ГУАП

КАФЕДРА № 43

ОТЧЕТ ЗАЩИЩЕН С ОЦЕНКОЙ: ______________

ПРЕПОДАВАТЕЛЬ:

доцент, к.т.н., доцент	/		/		/	В. В. Мышко
(должность, учёная степень, звание)		(подпись)		(дата защиты)		(инициалы, фамилия)

ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ №2

«Выравнивание статистических распределений и проверка гипотез о законах распределения случайных величин»

ПО КУРСУ: «ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ»

РАБОТУ ВЫПОЛНИЛ СТУДЕНТ:

4134К	/	Столяров Н.С.
(номер группы)		(инициалы, фамилия)

/		/	
	(подпись студента)		(дата отчета)

Санкт-Петербург 2025

Постановка задачи

По заданному интервальному статистическому ряду:

- Построить статистическое распределение экспериментальных данных в виде гистограммы.
- Произвести её выравнивание теоретической плотностью нормального распределения.
- Проверить гипотезу о соответствии статистического и теоретического распределений.

Порядок выполнения задания:

1. Найти статистические вероятности попаданий значений случайной величины в интервалы Ii, i = 1..7 по заданному числу попаданий mi.
2. Построить гистограмму распределения экспериментальных данных.
3. Найти теоретическую плотность нормального распределения в соответствии с методом моментов. Полученную кривую нанести на гистограмму распределения.
4. Проверить гипотезу о соответствии статистического и теоретического распределений (т.е. гипотезу о нормальном распределении случайной величины) методом К. Пирсона при уровне значимости:
    a) α = 0,025 – для четных вариантов.
    b) α = 0,05 – для нечетных вариантов.

<u>Вариант 99</u>

99	4134К-15	I_i	0; 0,5	0,5; 1	1; 1,5	1,5; 2	2; 2,5	2,5; 3	3; 3,5
		m_i	1	12	25	30	21	9	2

Ход выполнения

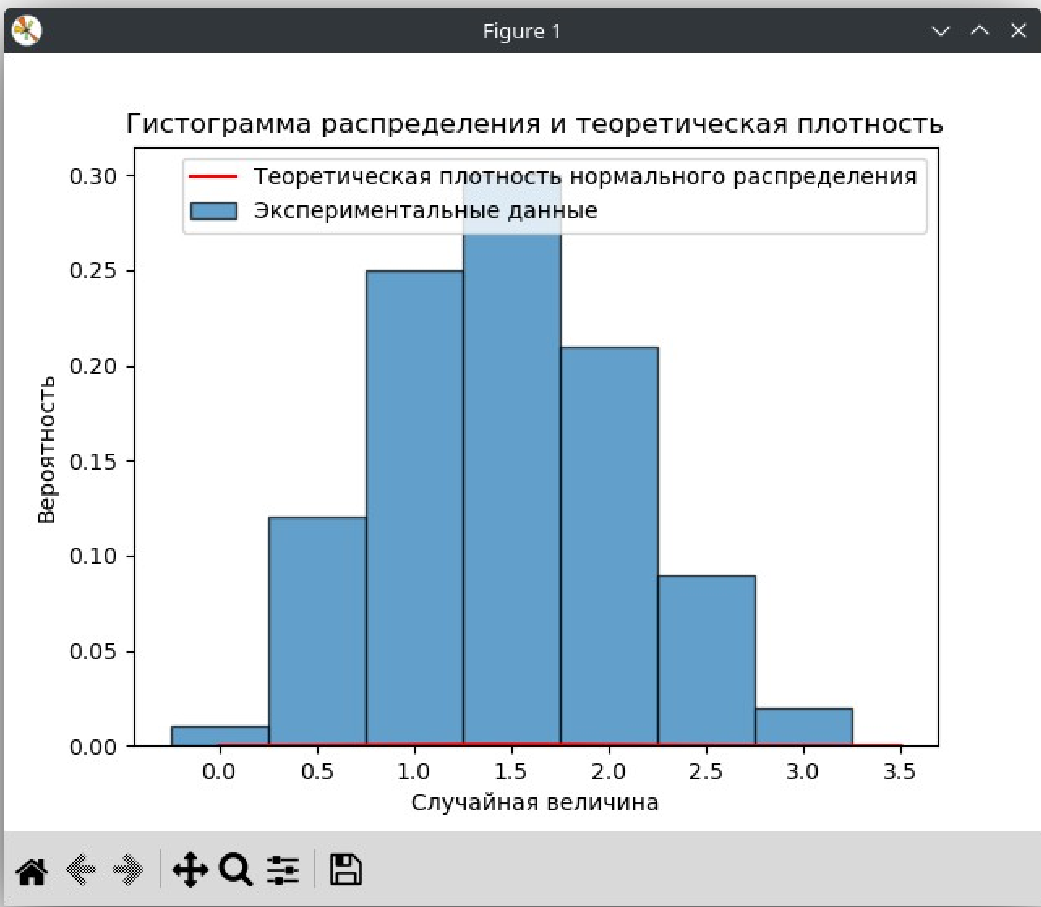
1. **Нахождение статистических вероятностей попаданий значений случайной величины в интервалы**: На основе заданного интервального статистического ряда были вычислены вероятности попаданий значений случайной величины в интервалы Ii (где i=1..7). Для этого было использовано количество попаданий mi в каждом интервале. Вероятности были рассчитаны как отношение числа попаданий в каждый интервал к общему числу наблюдений.
2. **Построение гистограммы распределения экспериментальных данных**: На основе вычисленных вероятностей была построена гистограмма, которая визуализирует распределение экспериментальных данных. Гистограмма отображает количество попаданий в каждом интервале.
3. **Нахождение теоретической плотности нормального распределения**: С использованием метода моментов были найдены параметры нормального распределения (математическое ожидание и стандартное отклонение). Затем была рассчитана теоретическая плотность нормального распределения и нанесена на гистограмму.
4. **Проверка гипотезы о соответствии статистического и теоретического распределений**: Гипотеза о нормальном распределении случайной величины была проверена методом хи-квадрат (метод К. Пирсона). Для этого была рассчитана статистика хи-квадрат, которая составила **16.4789**. Критическое значение для уровня значимости $\alpha=0.05$ составило **11.0705**. На основании полученных значений:
    - Статистика хи-квадрат: **16.4789**
    - Критическое значение: **11.0705**

