МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ

ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ

УНИВЕРСИТЕТ им. В. Г. ШУХОВА» (БГТУ им. В.Г. Шухова)

Кафедра программного обеспечения

вычислительной техники и автоматизированных

систем

**Лабораторная работа №0**

по дисциплине: Анализ Данных

Выполнил: студент группы ВТ-231

Масленников Д. А.

Проверили:  
Кабалянц П. С.

Белгород 2025

**Лабораторная работа № 0**

**Задание**:1.Дискретная случайная величина ξ принимает значения 0, 1, 2 и 3 с вероятностями соответствующими данным варианта. Используя генератор псевдослучайных чисел, получить 10000 чисел, имитирующих поведение ξ. Вывести значения частот появления значений 0, 1, 2 и 3. Проверить на уровне значимости 0,05 соответствие полученных данных распределению ξ, используя критерий сравнения долей. Вероятности: р0=i/2\*(i+j+k), р1=(j+k)/2\*(i+j+k), р2=(i+k)/2\*(i+j+k), р3=j/2\*(i+j+k). Значения i, j, k являются последними тремя цифрами зачетки. Если какая-то из этих трех цифр равна 0, то вместо 0 берется 10.

2. Непрерывная случайная величина ξ распределена по экспоненциальному закону с параметром λ соответствующим данным варианта. Построив генератор псевдослучайных чисел, получить 10000 чисел, имитирующих поведение ξ. Разбить выборку на интервалы и построить гистограмму частот. Вывести средние арифметическое значений выборки и проверить  на уровне значимости 0,05 соответствие полученных данных показательному распределению с параметром λ. Здесь λ определяется номер по списку разделить на 4.

**Решение:**

**import numpy as np**

**import matplotlib.pyplot as plt**

**import scipy.stats as stats**

**# Заданные значения**

**i, j, k = 2, 10, 7**

**# Вычисляем вероятности для дискретной случайной величины**

**denominator = 2 \* (i + j + k)**

**p0 = i / denominator**

**p1 = (j + k) / denominator**

**p2 = (i + k) / denominator**

**p3 = j / denominator**

**# Генерация дискретной случайной величины**

**n\_samples = 10000**

**samples\_discrete = np.random.choice([0, 1, 2, 3], size=n\_samples, p=[p0, p1, p2, p3])**

**# Подсчёт частот**

**unique, counts = np.unique(samples\_discrete, return\_counts=True)**

**frequencies = counts / n\_samples**

**# Вывод результатов**

**print("Заданные вероятности:")**

**print(f"P(0) = {p0:.4f}, P(1) = {p1:.4f}, P(2) = {p2:.4f}, P(3) = {p3:.4f}")**

**print("\nСгенерированные частоты:")**

**for value, freq in zip(unique, frequencies):**

**print(f"Частота {value}: {freq:.4f}")**

**# Вторая часть: генерация экспоненциального распределения методом обратной функции**

**lambda\_param = 14 / 4**

**# Генератор экспоненциального распределения (НЕ ИСПОЛЬЗУЕМ random.expovariate)**

**uniform\_samples = np.random.uniform(0, 1, n\_samples)**

**samples\_exponential = -np.log(uniform\_samples) / lambda\_param**

**# Построение гистограммы**

**plt.hist(samples\_exponential, bins=30, density=True, alpha=0.6, color='b', edgecolor='black')**

**# Теоретическая плотность распределения**

**x = np.linspace(0, np.max(samples\_exponential), 1000)**

**plt.plot(x, lambda\_param \* np.exp(-lambda\_param \* x), 'r', label='Теоретическая плотность')**

**plt.title("Гистограмма и теоретическая плотность распределения")**

**plt.xlabel("Значение")**

**plt.ylabel("Плотность вероятности")**

**plt.legend()**

**plt.show()**

**# Среднее арифметическое**

**mean\_sample = np.mean(samples\_exponential)**

**# Проверка соответствия показательному распределению (К-С тест)**

**ks\_stat, ks\_p\_value = stats.kstest(samples\_exponential, 'expon', args=(0, 1/lambda\_param))**

**# Вывод результатов**

**print("\nСреднее значение выборки:", mean\_sample)**

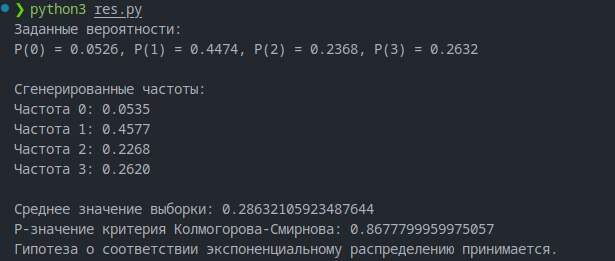
**print("P-значение критерия Колмогорова-Смирнова:", ks\_p\_value)**

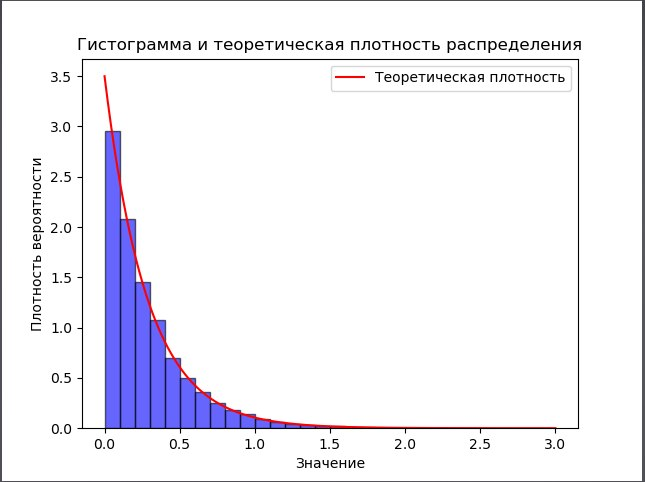
**if ks\_p\_value > 0.05:**

**print("Гипотеза о соответствии экспоненциальному распределению принимается.")**

**else:**

**print("Гипотеза о соответствии экспоненциальному распределению отвергается.")**

****

****

**Вывод:** Результаты анализа показали, что сгенерированные данные для дискретной случайной величины соответствуют заданному распределению (p > 0.05), а для экспоненциального распределения гипотеза о соответствии также не отвергается (p > 0.05). Таким образом, оба распределения успешно смоделированы и проверены на уровне значимости 0,05.