

Санкт-Петербургский  
Политехнический университет Петра Великого

**Отчет по лабораторным работам №1-2  
по дисциплине  
"Математическая статистика"**

Студент:	Скворцов Владимир Сергеевич
Преподаватель:	Баженов Александр Николаевич
Группа:	5030102/10201

Санкт-Петербург  
2024

# Содержание

<b>1</b>	<b>Постановка задачи</b>	<b>2</b>
1.1	Описательная статистика . . . . .	2
1.2	Точечное оценивание характеристик положения и рассеяния . . . . .	2
<b>2</b>	<b>Теоретическое обоснование</b>	<b>2</b>
2.1	Функции распределения . . . . .	2
2.2	Характеристики положения и рассеяния . . . . .	3
<b>3</b>	<b>Описание работы</b>	<b>3</b>
<b>4</b>	<b>Результаты</b>	<b>4</b>
4.1	Гистограммы и графики плотности распределения . . . . .	4
4.2	Характеристики положения и рассеяния . . . . .	6
<b>5</b>	<b>Выводы</b>	<b>8</b>

# 1 Постановка задачи

## 1.1 Описательная статистика

Для 5 распределений:

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

Сгенерировать выборки размером 10, 50, 1000 элементов.

Построить на одном рисунке гистограмму и график плотности распределения.

## 1.2 Точечное оценивание характеристик положения и рассеяния

Сгенерировать выборки размером 10, 50, 1000 элементов.

Для каждой выборки вычислить следующие статистические характеристики положения данных:  $\bar{x}$ ,  $med\ x$ ,  $z_Q$ ,  $z_R$ ,  $z_{tr}$ . Повторить такие вычисления 1000 раз для каждой выборки и найти среднее характеристик положения и их квадратов:  $E(z) = \bar{z}$ . Вычислить оценку дисперсии по формуле  $D(z) = \overline{z^2} - \bar{z}^2$ .

# 2 Теоретическое обоснование

## 2.1 Функции распределения

- Нормальное распределение

$$N(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \quad (1)$$

- Распределение Коши

$$C(x, 0, 1) = \frac{1}{\pi} \frac{1}{x^2 + 1} \quad (2)$$

- Распределение Стьюдента  $t(x, 0, 3)$  с тремя степенями свободы

$$t(x, 0, 3) = \frac{6\sqrt{3}}{\pi(3 + t^2)^2} \quad (3)$$

- Распределение Пуассона

$$P(k, 10) = \frac{10^k}{k!} e^{-10} \quad (4)$$

- Равномерное распределение

$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{при } |x| \leq \sqrt{3} \\ 0 & \text{при } |x| > \sqrt{3} \end{cases} \quad (5)$$

## 2.2 Характеристики положения и рассеяния

- Выборочное среднее

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (6)$$

- Выборочная медиана

$$\text{med } x = \begin{cases} x_{(l+1)} & \text{при } n = 2l + 1 \\ \frac{x_{(l)} + x_{(l+1)}}{2} & \text{при } n = 2l \end{cases} \quad (7)$$

- Полусумма экстремальных выборочных элементов

$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \quad (8)$$

- Полусумма квартилей

Выборочная квартиль  $z_p$  порядка  $p$  определяется формулой

$$z_p = \begin{cases} x_{([np]+1)} & \text{при } np \text{ дробном} \\ x_{(np)} & \text{при } np \text{ целом} \end{cases} \quad (9)$$

Полусумма квартилей

$$z_Q = \frac{z_{1/4} + z_{3/4}}{2} \quad (10)$$

- Усечённое среднее

$$z_{tr} = \frac{1}{n-2r} \sum_{i=r+1}^{n-r} x_{(i)}, \quad r \approx \frac{n}{4} \quad (11)$$

- Среднее характеристики

$$E(z) = \bar{z} \quad (12)$$

- Оценка дисперсии

$$D(z) = \overline{z^2} - \bar{z}^2 \quad (13)$$

## 3 Описание работы

Лабораторные работы выполнены с использованием Python и его сторонних библиотек `numpy`, `pandas`, `matplotlib`, `seaborn` были построены гистограммы распределений и посчитаны характеристики положения.

