|  |  |  |
| --- | --- | --- |
| 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 | / | Самарин Д. В. |
|  | (номер группы) |  | (инициалы, фамилия) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | / |  | / | 18.02.2025 |
|  |  | (подпись студента) |  | (дата отчета) |

Цель

Целью данной работы является построение статистического распределения экспериментальных данных в виде гистограммы, выравнивание её с помощью теоретической плотности нормального распределения, а также проверка гипотезы о соответствии экспериментального распределения нормальному закону распределения с использованием критерия хи-квадрат.

Задание на лабораторную работу

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 (таблица 2.1);

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:

а. α = 0,025 – для четных вариантов;

б. α = 0,05 – для нечетных вариантов.

Ход работы

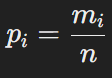
Вариант 96



В данном варианте уровень значимости выбран α=0.025.

Решение:

**Статистическая вероятность попадания в интервал:**

****

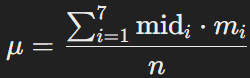
**Плотность вероятности для гистограммы (При длине интервала Δx):**



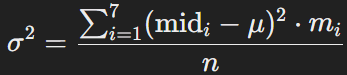
**Метод моментов для нормального распределения:**

Вычисление параметров:

Математическое ожидание:

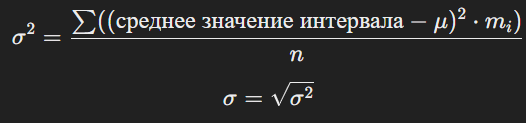


Дисперсия:



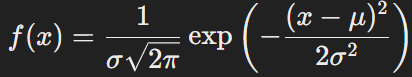
где midi​ – середина интервала Ii.

**Стандартное отклонение:**

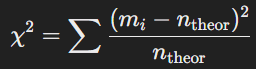


Стандартное отклонение (sigma) = 0.1398

**Теоретическая плотность нормального распределения:**

****

**Критерий хи-квадрат для проверки гипотезы о нормальности распределения:**



где piтеор = F(bi) − F(ai) – теоретическая вероятность попадания в интервал (ai, bi) при нормальном распределении с параметрами μ и σ. Степени свободы: df=k−1−r, где r – число оценённых параметров (для нормального распределения r=2).

**Результаты выполнения работы**

В ходе работы был реализован алгоритм

С работоспособностью алгоритма можно ознакомиться по ссылке ниже:

<https://seclar.streamlit.app/>

Шаг 1: Статистические вероятности

На основе предоставленных данных для каждого интервала были рассчитаны статистические вероятности попадания значений случайной величины в интервалы:

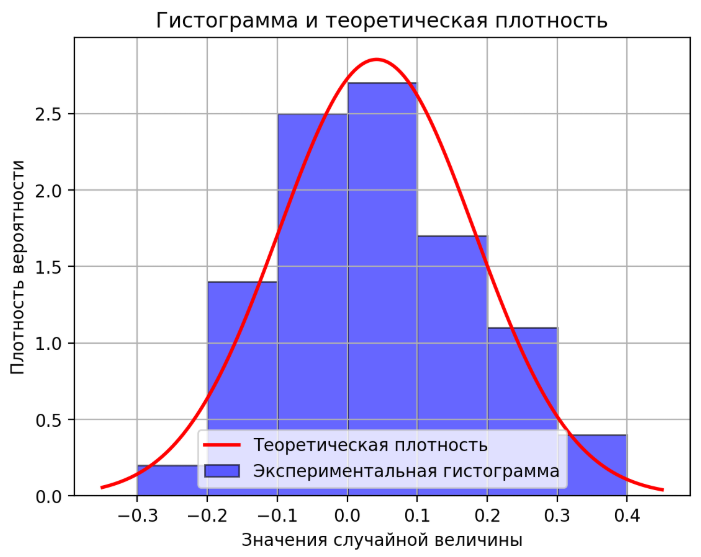
* Интервалы:

(−0.3,−0.2), (−0.2,−0.1), (−0.1,0), (0,0.1), (0.1,0.2), (0.2,0.3), (0.3,0.4)

* Частоты попаданий: 2,14,25,27,17,11,4
* Статистические вероятности: p1=0.040, p2=0.280, p3=0.500, p4=0.540, p5=0.340, p6=0.220, p7=0.080

Шаг 2: Построение гистограммы

Гистограмма экспериментальных данных была построена на основе рассчитанных вероятностей попадания в интервалы. Гистограмма отображает экспериментальные данные, сравнивая их с теоретической плотностью нормального распределения.

