МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования

«Омский государственный технический университет»

Факультет информационных технологий и компьютерных систем Кафедра «Прикладная математика и фундаментальная информатика»

Расчётно-графическая работа

|  |  |
| --- | --- |
| по дисциплине | Теория вероятностей и  математическая статистика |

|  |  |
| --- | --- |
| Студента |  |
|  | фамилия, имя, отчество полностью |
| Курс | 2 Группа ФИТ-231 |
| Направление | 02.03.02 Фундаментальная информатика |
|  | и информационные технологии |
|  | код, наименование |
| Руководитель | доц., канд. техн. наук, доцент |
|  | должность, ученая степень, звание |
|  | Болдовская Т.Е. |
|  | фамилия, инициалы |
| Выполнил |  |
|  | дата, подпись студента |
| баллы |  |
|  | дата, подпись руководителя |

Омск-2025

Оглавление

[Введение 3](#_Toc196342527)

[Выборка 4](#_Toc196342528)

[Интервальный статистический ряд 4](#_Toc196342529)

[Гипотеза о распределении 7](#_Toc196342530)

[Проверка правила трёх сигм 8](#_Toc196342531)

[Проверка критерием Пирсона 9](#_Toc196342532)

[Доверительные интервалы 10](#_Toc196342533)

# Введение

Современная статистика представляет собой мощный инструментарий для анализа случайных величин, позволяющий на основе ограниченной выборки делать выводы о свойствах всей генеральной совокупности. Одной из фундаментальных задач математической статистики является проверка гипотез о законе распределения случайной величины, что имеет важное значение для последующего моделирования и прогнозирования.

В данной лабораторной работе проводится комплексное исследование выборочных данных, включающее:

* Первичную обработку и систематизацию данных через построение вариационного и интервального рядов
* Визуализацию распределения с использованием гистограммы, полигона частот и эмпирической функции распределения
* Расчет основных числовых характеристик (моментов распределения)
* Выдвижение и проверку гипотезы о виде теоретического распределения
* Оценку параметров предполагаемого распределения
* Проверку адекватности модели с помощью критерия согласия Пирсона
* Построение доверительных интервалов для параметров распределения

# Выборка

В соответствии с выданным преподавателем вариантом в данной работе будут производиться расчёты над данной выборкой:

18,5; 19,0; 20,6; 18,4; 18,3; 19,2 18,5; 20,3; 18,6; 20,6; 18,5; 18,3; 19,1; 21,0; 18,6; 19,1; 18,4; 19,9; 18,7; 19,5; 18,4; 21,6; 18,9; 19,5; 20,1; 19,4; 19,0; 19,2; 19,7; 19,9; 19,5; 20,0; 18,4; 18,3; 19,6; 18,8; 23,1; 19,6; 18,5; 20,7; 18,7; 18,7; 22,8; 18,9; 20,2; 19,0; 19,2; 19,6; 18,9; 20,3; 21,0; 18,9; 20,3; 18,3; 19,5; 18,5;18,5; 18,6; 19,3; 18,6

Выборка в отсортированном виде представляет собой вариационный ряд:

18.3, 18.3, 18.3, 18.3, 18.4, 18.4, 18.4, 18.4, 18.5, 18.5, 18.5, 18.5, 18.5, 18.5, 18.6, 18.6, 18.6, 18.6, 18.7, 18.7, 18.7, 18.8, 18.9, 18.9, 18.9, 18.9, 19.0, 19.0, 19.0, 19.1, 19.1, 19.2, 19.2, 19.2, 19.3, 19.4, 19.5, 19.5, 19.5, 19.5, 19.6, 19.6, 19.6, 19.7, 19.9, 19.9, 20.0, 20.1, 20.2, 20.3, 20.3, 20.3, 20.6, 20.6, 20.7, 21.0, 21.0, 21.6, 22.8, 23.1

# Интервальный статистический ряд

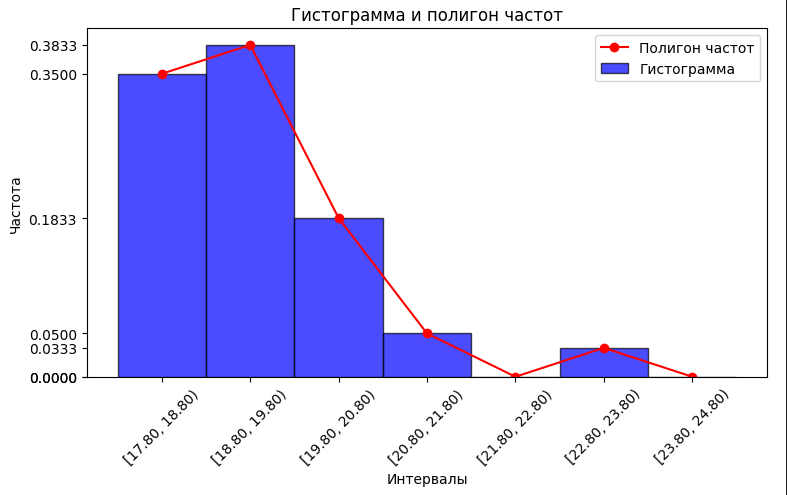
Для расчёта воспользуемся формулой Стёрджесса:

Где h – длина частичного интервала, R – размах выборки, m – число интервалов, n – объём выборки.

