

Санкт-Петербургский
Политехнический университет Петра Великого

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

Студент:	Афанасьев Андрей Валерьевич
Преподаватель:	Баженов Александр Николаевич
Группа:	5030102/10201

Санкт-Петербург
2024

Содержание

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

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} , $medx$, 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

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

- Оценка дисперсии

$$D(z) = \overline{z^2} - \bar{z}^2 \quad (12)$$

3 Описание работы

Лабораторные работы выполнены на языке программирования Python. С использованием сторонних библиотек `numpy`, `matplotlib`, `pandas`, `IPython` были построены гistogramмы распределений и посчитаны характеристики положения.

Ссылка на GitHub репозиторий: https://github.com/afafos/Statistics/tree/main/lab_1

4 Результаты

4.1 Гистограммы и графики плотности распределения

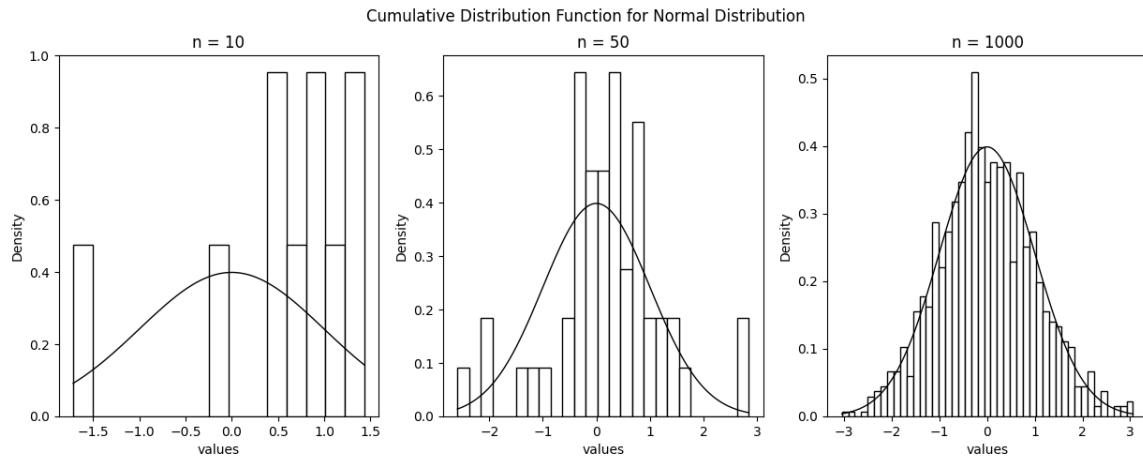


Рис. 1: Нормальное распределение

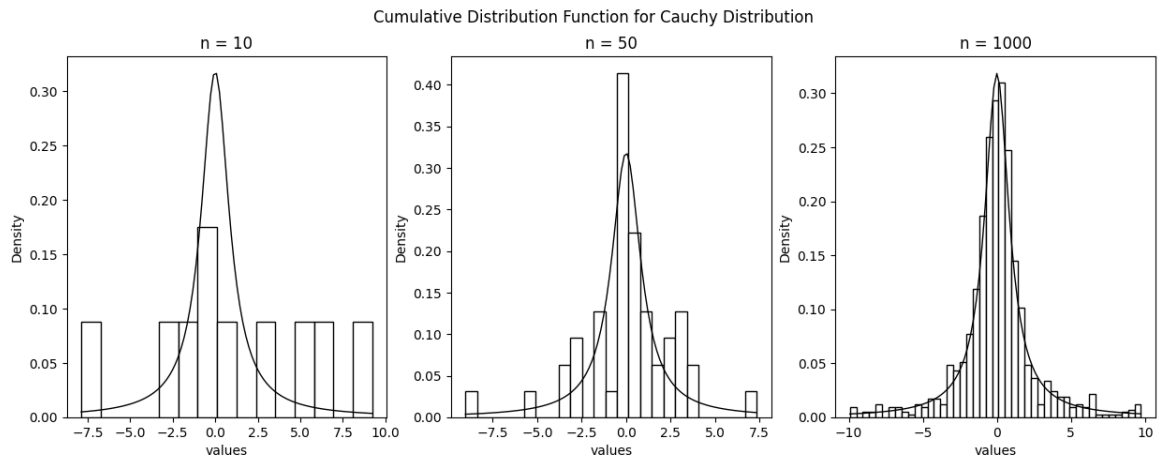


Рис. 2: Распределение Коши

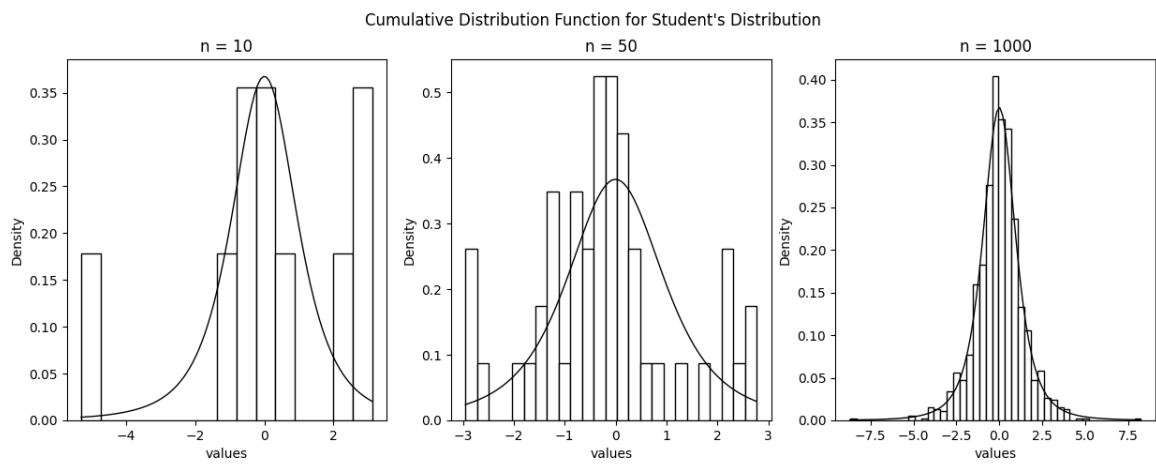


Рис. 3: Распределение Стьюдента

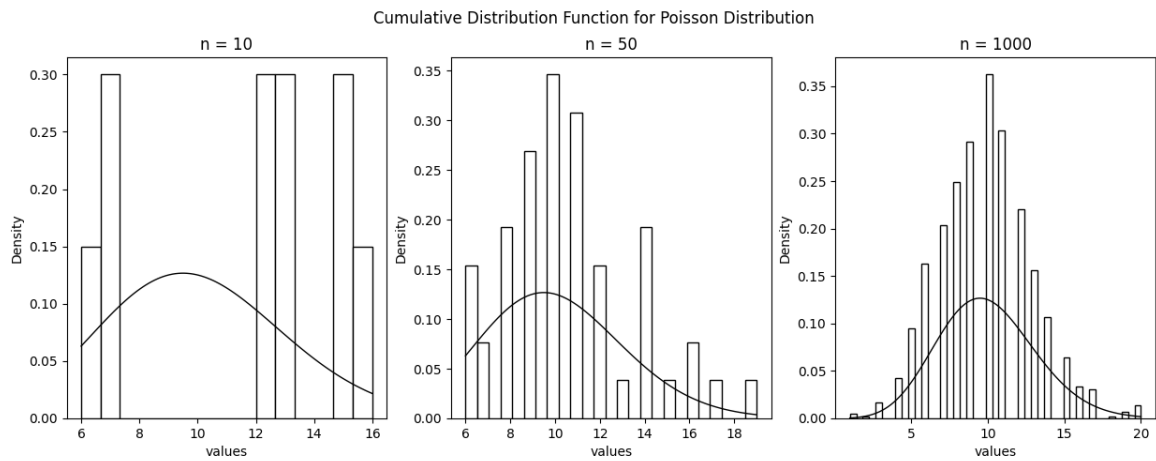


Рис. 4: Распределение Пуассона

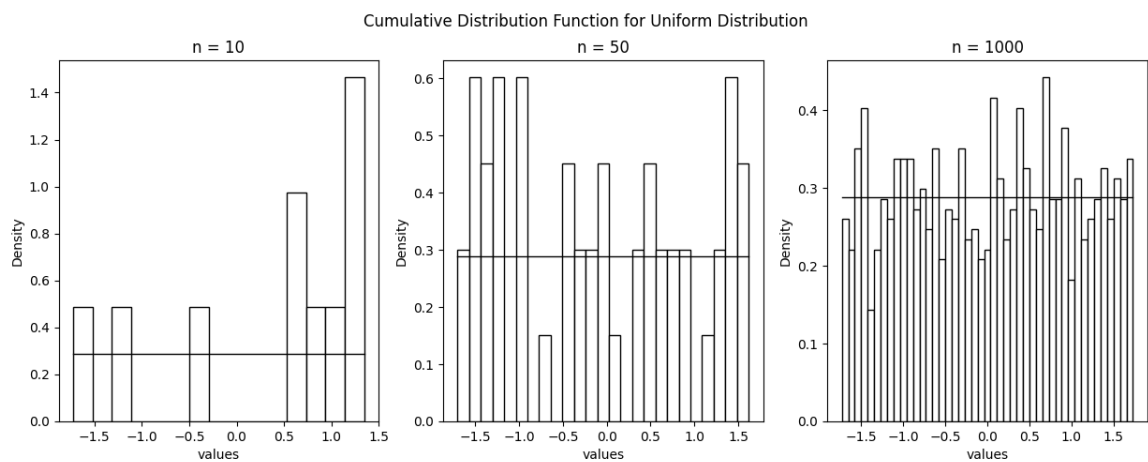


