

Санкт-Петербургский Политехнический Университет Петра
Великого
Физико-механический институт

Высшая школа прикладной математики и вычислительной
физики

Лабораторная работа №1
по дисциплине “Математическая статистика”

Обучающаяся:

А.Д. Балакшина
(группа 5030102/20101)

Преподаватель:

А.Н. Баженов

Санкт-Петербург

2025

Содержание

1	Формулировка задания	3
1.1	Задача 1	3
1.2	Задача 2	3
2	Формализация	4
3	Выполнение работы	4
4	Результаты	5
4.1	Плотности вероятности и гистограммы	5
4.2	Характеристики распределений	7
5	Вывод	9

1 Формулировка задания

Для 4 распределений:

- Нормальное распределение $N(x, 0, 1)$
- Распределение Коши $C(x, 0, 1)$
- Распределение Пуассона $P(k, 10)$
- Равномерное распределение $U(x, -\sqrt{3}, \sqrt{3})$

1.1 Задача 1

Сгенерировать выборки размером 10, 50 и 1000 элементов. Построить на одном рисунке гистограмму и график плотности распределения.

1.2 Задача 2

Сгенерировать выборки размером 10, 100 и 1000 элементов. Для каждой выборки вычислить следующие статистические характеристики положения данных: \bar{x} , $medx$, z_Q . Повторить такие вычисления 1000 раз для каждой выборки и найти среднее характеристик положения и их квадратов. Вычислить оценку дисперсии. Представить полученные данные в виде таблиц.

2 Формализация

- Выборочное среднее:

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i \quad (1)$$

- Выборочная медиана:

$$medx = \begin{cases} z_{\frac{n+1}{2}} & n - \text{нечётное} \\ \frac{z_{\frac{n}{2}} + z_{\frac{n+2}{2}}}{2} & n - \text{чётное} \end{cases} \quad (2)$$

- Полусумма квартилей:

$$z_Q = \frac{z_{\frac{1}{4}} + z_{\frac{3}{4}}}{2} \quad (3)$$

- Среднее:

$$E(z) = \bar{z} \quad (4)$$

- Оценка дисперсии:

$$D(z) = \bar{z}^2 - \bar{z}^2 \quad (5)$$

3 Выполнение работы

Лабораторная работа выполнена на языке программирования Python 3.12 с использованием библиотек numpy, scipy, pandas, matplotlib.

Были сгенерированы выборки, построены графики и гистограммы (сохранялись в виде файлов png), оценены характеристики распределений (выводились в консоль в формате таблиц LATEX). Программа отработала корректно.

4 Результаты

4.1 Плотности вероятности и гистограммы

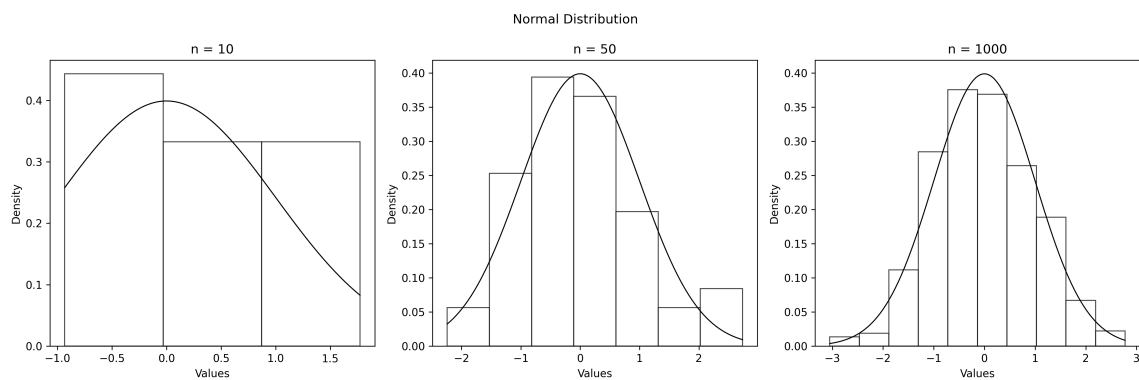


Рис. 1: Нормальное распределение.