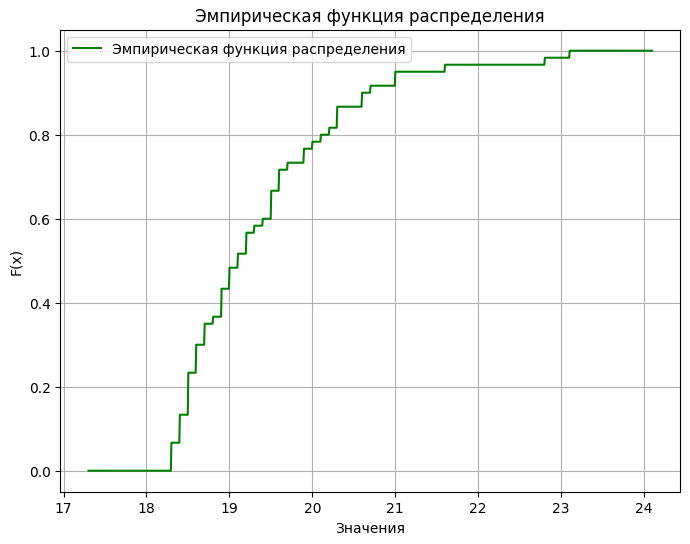
По данной формуле оптимальное количество интервалов – 7 c шагом 1. Рассчитаем начальную точку первого интервала по формуле:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Интервал | Фактические частоты | Относительные частоты |
| 1 | [17.80, 18.80) | 21 | 0.35 |
| 2 | [18.80, 19.80) | 23 | 0.383 |
| 3 | [19.80, 20.80) | 11 | 0.183 |
| 4 | [20.80, 21.80) | 3 | 0.05 |
| 5 | [21.80, 22.80) | 0 | 0 |
| 6 | [22.80, 23.80) | 2 | 0.03 |
| 7 | [23.80, 24.80) | 0 | 0 |

**Полигон и гистограмма относительных частот**

****

**График эмпирической функции распределения**

****

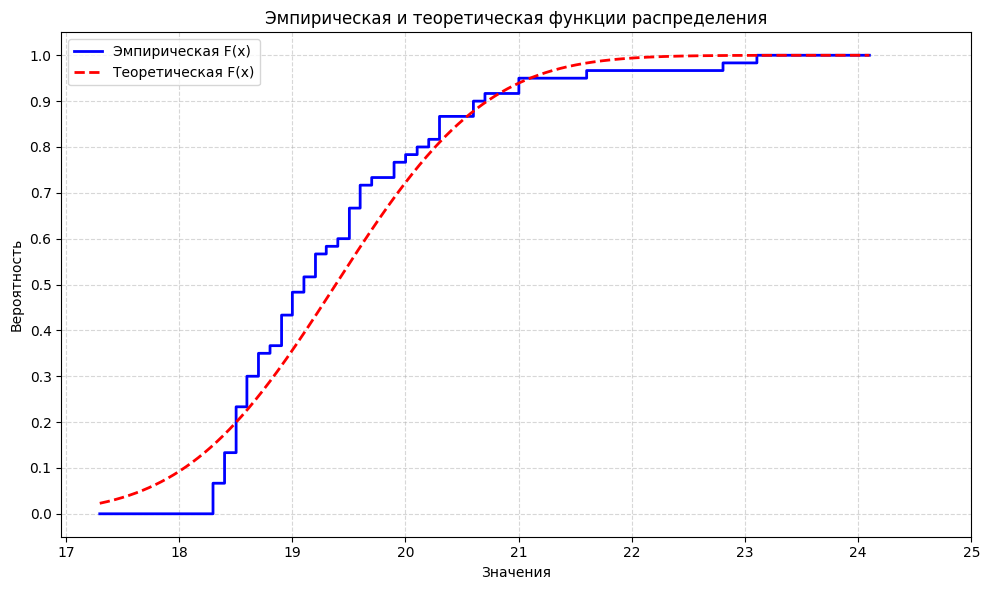
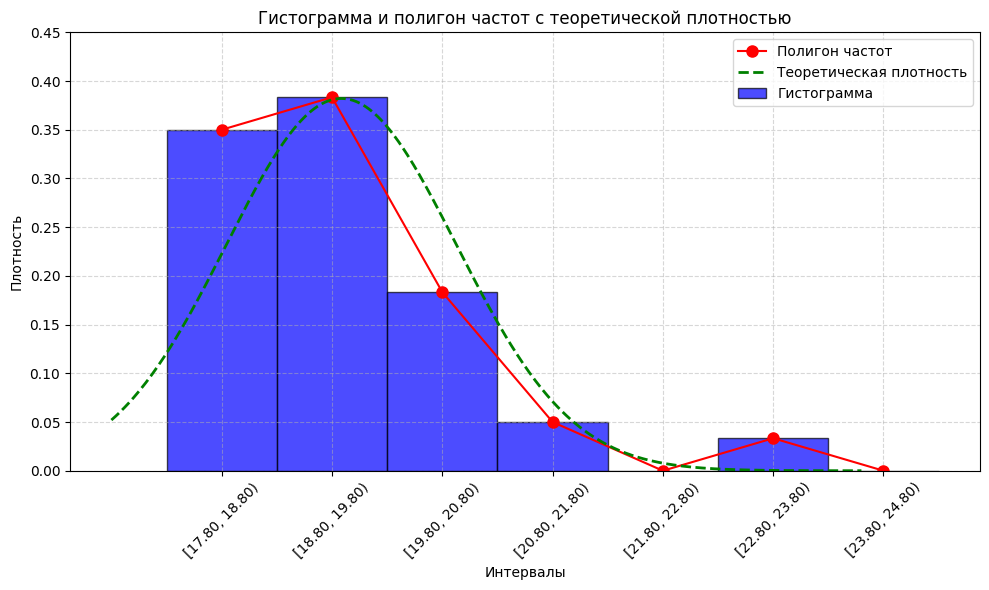
**Числовые характеристики**