Рис. 5: Равномерное распределение

4.2 Характеристики положения и рассеяния

n = 10					
	\bar{x} (6)	$med\ x$ (7)	z_R (8)	z_Q (10)	z_{tr} (11)
$E(z)$	0.005893	0.004250	0.008004	0.004354	0.005270
$D(z)$ (12)	0.098927	0.134744	0.180383	0.125228	0.114211
n = 50					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	0.005048	0.012418	-0.005397	0.005113	0.010137
$D(z)$	0.020863	0.030744	0.113293	0.025061	0.024081
n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	0.000045	0.000098	-0.012101	-0.001734	0.000118
$D(z)$	0.000997	0.001516	0.061306	0.001217	0.001194

Таблица 1: Нормальное распределение

n = 10					
	\bar{x} (6)	$med\ x$ (7)	z_R (8)	z_Q (10)	z_{tr} (11)
$E(z)$	0.810695	-0.020504	4.179022	-0.014779	-0.020792
$D(z)$ (12)	878.433932	0.273504	21722.377106	3.003419	0.648409
n = 50					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	-0.531111	0.01267	-13.367328	0.003952	0.008745
$D(z)$	181.463944	0.05363	106478.010121	0.116777	0.059709
n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	0.362018	0.000957	1.668537e+02	-0.001921	0.001206
$D(z)$	1073.614829	0.002229	2.577761e+08	0.004664	0.002272

Таблица 2: Распределение Коши

n = 10					
	\bar{x} (6)	$med\ x$ (7)	z_R (8)	z_Q (10)	z_{tr} (11)
$E(z)$	0.051233	0.024060	0.131840	0.026916	0.027702