Ссылка на GitHub репозиторий: <https://github.com/vladimir-skvortsov/spbstu-mathematical-statistics>

## 4 Результаты

### 4.1 Гистограммы и графики плотности распределения

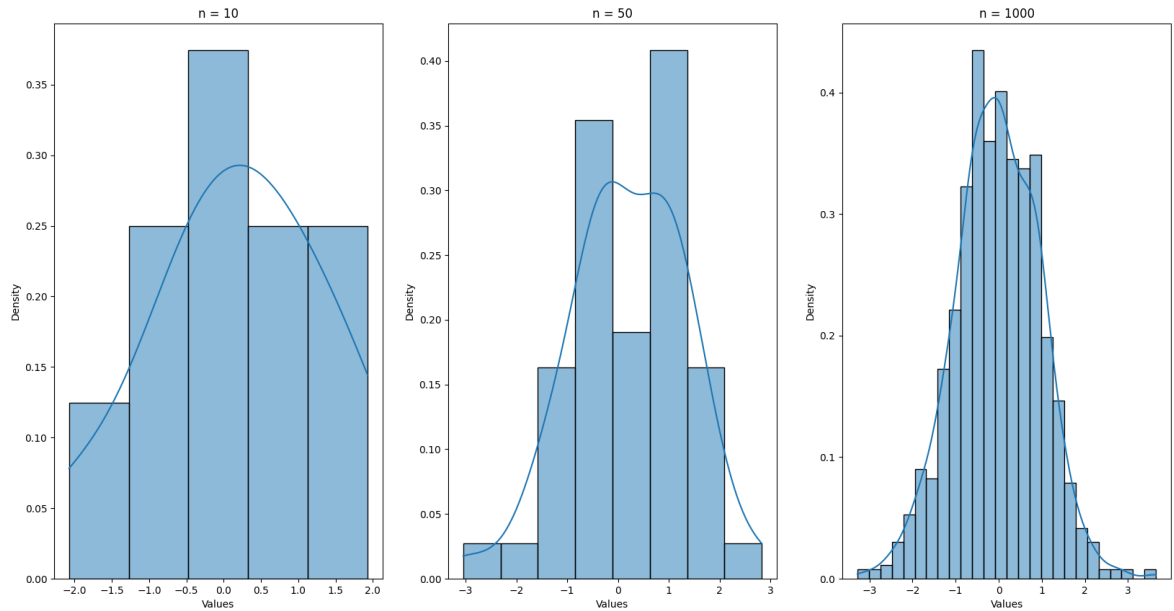


Рис. 1: Нормальное распределение (1)