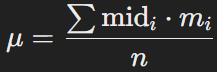


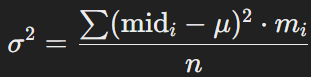
Шаг 3: Теоретическая плотность нормального распределения

Математическое ожидание (mu) и стандартное отклонение (sigma) были рассчитаны следующим образом:

* Математическое ожидание (mu): 0.0420
* Стандартное отклонение (sigma): 0.1398

Эти значения получены по формулам:





На интервале значений (от -0.3 до 0.4) с использованием функции norm.pdf построена теоретическая плотность нормального распределения с параметрами μ и σ. Кривая нанесена на гистограмму экспериментальных данных для сравнения.

**Шаг 4: Проверка гипотезы о нормальном распределении**

Для каждого интервала вычисляется вероятность попадания по нормальному закону:

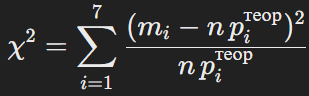


Затем теоретические частоты:



Вычисление статистики хи-квадрат:

Статистика:



При k=7 интервалах и r=2 оценённых параметрах степени свободы:

df=7−1−2=4.

Полученная статистика: χ2=3.0471.

При уровне значимости α=0.025 критическое значение хи-квадрат: 11.1433.

Поскольку статистика хи-квадрат меньше критического значения, гипотеза о нормальности распределения принимается.

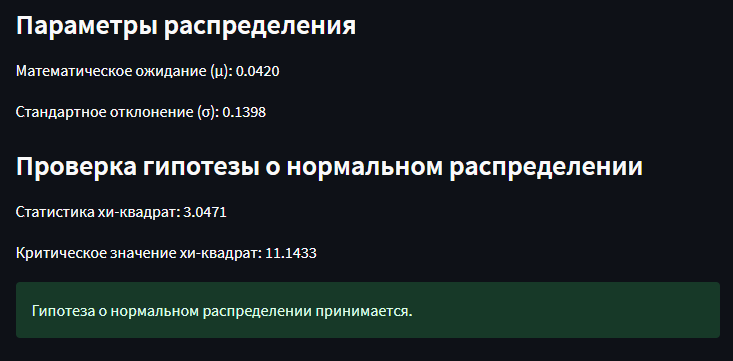
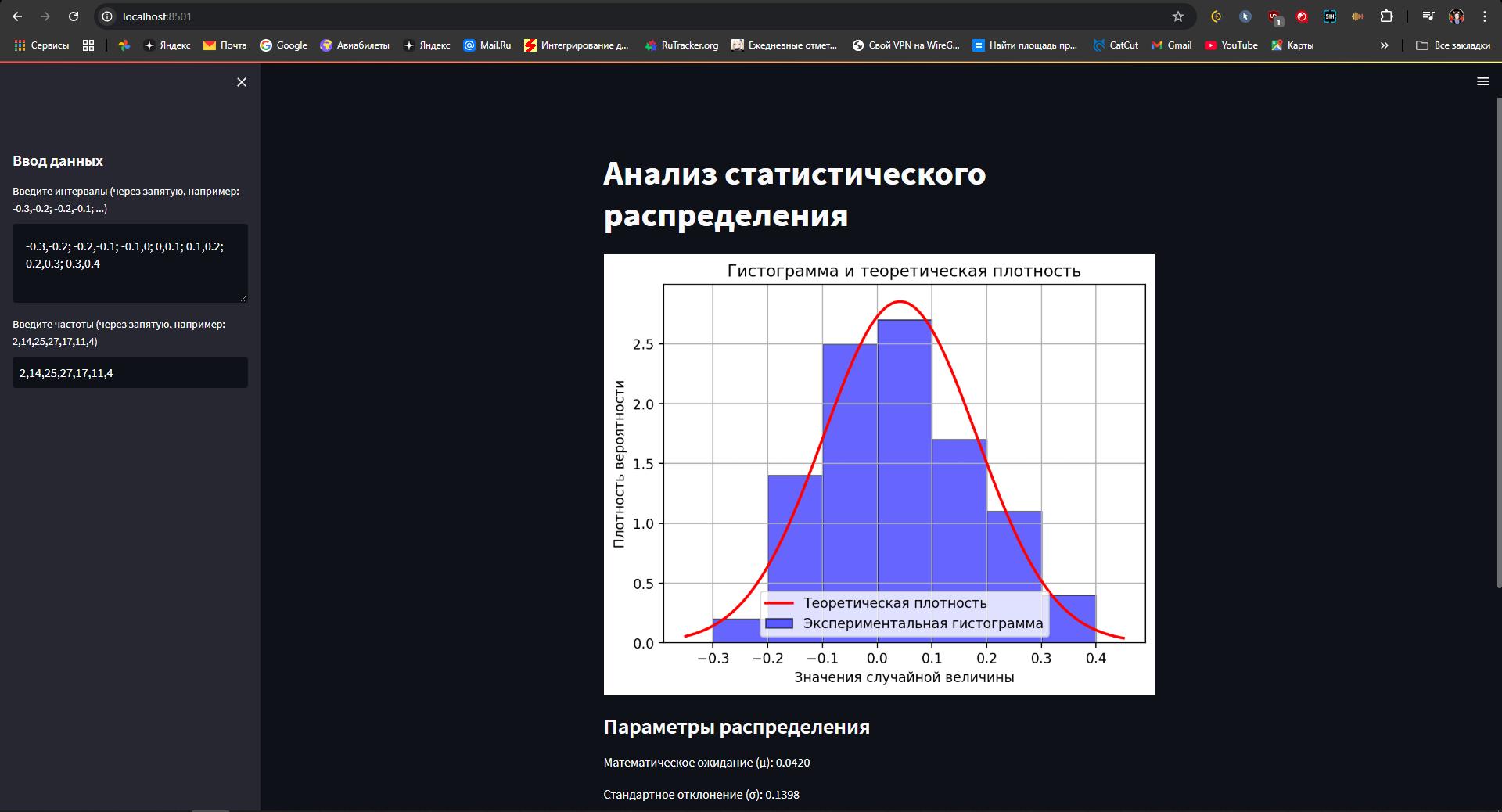
Полученные результаты в ходе вычислений:

* Математическое ожидание (mu): 0.0420
* Стандартное отклонение (sigma): 0.1398
* Статистика хи-квадрат: 3.0471
* Критическое значение хи-квадрат: 11.1433
* Гипотеза о нормальном распределении принимается.

Это означает, что экспериментальные данные, представленные интервальным статистическим рядом, хорошо описываются нормальным распределением с параметрами μ=0.0420 и σ=0.1398.

Также на графике построены экспериментальная гистограмма (синие столбики) и теоретическая плотность нормального распределения (красная кривая), что позволяет визуально оценить соответствие распределений.

Работа программы:



**Анализ полученных результатов:**

В результате анализа интервального статистического ряда экспериментальных данных была получена следующая оценка параметров нормального распределения методом моментов:

* Математическое ожидание (μ) = 0.0420,
* Стандартное отклонение (σ) = 0.1398.

Построенная гистограмма, полученная на основе экспериментальных частот, выровнена теоретической кривой плотности нормального распределения, что позволяет визуально оценить соответствие. Расчёт критерия хи-квадрат показал, что наблюдаемая статистика (χ2=3.0471) значительно меньше критического значения (11.1433) при уровне значимости α=0.025. Это позволяет принять гипотезу о нормальном распределении исследуемой случайной величины.

Таким образом, экспериментальные данные согласуются с теоретической моделью нормального распределения, что подтверждается как визуальным анализом (гистограмма и наложенная теоретическая кривая), так и формальной проверкой гипотезы с использованием критерия хи-квадрат.

1. Экспериментальные данные соответствуют нормальному распределению, что подтверждается результатами проверки гипотезы.
2. Гистограмма экспериментальных данных хорошо согласуется с теоретической плотностью нормального распределения.
3. Результаты проверки гипотезы с использованием критерия хи-квадрат подтверждают гипотезу о нормальности распределения случайной величины.

**Вывод:**

В ходе выполнения лабораторной работы была проведена проверка гипотезы о нормальном распределении случайной величины на основе экспериментальных данных. Для этого была использована методология выравнивания статистических распределений и тестирование гипотезы с помощью критерия хи-квадрат. Экспериментальные данные соответствуют нормальному распределению, что подтверждается проведенным тестом.

