

Санкт-Петербургский Политехнический Университет Петра
Великого
Физико-механический институт

Высшая школа прикладной математики и вычислительной
физики

Лабораторная работа №1
по дисциплине “Математическая статистика”

Преподаватель:

А.Н. Баженов

Обучающаяся:

А.Д. Балакшина
(группа 5030102/20101)

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:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

- Выборочная медиана:

$$medx = \begin{cases} x_{\frac{n+1}{2}} & n - \text{нечётное} \\ \frac{x_{\frac{n}{2}} + x_{\frac{n+2}{2}}}{2} & n - \text{чётное} \end{cases}$$

- Полусумма квартилей:

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

- Оценка дисперсии:

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

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

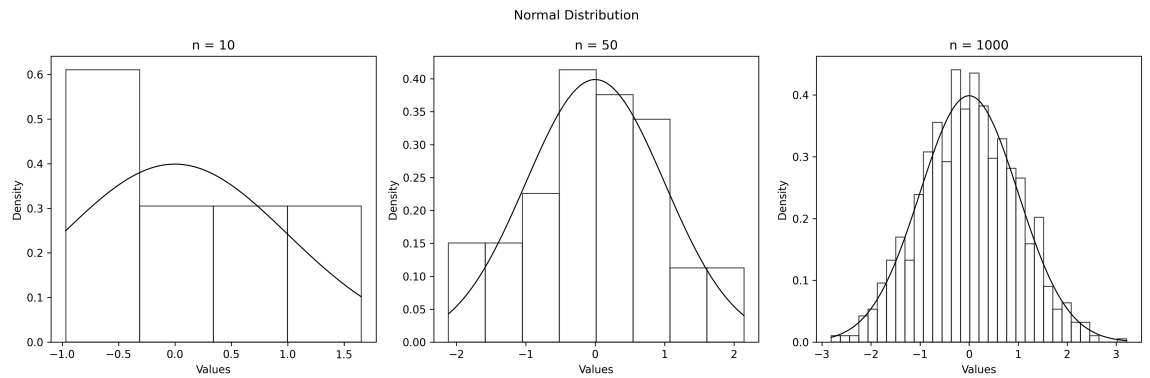


Рис. 1: Нормальное распределение.

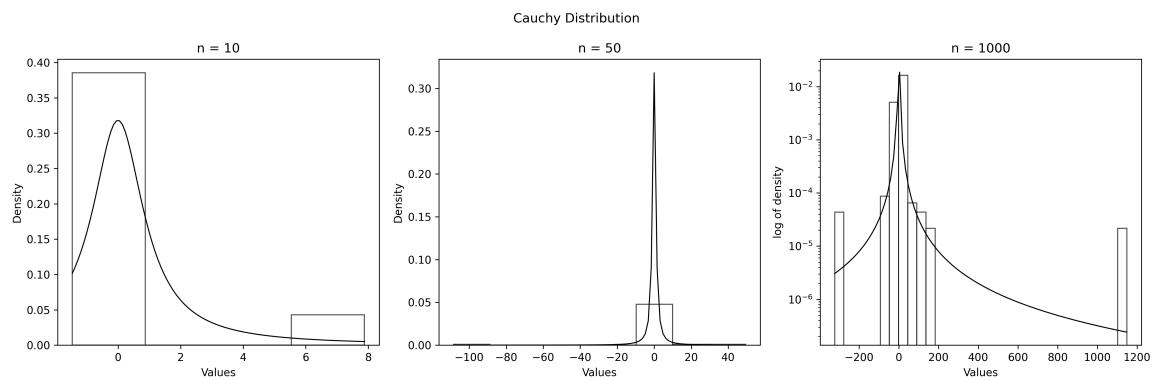


Рис. 2: Распределение Коши.

