Министерство науки и высшего образования Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
ИТМО

Лабораторная работа №1

по дисциплине

«Статистика и анализ данных»

Семестр II

Выполнили:

Урядов Валерий Сергеевич
ИСУ: **467812**
Практический поток: **J3111**

Мокийчук Никита Артурович
ИСУ: **466748**
Практический поток: **J3111**

Кек Герман Вадимович
ИСУ: **466149**
Практический поток: **J3111**

Отчет сдан:
16.03.2025

Санкт-Петербург
2025

Ход выполнения работы

В процессе выполнения лабораторной работы были выполнены следующие этапы:

1. **Анализ исходных данных.** Определение экспериментальной области: квадрат со стороной $2a$ с центром в начале координат и круг с центром в начале координат и радиусом r. Выбраны 5 значений радиуса в интервале $(0, a]$.

2. **Расчёт истинной геометрической вероятности.** Для каждого значения r вычислена истинная вероятность попадания точки в круг по формуле
$$p = \frac{\pi r^2}{4},$$
где 4 — площадь квадрата.

3. **Генерация случайных точек.** С использованием генератора случайных чисел из библиотеки `numpy` были сгенерированы координаты точек, равномерно распределённых по квадрату.

4. **Проверка попадания в круг.** Функция `is_in_circle` определяет, принадлежит ли точка кругу.

5. **Построение графиков.** Для каждого значения r построены:
   - График сходимости оценки $\hat{p}(n)$ к истинной вероятности p,
   - График зависимости абсолютной ошибки $\varepsilon(n) = |\hat{p}(n) - p|$,
   - График зависимости необходимого числа точек $N(\varepsilon)$ для достижения заданной точности.

6. **Определение необходимого количества точек.** Функция `find_N_for_epsilon` вычисляет среднее значение N, при котором абсолютная ошибка становится не больше заданного ϵ.

Основная часть

В данном разделе описывается выполнение ключевых шагов и анализ промежуточных результатов.

1. Анализ исходных данных и расчёт истинной вероятности

Исходные данные задавались следующим образом:

- Квадрат с координатами от $-a$ до a, где $a = 1$.
- Круг с центром в начале координат и радиусом r, для которого истинная вероятность попадания точки определяется как
$$p = \frac{\pi r^2}{4}.$$

Выбранные значения радиуса (например, 1.0, 0.5, 0.33, 0.25, 0.2) позволяют оценить зависимость вероятности от размера круга.

2. Генерация случайных точек и проверка попадания в круг

Для генерации точек использовалась функция `numpy.random.Generator.uniform` с установленным зерном на основе текущего времени. Каждая точка проверялась на принадлежность кругу с помощью условия $x^2 + y^2 \leq r^2$. Кумулятивная сумма попаданий использовалась для вычисления оценки $\hat{p}(n)$ по мере увеличения числа точек.

3. Построение графиков и анализ ошибок

Были построены три графика для каждого значения r:

1. **Сходимость оценки $\hat{p}(n)$.** График демонстрирует, как с увеличением числа точек оценка приближается к истинной вероятности.

2. **Динамика абсолютной ошибки.** График зависимости ошибки $\varepsilon(n)$ от числа точек позволяет увидеть, как ошибка уменьшается с ростом выборки. Построение графика в логарифмическом масштабе подчёркивает закономерность сходимости.

3. **Зависимость $N(\varepsilon)$.** На этом графике показано, какое количество точек необходимо для достижения заданной точности (значения $\epsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$).

Полученные графики (см. приложенные изображения) демонстрируют ожидаемую сходимость оценки и снижение абсолютной ошибки по мере увеличения количества точек.

4. Промежуточные результаты

- При каждом выбранном значении r оценка $\hat{p}(n)$ сходится к теоретическому значению p.

- Графики динамики ошибки подтверждают, что с увеличением числа точек абсолютная ошибка стремится к нулю.

- Зависимость $N(\varepsilon)$ позволяет оценить необходимый объём выборки для достижения требуемой точности, что соответствует теоретическим ожиданиям.

Заключение

В результате выполнения лабораторной работы были достигнуты следующие выводы:

- Реализованный алгоритм корректно моделирует процесс определения геометрической вероятности методом Монте-Карло.

- Построенные графики сходимости оценки, динамики ошибки и зависимости $N(\varepsilon)$ демонстрируют, что при достаточном количестве случайных точек полученная оценка приближается к истинному значению вероятности.

- Экспериментальные результаты подтверждают теоретическую модель: с ростом числа точек уменьшается погрешность оценки.

- Применённые методы генерации случайных чисел и анализа данных оказались эффективными и позволяют оценивать параметры с требуемой точностью.