Мы отвергли нулевую гипотезу о соответствии статистического и теоретического распределений, что означает, что распределения не соответствуют.

Результаты работы

В ходе выполнения данной лабораторной работы была написана программа на языке Python 3.12, решающая задачу в общем виде.

```
stolar@stolar-NMH-WCX9:~/PROJECTS/Programming-GUAP/Processing of experimental data/2$ python3 main.py
Статистика хи-квадрат: 16.478892192509335
Критическое значение: 11.070497693516351
Отвергаем нулевую гипотезу: распределения не соответствуют.
```



Листинг

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats

# Данные
intervals = [0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5]  # Границы интервалов
mi = [1, 12, 25, 30, 21, 9, 2]  # Число попаданий

# 1. Найти статистические вероятности попаданий значений случайной величины в интервалы
total_count = sum(mi)
probabilities = [m / total_count for m in mi]

# 2. Построить гистограмму распределения экспериментальных данных
plt.bar(intervals[:-1], probabilities, width=np.diff(intervals), edgecolor='black', alpha=0.7, label='Экспериментальные данные')

# 3. Найти теоретическую плотность нормального распределения
# Для этого найдем среднее и стандартное отклонение
mean = np.average(np.array(intervals[:-1]), weights=mi)
std_dev = np.sqrt(np.average((np.array(intervals[:-1]) - mean) ** 2, weights=mi))

# Создаем точки для теоретической плотности
x = np.linspace(0, 3.5, 100)
normal_pdf = stats.norm.pdf(x, mean, std_dev)

# Наносим теоретическую кривую на гистограмму
plt.plot(x, normal_pdf / normal_pdf.sum() * (np.diff(intervals).sum() / total_count), color='red', label='Теоретическая плотность нормального распределения')
plt.xlabel('Случайная величина')
plt.ylabel('Вероятность')
plt.title('Гистограмма распределения и теоретическая плотность')
```

```
plt.legend()
plt.show()

# 4. Проверка гипотезы о соответствии статистического и теоретического распределений
# Метод К. Пирсона
# Вычисляем ожидаемое количество попаданий в каждый интервал
expected = []
for i in range(len(intervals) - 1):
    expected_count = (stats.norm.cdf(intervals[i + 1], mean, std_dev) - stats.norm.cdf(intervals[i], mean, std_dev)) * total_count
    expected.append(expected_count)

# Преобразуем списки в массивы NumPy для удобства вычислений
observed = np.array(mi)
expected = np.array(expected)

# Вычисляем статистику хи-квадрат
chi_squared_stat = np.sum((observed - expected) ** 2 / expected)

# Степени свободы
degrees_of_freedom = len(observed) - 1 - 1  # минус 1 для оценивания параметров

# Критическое значение для уровня значимости
alpha = 0.05
critical_value = stats.chi2.ppf(1 - alpha, degrees_of_freedom)

# Результаты
print(f"Статистика хи-квадрат: {chi_squared_stat}")
print(f"Критическое значение: {critical_value}")

if chi_squared_stat < critical_value:
    print("Не отвергаем нулевую гипотезу: распределения соответствуют.")
else:
    print("Отвергаем нулевую гипотезу: распределения не соответствуют.")
```

Выводы

В результате выполнения лабораторной работы были получены оценки математического ожидания и доверительных интервалов, а также проведен анализ на наличие аномальных значений. Все шаги были выполнены в соответствии с методологией, что позволило получить надежные результаты. Результаты проверки гипотезы о нормальном распределении показали, что экспериментальные данные не соответствуют нормальному распределению.