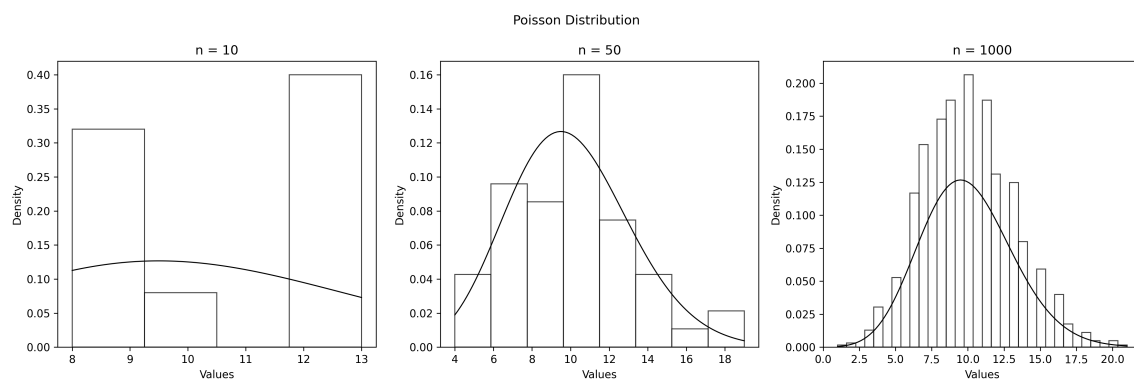


Рис. 3: Распределение Пуассона.

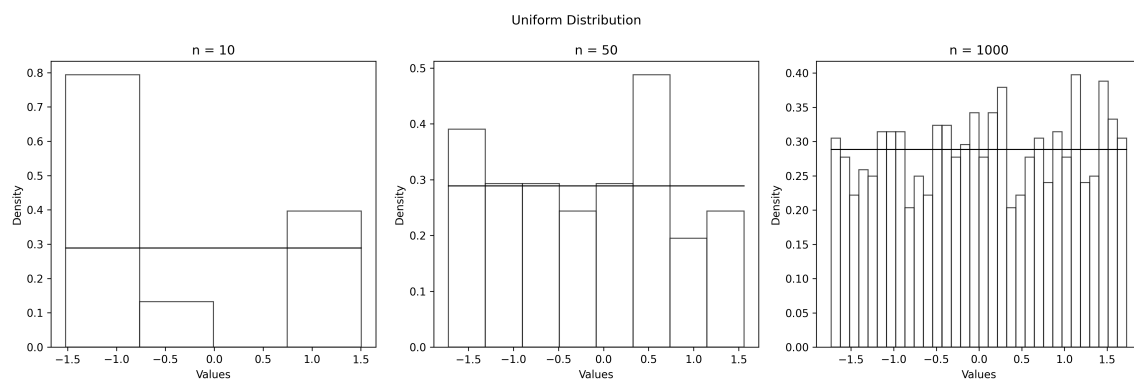


Рис. 4: Равномерное распределение.

4.2 Характеристики распределений

Normal distribution

Sample size 10

	x	med x	z_Q
E(z)	-0.022434	-0.021191	-0.025037
D(z)	0.102296	0.141397	0.118712

Sample size 100

	x	med x	z_Q
E(z)	-0.003147	-0.001905	-0.019981
D(z)	0.009557	0.015348	0.012158

Sample size 1000

	x	med x	z_Q
E(z)	-0.000839	-0.000362	-0.001753
D(z)	0.001016	0.001632	0.001270

Cauchy distribution

Sample size 10

	x	med x	z_Q
E(z)	-0.944701	-0.018151	-0.065972
D(z)	480.215854	0.308154	1.146016

Sample size 100

	x	med x	z_Q
E(z)	0.905721	-0.009459	-0.043541
D(z)	4195.020155	0.025563	0.054394

Sample size 1000

	x	med x	z_Q
E(z)	-3.988564	-0.000053	-0.003675
D(z)	13606.759490	0.002649	0.005158

Poisson distribution

Sample size 10

	x	med x	z_Q
E(z)	10.030400	9.916000	9.947500
D(z)	1.016536	1.411444	1.217494

Sample size 100

	x	med x	z_Q
E(z)	10.001790	9.843000	9.865000
D(z)	0.104111	0.201851	0.171775

Sample size 1000

	x	med x	z_Q
E(z)	9.998851	9.997000	9.996000
D(z)	0.009833	0.002491	0.002484

Uniform distribution

Sample size 10

	x	med x	z_Q
E(z)	-0.003724	-0.011343	-0.001972
D(z)	0.100455	0.243433	0.131828

Sample size 100

	x	med x	z_Q
E(z)	0.000415	-0.003007	-0.016770
D(z)	0.010058	0.029374	0.015065

Sample size 1000

	x	med x	z_Q
E(z)	0.000282	0.000325	-0.001374
D(z)	0.001100	0.003325	0.001554

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.