	0,41774
0,30	0,11791	0,67	0,24857	1,03	0,34849	1,40	0,41924
0,31	0,12172	0,675	0,25016	1,04	0,35083	1,41	0,42073
0,32	0,12552	0,68	0,25175	1,05	0,35314	1,42	0,42220
0,33	0,12930	0,69	0,25490	1,06	0,35543	1,43	0,42364
0,34	0,13307	0,70	0,25804	1,07	0,35769	1,44	0,42507
0,35	0,13683	0,71	0,26115	1,08	0,35993	1,45	0,42647
0,36	0,14058	0,72	0,26424	1,09	0,36214	1,46	0,42785

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,47	0,42922	1,87	0,46926	2,27	0,48840	2,67	0,49621
1,48	0,43056	1,88	0,46995	2,28	0,48870	2,68	0,49632
1,49	0,43189	1,89	0,47062	2,29	0,48899	2,69	0,49643
1,50	0,43319	1,90	0,47128	2,30	0,48928	2,70	0,49653
1,51	0,43448	1,91	0,47193	2,31	0,48956	2,71	0,49664
1,52	0,43574	1,92	0,47257	2,32	0,48983	2,72	0,49674
1,53	0,43699	1,93	0,47320	2,33	0,49010	2,73	0,49683
1,54	0,43822	1,94	0,47381	2,34	0,49036	2,74	0,49693
1,55	0,43943	1,95	0,47441	2,35	0,49061	2,75	0,49702
1,56	0,44062	1,96	0,47500	2,36	0,49086	2,76	0,49711
1,57	0,44179	1,97	0,47558	2,37	0,49111	2,77	0,49720
1,58	0,44295	1,98	0,47615	2,38	0,49134	2,78	0,49728
1,59	0,44408	1,99	0,47670	2,39	0,49158	2,79	0,49736
1,60	0,44520	2,00	0,47725	2,40	0,49180	2,80	0,49744
1,61	0,44630	2,01	0,47778	2,41	0,49202	2,81	0,49752
1,62	0,44738	2,02	0,47831	2,42	0,49224	2,82	0,49760
1,63	0,44845	2,03	0,47882	2,43	0,49245	2,83	0,49767
1,64	0,44950	2,04	0,47932	2,44	0,49266	2,84	0,49774
1,65	0,45053	2,05	0,47982	2,45	0,49286	2,85	0,49781
1,66	0,45154	2,06	0,48030	2,46	0,49305	2,86	0,49788
1,67	0,45254	2,07	0,48077	2,47	0,49324	2,87	0,49795
1,68	0,45352	2,08	0,48124	2,48	0,49343	2,88	0,49801
1,69	0,45449	2,09	0,48169	2,49	0,49361	2,89	0,49807
1,70	0,45543	2,10	0,48214	2,50	0,49379	2,90	0,49813
1,71	0,45637	2,11	0,48257	2,51	0,49396	2,91	0,49819
1,72	0,45728	2,12	0,48300	2,52	0,49413	2,92	0,49825
1,73	0,45818	2,13	0,48341	2,53	0,49430	2,93	0,49831
1,74	0,45907	2,14	0,48382	2,54	0,49446	2,94	0,49836
1,75	0,45994	2,15	0,48422	2,55	0,49461	2,95	0,49841
1,76	0,46080	2,16	0,48461	2,56	0,49477	2,96	0,49846
1,77	0,46164	2,17	0,48500	2,57	0,49492	2,97	0,49851
1,78	0,46246	2,18	0,48537	2,58	0,49506	2,98	0,49856
1,79	0,46327	2,19	0,48574	2,59	0,49520	2,99	0,49891
1,80	0,46407	2,20	0,48610	2,60	0,49534	3,00	0,49865
1,81	0,46485	2,21	0,48645	2,61	0,49547	3,01	0,49869
1,82	0,46562	2,22	0,48679	2,62	0,49560	3,02	0,49874
1,83	0,46638	2,23	0,48713	2,63	0,49573	3,03	0,49878
1,84	0,46712	2,24	0,48745	2,64	0,49585	3,04	0,49882
1,85	0,46784	2,25	0,48778	2,65	0,49598	3,05	0,49886
1,86	0,46856	2,26	0,48809	2,66	0,49609	3,06	0,49889