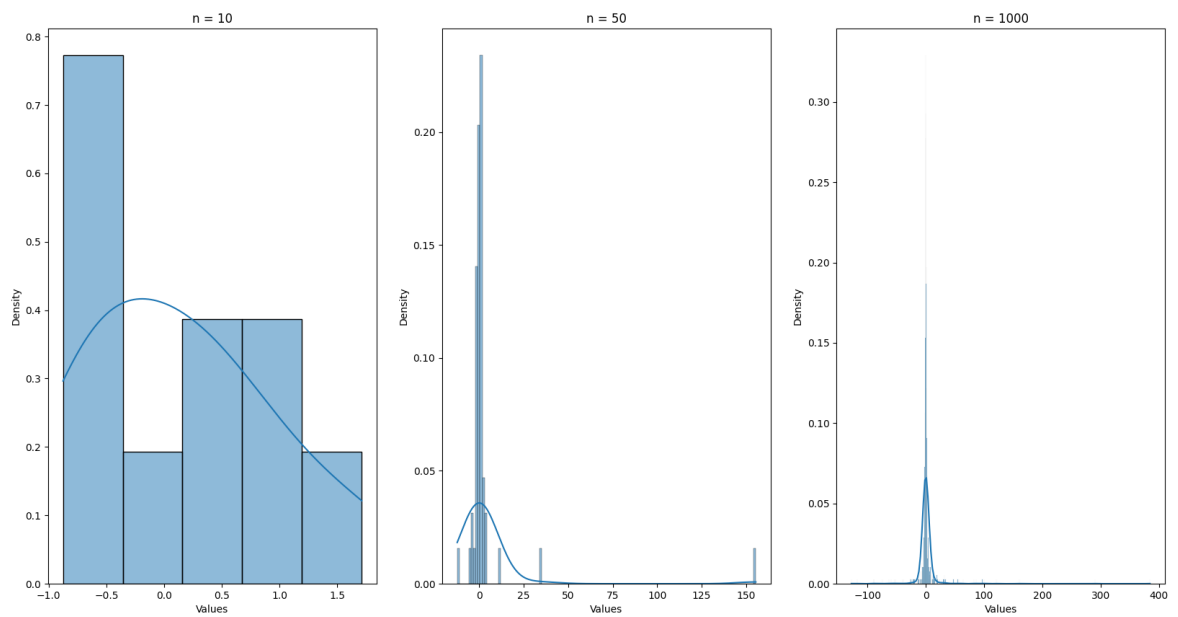


Рис. 2: Распределение Коши (2)

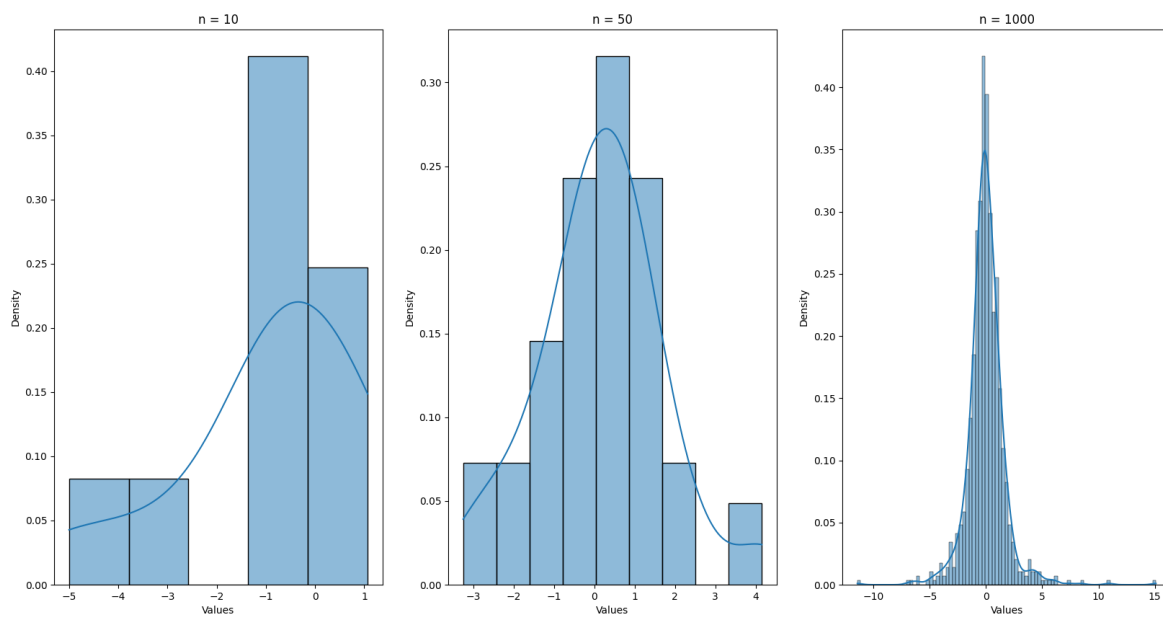


Рис. 3: Распределение Стьюдента (3)