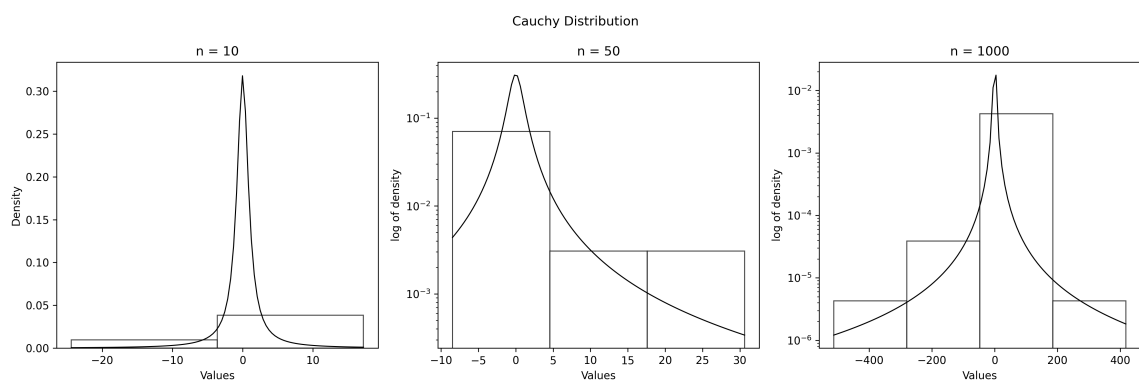


Рис. 2: Распределение Коши.

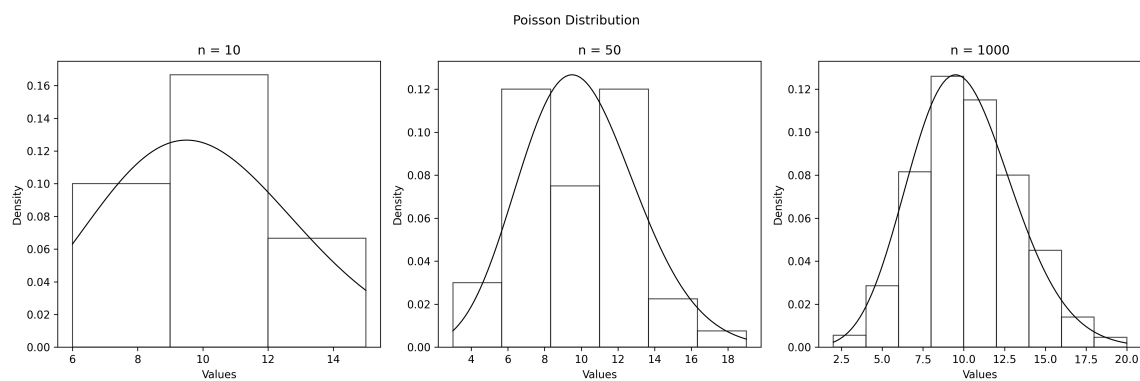


Рис. 3: Распределение Пуассона.

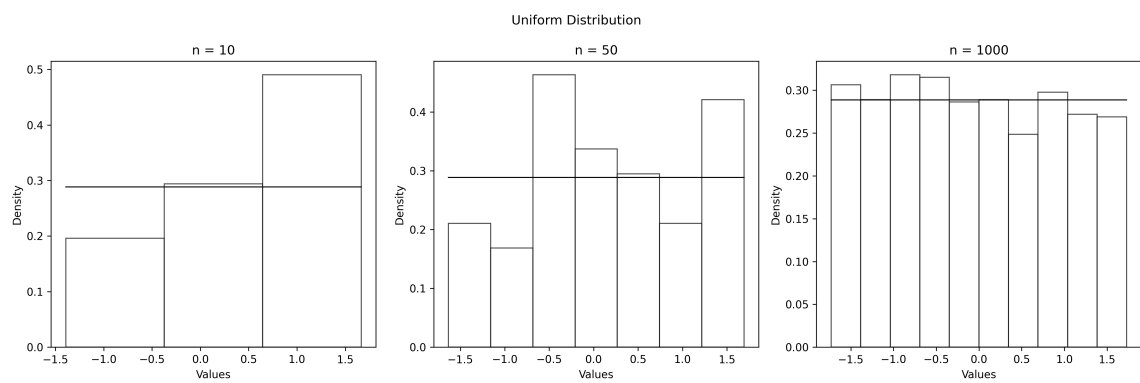


Рис. 4: Равномерное распределение.