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
3,07	0,49893	3,13	0,49913	3,19	0,49929	3,70	0,49989
3,08	0,49896	3,14	0,49916	3,20	0,49931	3,80	0,49993
3,09	0,49900	3,15	0,49918	3,30	0,49952	3,90	0,49995
3,10	0,49903	3,16	0,49921	3,40	0,49966	4,00	0,499968
3,11	0,49906	3,17	0,49924	3,50	0,49977	4,50	0,499997
3,12	0,49910	3,18	0,49926	3,60	0,49984	5,00	0,499999

Таблица А.4 – Значение χ^2 распределения

В таблице представлены значения $\chi^2_{\alpha,\nu}$ в зависимости от числа степеней свободы ν и вероятности α

ν / α	0,2	0,1	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13,815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19,669	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,783	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23,542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820
20	25,038	28,412	31,410	35,020	37,556	45,315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,553	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,556	44,314	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	38,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

Таблица А.5 – Распределение Стьюдента

Значения $t_{\alpha,\nu}$ удовлетворяют условию $P\left(t \geq t_{\alpha,\nu}\right) = \int\limits_{t_{\alpha,\nu}}^{\infty} S(t,\nu)\,dt = \alpha$.

В таблице представлены значения квантилей $t_{\alpha,\nu}$ в зависимости от числа степеней свободы ν и вероятности α.

ν \ α	0,40	0,30	0,20	0,10	0,050	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	381,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	5,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,765	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	6,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	3,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,562	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,254	0,527	0,848	1,293	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,526	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Таблица А.6 – Распределение Фишера

В таблице приведены критические значения (квантили) F_{α,ν_1,ν_2} в зависимости от числа степеней свободы ν_1 и ν_2 для значения $\alpha = 0,05$:

$$P\left(F \geq F_{\alpha,\nu_1,\nu_2}\right) = 0,05 .$$

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,512	18,999	19,163	19,248	19,298	19,329	19,371	19,414	19,453	19,496
3	10,129	9,552	9,276	9,118	9,014	8,941	8,844	8,774	8,638	8,527
4	7,710	6,945	6,591	6,388	6,257	6,164	6,041	5,912	5,774	5,628
5	6,607	5,785	5,410	5,192	5,050	4,950	4,818	4,678	4,527	4,365
6	5,987	5,143	4,756	4,534	4,388	4,284	4,147	4,000	3,841	3,669
7	5,591	4,737	4,347	4,121	3,972	3,866	3,725	3,574	3,410	3,230
8	5,317	4,459	4,067	3,838	3,688	3,580	3,438	3,284	3,116	2,928
9	5,117	4,256	3,863	3,633	3,482	3,374	3,230	3,073	2,900	2,707
10	4,965	4,103	3,708	3,478	3,326	3,217	3,072	2,913	2,737	2,538
11	4,844	3,982	3,587	3,357	3,204	3,094	2,948	2,788	2,609	2,405
12	4,747	3,885	3,490	3,259	3,106	2,999	2,848	2,686	2,505	2,296
13	4,667	3,805	3,410	3,179	3,025	2,915	2,767	2,604	2,420	2,207
14	4,600	3,739	3,344	3,112	2,958	2,848	2,699	2,534	2,349	2,131
15	4,543	3,683	3,287	3,056	2,901	2,790	2,641	2,475	2,288	2,066
16	4,494	3,634	3,239	3,007	2,853	2,741	2,591	2,424	2,235	2,010
17	4,451	3,592	3,197	2,965	2,810	2,699	2,548	2,381	2,190	1,961
18	4,414	3,555	3,160	2,928	2,773	2,661	2,510	2,342	2,150	1,917
19	4,381	3,522	3,127	2,895	2,740	2,629	2,477	2,308	2,114	1,878
20	4,351	3,493	3,098	2,866	2,711	2,599	2,447	2,278	2,083	1,843
21	4,325	3,467	3,072	2,840	2,685	2,573	2,421	2,250	2,064	1,812
22	4,301	3,443	3,049	2,817	2,661	2,549	2,397	2,226	2,028	1,783
23	4,279	3,422	3,028	2,795	2,640	2,528	2,375	2,203	2,005	1,757
24	4,260	3,403	3,009	2,777	2,621	2,508	2,355	2,183	1,984	1,733
25	4,242	3,385	2,991	2,759	2,603	2,490	2,337	2,165	1,965	1,711
26	4,225	3,369	2,975	2,743	2,587	2,474	2,321	2,148	1,947	1,691
27	4,210	3,354	2,961	2,728	2,572	2,459	2,305	2,132	1,930	1,672
28	4,196	3,340	2,947	2,714	2,558	2,445	2,992	2,118	1,915	1,654
29	4,183	3,328	2,934	2,702	2,545	2,432	2,278	2,104	1,901	1,638
30	4,171	3,316	2,922	2,690	2,534	2,421	2,266	2,092	1,887	1,622
40	4,085	3,232	2,839	2,605	2,449	2,336	2,180	2,004	1,793	1,509
60	4,001	3,151	2,758	2,525	2,368	2,254	2,097	1,918	1,700	1,389
120	3,920	3,072	2,680	2,447	2,290	2,175	2,106	1,834	1,608	1,254
∞	3,841	2,996	2,605	2,372	2,214	2,098	1,938	1,752	1,517	1,000