Листинг кода:

|  |
| --- |
| import numpy as np  import matplotlib.pyplot as plt  from scipy.stats import norm, chi2  import streamlit as st  # Заголовок приложения  st.title("Анализ статистического распределения")  # Ввод данных через интерфейс Streamlit  st.sidebar.header("Ввод данных")  intervals\_input = st.sidebar.text\_area(      "Введите интервалы (через запятую, например: -0.3,-0.2; -0.2,-0.1; ...)",      "-0.3,-0.2; -0.2,-0.1; -0.1,0; 0,0.1; 0.1,0.2; 0.2,0.3; 0.3,0.4"  )  m\_i\_input = st.sidebar.text\_input(      "Введите частоты (через запятую, например: 2,14,25,27,17,11,4)",      "2,14,25,27,17,11,4"  )  # Преобразование ввода в списки  try:      intervals = [tuple(map(float, pair.split(','))) for pair in intervals\_input.split(';')]      m\_i = list(map(int, m\_i\_input.split(',')))      n = sum(m\_i)  # Общее количество наблюдений  except Exception as e:      st.error(f"Ошибка ввода данных: {e}")      st.stop()  # Шаг 1: Статистические вероятности  p\_i = [m / n for m in m\_i]  # Шаг 2: Плотности вероятности для гистограммы  delta\_x = intervals[0][1] - intervals[0][0]  # Длина интервала  f\_i = [p / delta\_x for p in p\_i]  # Середины интервалов  midpoints = [(a + b) / 2 for a, b in intervals]  # Построение гистограммы  fig, ax = plt.subplots()  ax.bar(midpoints, f\_i, width=delta\_x, alpha=0.6, color='blue', edgecolor='black', label='Экспериментальная гистограмма')  # Шаг 3: Теоретическая плотность нормального распределения  mu = sum(mid \* m for mid, m in zip(midpoints, m\_i)) / n  sigma\_squared = sum(((mid - mu) \*\* 2) \* m for mid, m in zip(midpoints, m\_i)) / n  sigma = np.sqrt(sigma\_squared)  # Теоретическая плотность нормального распределения  x\_theory = np.linspace(min(midpoints) - delta\_x, max(midpoints) + delta\_x, 500)  f\_theory = norm.pdf(x\_theory, loc=mu, scale=sigma)  # Добавление теоретической кривой на график  ax.plot(x\_theory, f\_theory, color='red', linewidth=2, label='Теоретическая плотность')  # Настройка графика  ax.set\_title("Гистограмма и теоретическая плотность")  ax.set\_xlabel("Значения случайной величины")  ax.set\_ylabel("Плотность вероятности")  ax.legend()  ax.grid()  # Отображение графика в Streamlit  st.pyplot(fig)  # Вывод параметров распределения  st.subheader("Параметры распределения")  st.write(f"Математическое ожидание (μ): {mu:.4f}")  st.write(f"Стандартное отклонение (σ): {sigma:.4f}")  # Шаг 4: Проверка гипотезы о нормальном распределении (критерий хи-квадрат)  theoretical\_probabilities = [      norm.cdf(b, loc=mu, scale=sigma) - norm.cdf(a, loc=mu, scale=sigma) for a, b in intervals  ]  theoretical\_frequencies = [p \* n for p in theoretical\_probabilities]  # Статистика хи-квадрат  chi2\_stat = sum(((m - n\_theor) \*\* 2) / n\_theor for m, n\_theor in zip(m\_i, theoretical\_frequencies))  degrees\_of\_freedom = len(intervals) - 1 - 2  # k = n - 1 - r (r = 2 для нормального распределения)  alpha = 0.025  # Уровень значимости для четных вариантов  chi2\_critical = chi2.ppf(1 - alpha, degrees\_of\_freedom)  # Вывод результатов проверки гипотезы  st.subheader("Проверка гипотезы о нормальном распределении")  st.write(f"Статистика хи-квадрат: {chi2\_stat:.4f}")  st.write(f"Критическое значение хи-квадрат: {chi2\_critical:.4f}")  if chi2\_stat < chi2\_critical:      st.success("Гипотеза о нормальном распределении принимается.")  else:      st.error("Гипотеза о нормальном распределении отвергается.") |