4.2 Характеристики распределений

Normal distribution

Sample size 10

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	0.00013	0.0074	-0.0019
$D(z)^{(5)}$	0.1085	0.1437	0.1309

Sample size 100

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	-0.0028	0.0029	-0.018
$D(z)^{(5)}$	0.01005	0.01622	0.01255

Sample size 1000

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	9.9e-05	0.00055	-0.0012
$D(z)^{(5)}$	0.0009126	0.001500	0.001180

Cauchy distribution

Sample size 10

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	-0.39	-0.015	0.021
$D(z)^{(5)}$	414.0	0.3318	1.077

Sample size 100

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	1.6	0.0084	-0.033
$D(z)^{(5)}$	5911	0.02428	0.0540

Sample size 1000

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	-0.00087	-0.0019	-0.0069
$D(z)^{(5)}$	1534	0.002519	0.004812

Poisson distribution

Sample size 10

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	10	9.8	9.9
$D(z)^{(5)}$	1.024	1.541	1.241

Sample size 100

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	10	9.8	9.9
$D(z)^{(5)}$	0.09076	0.2092	0.1575

Sample size 1000

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	10	10	10
$D(z)^{(5)}$	0.009762	0.005964	0.003220

Uniform distribution

Sample size 10

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	0.00048	-0.0045	0.0073
$D(z)^{(5)}$	0.09636	0.2190	0.1309

Sample size 100

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	9.7e-05	0.0020	-0.015
$D(z)^{(5)}$	0.009923	0.02848	0.01523

Sample size 1000

	$x^{(1)}$	med $x^{(2)}$	$z_Q^{(3)}$
$E(z)^{(4)}$	0.00072	0.00058	-0.00069
$D(z)^{(5)}$	0.0009012	0.002697	0.001381

В сносках указаны номера формул, по которым происходило вычисление/ оценка.

Дисперсия оценивалась с точностью до четырёх значащих цифр, среднее - до двух.

5 Вывод

В ходе лабораторной работы было изучено четыре распределения: нормальное, Пуассона, Коши, равномерное. Для каждого были сгенерированы выборки размеров, указанных в соответствующих пунктах задания.

Для каждого распределения и размера выборки были построены гистограммы и графики плотности вероятности. Было замечено, что при увеличении числа элементов в выборке, гистограмма становится более похожей на график функции плотности.

Также были оценены различные характеристики каждой выборки.

В случае нормального распределения можно заметить, что характеристики положения и рассеивания с увеличением выборки приближаются к нулю, что объясняется симметричностью распределения.

Распределение Коши имеет особое поведение: выборочное среднее не имеет конечного математического ожидания или дисперсии, поэтому значение оказывается нестабильным. Характеристики положения и рассеивания медианы и полусуммы квартилей в случае распределения Коши приближаются к нулю (также в силу симметричности).

В случае распределения Пуассона характеристики положения оказались примерно равными 10. Оценка дисперсии стремится к нулю при увеличении числа элементов, что также согласуется с теорией (здесь, например, $D(\bar{x}) = \frac{10}{n}$).

Для равномерного распределения оценки также устремляются к нулю,

что согласуется с их теоретической оценкой (в силу симметричности интервала математическое ожидание величин должно быть равно нулю, а в силу обратной зависимости от количества элементов дисперсии оно также приближается к нулю).

Также отметим, что выборочное среднее будет иметь самую низкую дисперсию среди всех оценок, не считая распределения Коши, поскольку для них оно наилучшим образом приближает мат ожидание.

Таким образом, в ходе данной лабораторной работы были изучены свойства основных распределений и их оценки.