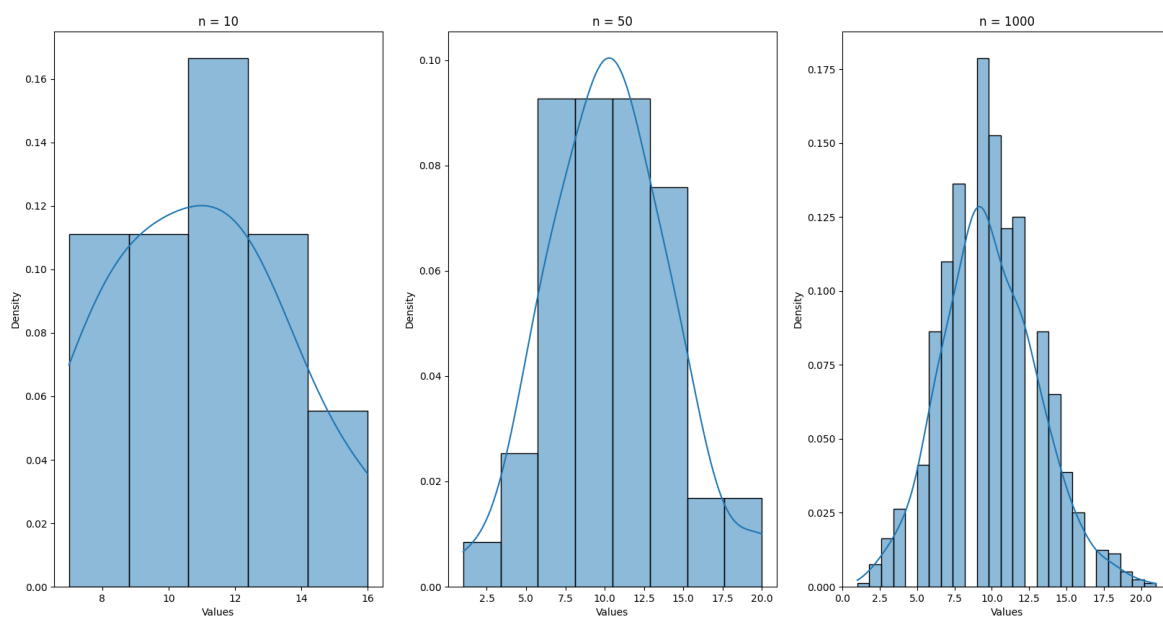


Рис. 4: Распределение Пуассона (4)

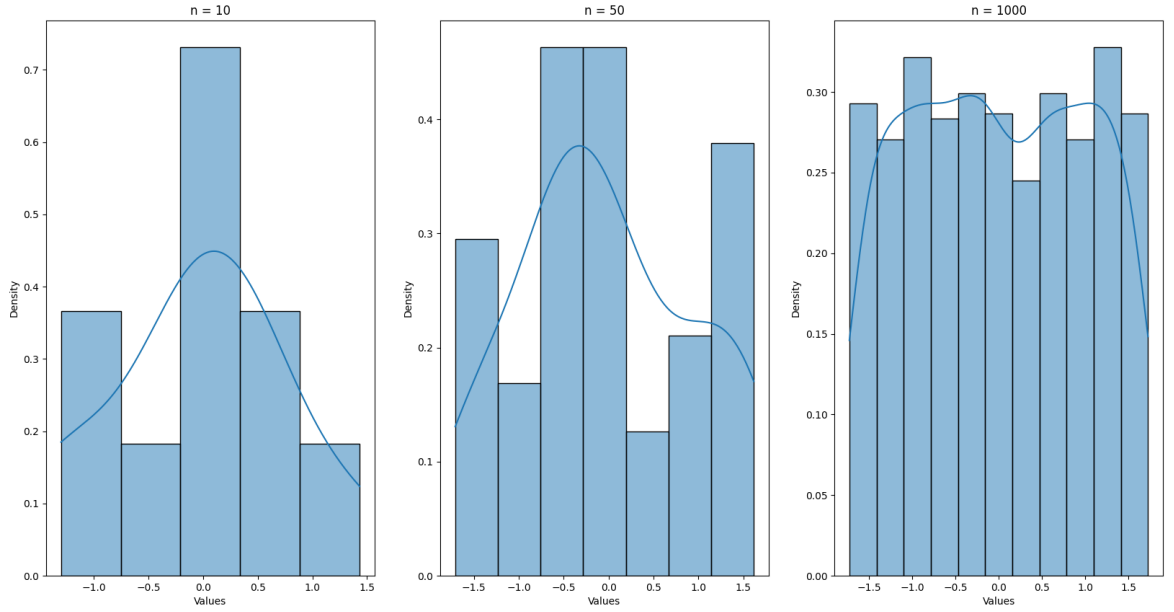


Рис. 5: Равномерное распределение (5)