В этом разделе будут рассчитаны основные числовые характеристики выборки.

|  |  |  |
| --- | --- | --- |
| Характеристика | Формула | Значение |
| Выборочное среднее |  | 19.38 |
| Исправленная выборочная дисперсия |  | 1.09 |
| Мода |  | 18.94 |
| Медиана |  | 19.19 |
| Эксцесс |  | 3 |
| Коэффициент асимметрии |  | 1.59 |

# Гипотеза о распределении

На основе анализа гистограммы и эмпирической функции распределения предполагаем, что генеральная совокупность имеет нормальное распределение с параметрами, близкими к выборочному среднему ≈ 19.385 и исправленной дисперсии ≈ 1.090. Однако положительные значения асимметрии 1.589 и эксцесса 3.003 указывают на отклонение от симметрии, что может свидетельствовать о смещенном нормальном распределении или логарифмически нормальном распределении.



# Проверка правила трёх сигм

Из расчетов выше, среднее выборочное равняется 19.38, сигма равна . Тогда диапазон для проверки равняется [16.25, 22.52]. В интервал попадает 58 значений из 60. Правило выполняется приближённо

# Проверка критерием Пирсона

Для теста возьмём интервальный статистический ряд. Рассчитаем теоретические частоты при помощи формулы

Тогда:

|  |  |  |
| --- | --- | --- |
| Интервал |  |  |
| [17.80, 18.80) |  | 13.45016 |
| [18.80, 19.80) |  | 22.01869 |
| [19.80, 20.80) |  | 15.41009 |
| [20.80, 21.80) |  | 4.60085 |
| [21.80, 22.80) |  | 0.58212 |
| [22.80, 23.80) |  | 0.03094 |
| [23.80, 24.80) |  | 0.00068 |

И тогда теоретические частоты будут равны [13.45016, 22.01869, 15.41009, 4.60085, 0.58212, 0.03094, 0.00068]. Фактические же наблюдаемые частоты - [21, 23, 11, 3, 0, 2, 0].

Теоретические частоты должны быть больше 5, поэтому объединим 4, 5 и 6 интервалы:

|  |  |  |
| --- | --- | --- |
| Интервал |  |  |
| [17.80, 18.80) | 21 | 13.45016 |
| [18.80, 19.80) | 23 | 22.01869 |
| [19.80, 20.80) | 11 | 15.41009 |
| [20.80, 23.80) | 5 | 5.21459 |

Рассчитаем статистики хи-квадрат:

Определим число степеней свободы .

Для уровня значимости и : .

# Доверительные интервалы

Для генерального среднего:

Степень свободы: 59

Критическое значение при

Теперь для генерального среднеквадратичного отклонения

Нижняя граница:

Верхняя граница:

Доверительный интервал:

# Заключение

Проведенное статистическое исследование позволило получить полную характеристику исследуемой выборки и сделать обоснованные выводы о законе ее распределения. На первом этапе были систематизированы исходные данные через построение вариационного и интервального рядов, что дало наглядное представление о структуре выборки.

Графические методы анализа (гистограмма, полигон частот, эмпирическая функция распределения) в сочетании с расчетом числовых характеристик (среднее значение 19.385, дисперсия 1.090, коэффициент асимметрии 1.589) позволили выдвинуть гипотезу о возможном нормальном распределении генеральной совокупности. Однако последующая проверка с помощью критерия согласия Пирсона (χ² = 5.55 при критическом значении 3.841) привела к отвержению этой гипотезы на 5%-ном уровне значимости.

Дополнительный анализ показал, что:

* Правило "трех сигм" выполняется приближенно
* Доверительные интервалы для параметров распределения составили:
  + Для математического ожидания: []
  + Для среднеквадратического отклонения: [0.878; 1.260]

Полученные результаты свидетельствуют о том, что, несмотря на некоторое сходство с нормальным распределением, исследуемая выборка имеет статистически значимые отклонения от нормального закона, что, вероятно, связано с правосторонней асимметрией распределения.