Таблица А.7 – Доверительные интервалы для σ

P / ν	0,99		0,98		0,95		0,90	
	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2
1	0,356	15,9	0,388	9,98	0,446	9,31	0,510	5,19
2	0,434	14,1	0,466	9,97	0,521	6,28	0,578	4,40
3	0,483	6,47	0,514	5,11	0,566	3,73	0,620	2,32
4	0,519	4,39	0,549	3,67	0,599	2,87	0,649	2,37
5	0,546	3,48	0,576	3,00	0,624	2,45	0,672	2,09
6	0,569	2,98	0,587	2,62	0,644	2,202	0,690	1,916
7	0,588	2,66	0,616	2,377	0,661	2,035	0,705	1,797
8	0,604	2,44	0,631	2,205	0,675	1,916	0,718	1,711
9	0,618	2,277	0,644	2,076	0,688	1,826	0,729	1,645
10	0,630	2,154	0,656	1,977	0,699	1,755	0,739	1,593
11	0,641	2,056	0,667	1,898	0,708	1,698	0,748	1,550
12	0,651	1,976	0,677	1,833	0,717	1,651	0,755	1,515
13	0,660	1,910	0,685	1,779	0,725	1,611	0,762	1,485
14	0,669	1,854	0,693	1,733	0,732	1,577	0,769	1,460
15	0,676	1,806	0,700	1,694	0,739	1,548	0,775	1,437
16	0,683	1,764	0,707	1,659	0,745	1,522	0,780	1,418
17	0,690	1,727	0,713	1,629	0,750	1,499	0,785	1,400
18	0,696	1,695	0,719	1,602	0,756	1,479	0,790	1,385
19	0,702	1,666	0,725	1,578	0,760	1,460	0,794	1,370
20	0,707	1,640	0,730	1,556	0,765	1,444	0,798	1,358
21	0,712	1,617	0,734	1,536	0,769	1,429	0,802	1,346
22	0,717	1,595	0,739	1,519	0,773	1,416	0,805	1,335
23	0,722	1,576	0,743	1,502	0,777	1,402	0,809	1,326
24	0,726	1,558	0,747	1,487	0,781	1,391	0,812	1,316
25	0,730	1,541	0,751	1,473	0,784	1,380	0,815	1,308
26	0,734	1,526	0,755	1,460	0,788	1,371	0,818	1,300
27	0,737	1,512	0,758	1,448	0,791	1,361	0,820	1,293
28	0,741	1,499	0,762	1,436	0,794	1,352	0,823	1,286
29	0,744	1,487	0,765	1,426	0,796	1,344	0,825	1,279
30	0,748	1,475	0,768	1,417	0,799	1,337	0,828	1,274
40	0,774	1,390	0,792	1,344	0,821	1,279	0,847	1,228
50	0,793	1,336	0,810	1,297	0,837	1,243	0,861	1,199
60	0,808	1,299	0,824	1,265	0,849	1,217	0,871	1,179
70	0,820	1,272	0,835	1,241	0,858	1,198	0,879	1,163
80	0,829	1,250	0,844	1,222	0,866	1,183	0,886	1,151
90	0,838	1,233	0,852	1,207	0,873	1,171	0,892	1,141
100	0,845	1,219	0,858	1,195	0,878	1,161	0,897	1,133
200	0,877	1,15	0,897	1,13	0,912	1,11	0,925	1,09