## 4.2 Характеристики положения и рассеяния

n = 10					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	-0.007982	-0.004386	-0.020996	-0.005516	-0.006624
$D(z)$ (13)	0.098372	0.143342	0.163284	0.113910	0.163167
n = 50					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.003888	0.006077	-0.000529	0.002100	0.005262
$D(z)$ (13)	0.010963	0.016229	0.092033	0.013306	0.019508
n = 1000					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.000841	0.001757	0.002664	0.000549	0.001120
$D(z)$ (13)	0.001003	0.001594	0.056454	0.001218	0.002025

Таблица 1: Нормальное распределение

n = 10					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.031223	0.000157	0.091224	0.010353	0.026922
$D(z)$ (13)	0.593557	0.276924	3.039111	0.431265	1.018721
n = 50					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	-0.009060	-0.001960	-0.009758	-0.005885	-0.001198
$D(z)$ (13)	0.059867	0.024296	1.246451	0.039833	0.119172
n = 1000					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.000632	-0.000889	0.005623	-0.001120	0.000454
$D(z)$ (13)	0.006238	0.002337	0.038649	0.004053	0.012341

Таблица 2: Распределение Коши

n = 10					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	-0.008120	0.003401	-0.051670	0.006376	-0.006949
$D(z)$ (13)	0.226164	0.168610	1.170873	0.181303	0.391251
n = 50					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.001555	-0.003052	0.021573	-0.002192	0.007673
$D(z)$ (13)	0.023879	0.018483	1.416933	0.018728	0.048591
n = 1000					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.002647	0.000765	0.005358	0.000688	0.000826
$D(z)$ (13)	0.002287	0.001833	0.584625	0.001828	0.004626

Таблица 3: Распределение Стьюдента

n = 10					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	10.060200	9.9315	10.343	9.97575	10.0495
$D(z)$ (13)	1.022076	1.364558	1.958351	1.159037	1.690466
n = 50					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	9.997760	9.856500	10.961500	9.904125	9.990620
$D(z)$ (13)	0.101923	0.209158	0.990268	0.147511	0.205333
n = 1000					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	10.001017	9.998000	11.683000	9.995250	10.006934
$D(z)$ (13)	0.010374	0.001996	0.741011	0.003759	0.020079

Таблица 4: Распределение Пуассона



n = 10					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.002619	0.013534	-0.007382	0.001860	0.003422
$D(z)$ (13)	0.096635	0.220092	0.046032	0.133492	0.163799
n = 50					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	0.001636	5.804625-05	-0.000114	0.002965	0.004812
$D(z)$ (13)	0.009758	0.029210	0.000521	0.014602	0.018929
n = 1000					
	$\bar{x}$ (6)	$med\ x$ (7)	$z_R$ (8)	$z_Q$ (10)	$z_{tr}$ (11)
$E(z)$ (12)	-0.000112	-0.000167	-1.956639-05	0.000345	-0.001010
$D(z)$ (13)	0.000967	0.002844	5.856189-06	0.001470	0.001981

Таблица 5: Равномерное распределение

## 5 Выводы

В процессе выполнения лабораторной работы был проведен анализ пяти уникальных распределений: нормальное, Коши, Стьюдента, Пуассона и равномерное. Были сгенерированы выборки разных объемов для каждого из них - 10, 50 и 1000 элементов. Были созданы гистограммы каждого распределения и нанесены на них графики плотности соответствующих распределений, что облегчило наглядное сопоставление формы распределения выборок с их теоретическими аналогами. Были также рассчитаны разные показатели положения и рассеяния для каждой выборки, включая выборочную среднюю величину, медиану, полусумму крайних элементов выборки, полусумму квартилей и усеченное среднее. Использовалась стандартная формула для оценки дисперсии.

На основании полученных данных были сделаны следующие выводы:

1. В случае нормального распределения, оценки показателей положения и рассеяния становятся ближе к их теоретическим значениям по мере увеличения размера выборки.
2. Для распределения Коши показатели положения и рассеяния менее стабильны и могут сильно отличаться от теоретических даже при больших размерах выборки.
3. Распределение Стьюдента при небольших размерах выборки также демонстрирует определенную нестабильность оценок, однако с увеличением размера выборки результаты становятся более точными.
4. Для распределения Пуассона и равномерного распределения, оценки показателей положения и рассеяния кажутся стабильными при любом объеме выборки.