Таблица А.8 – Таблица производных основных элементарных функций

1)	$C' = 0$, где $C = Const$;	2)	$(x^n)' = n \cdot x^{n-1}$
3)	$(a^x)' = a^x \cdot \ln a$, где $a > 0$, $a \neq 1$;	4)	$(e^x)' = e^x$
5)	$(\log_a x)' = \dfrac{1}{x \cdot \ln a}$, где $a > 0$, $a \neq 1$, $x > 0$;	6)	$(\ln x)' = \dfrac{1}{x}$, где $x > 0$;
7)	$(\sin x)' = \cos x$;	8)	$(\cos x)' = -\sin x$;
9)	$(\operatorname{tg} x)' = \dfrac{1}{\cos^2 x}$, где $x \neq \dfrac{\pi}{2} + \pi k$, $k \in Z$;	10)	$(\operatorname{ctg} x)' = -\dfrac{1}{\sin^2 x}$, где $x \neq \pi k$, $k \in Z$;
11)	$(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$, где $\lvert x \rvert < 1$;	12)	$(\arccos x)' = -\dfrac{1}{\sqrt{1-x^2}}$, где $\lvert x \rvert < 1$;
13)	$(\operatorname{arctg} x)' = \dfrac{1}{1+x^2}$;	14)	$(\operatorname{arcctg} x)' = -\dfrac{1}{1+x^2}$;
15)	$(\operatorname{sh} x)' = \operatorname{ch} x$;	16)	$(\operatorname{ch} x)' = \operatorname{sh} x$;
17)	$(\operatorname{th} x)' = \dfrac{1}{\operatorname{ch}^2 x}$;	18)	$(\operatorname{cth} x)' = -\dfrac{1}{\operatorname{sh}^2 x}$, где $x \neq 0$.

Таблица А.9 – Таблица основных неопределённых интегралов

1)	$\int x^n dx = \dfrac{x^{n+1}}{n+1} + C$, $n \neq -1$;	2)	$\int \dfrac{dx}{x} = \ln\lvert x \rvert + C$;
3)	$\int a^x dx = \dfrac{a^x}{\ln a} + C$, $a > 0$, $a \neq 1, x > 0$;	4)	$\int e^x dx = e^x + C$;
5)	$\int \sin x\, dx = -\cos x + C$;	6)	$\int \cos x\, dx = \sin x + C$;
7)	$\int \dfrac{dx}{\cos^2 x} = \operatorname{tg} x + C$;	8)	$\int \dfrac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$;
9)	$\int \dfrac{dx}{\sin x} = \ln\left\lvert \operatorname{tg}\dfrac{x}{2} \right\rvert + C$;	10)	$\int \dfrac{dx}{\cos x} = \ln\left\lvert \operatorname{tg}\left(\dfrac{x}{2} + \dfrac{\pi}{4}\right) \right\rvert + C$;
11)	$\int \dfrac{dx}{a^2 + x^2} = \dfrac{1}{a}\operatorname{arctg}\dfrac{x}{a} + C\ a \neq 0$;	12)	$\int \dfrac{dx}{\sqrt{a^2 - x^2}} = \arcsin\dfrac{x}{a} + C, \lvert x \rvert < \lvert a \rvert$;
13)	$\int \dfrac{dx}{a^2 - x^2} = \dfrac{1}{2a}\ln\left\lvert \dfrac{x+a}{x-a} \right\rvert + C$;	14)	$\int \dfrac{dx}{\sqrt{x^2 \pm a^2}} = \ln\left\lvert x + \sqrt{x^2 \pm a^2} \right\rvert + C$;
15)	$\int \operatorname{sh} x\, dx = \operatorname{ch} x + C$;	16)	$\int \operatorname{ch} x\, dx = \operatorname{sh} x + C$;
17)	$\int \dfrac{dx}{\operatorname{ch}^2 x} = \operatorname{th} x + C$;	18)	$\int \dfrac{dx}{\operatorname{sh}^2 x} = -\operatorname{cth} x + C$, где $x \neq 0$.

Учебное издание

МАТЕМАТИКА.
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Практикум

Составители:
Коваленко Александр Вильямович
Джежора Александр Александрович
Дмитриев Александр Петрович
Завацкий Юрий Александрович
Вардомацкая Елена Юрьевна

Редактор *Т.А. Осипова*
Корректор *А.В. Пухальская*
Компьютерная верстка *А.В. Коваленко*

Подписано к печати 09.01.2023. Формат 60x90 1/16. Усл. печ. листов 5,9.
Уч.-изд. листов 7,8. Тираж 99 экз. Заказ № 1.

Учреждение образования «Витебский государственный технологический университет»
210038, г. Витебск, Московский пр., 72.
Отпечатано на ризографе учреждения образования
«Витебский государственный технологический университет».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 1/172 от 12 февраля 2014 г.
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 3/1497 от 30 мая 2017 г.