$D(z)$ (12)	0.297739	0.175002	2.482403	0.184520	0.153276
n = 50					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	-0.001012	0.001274	-0.014033	0.000464	-0.000238
$D(z)$	0.054893	0.034997	3.892715	0.038857	0.029885
n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	-0.002679	-0.002163	-0.011349	-0.004241	-0.002402
$D(z)$	0.002924	0.001722	37.810339	0.001751	0.001430

Таблица 3: Распределение Стьюдента

n = 10					
	\bar{x} (6)	$med\ x$ (7)	z_R (8)	z_Q (10)	z_{tr} (11)
$E(z)$	9.973300	9.81550	10.318000	9.887000	9.844167
$D(z)$ (12)	1.038757	1.43121	1.882876	1.284731	1.182077
n = 50					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	9.994020	9.831000	10.783500	9.89550	9.838769
$D(z)$	0.190543	0.331939	1.097378	0.28533	0.230306
n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	9.995386	9.998000	11.667000	9.996000	9.853080
$D(z)$	0.009967	0.001996	0.663111	0.001984	0.011106

Таблица 4: Распределение Пуассона

n = 10					
	\bar{x} (6)	$med\ x$ (7)	z_R (8)	z_Q (10)	z_{tr} (11)
$E(z)$	0.004536	0.004262	-0.001677	0.004075	0.006316
$D(z)$ (12)	0.103443	0.231639	0.043614	0.144378	0.168062
n = 50					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	-0.009111	-0.013460	-0.000285	-0.011838	-0.012313
$D(z)$	0.018825	0.054305	0.002186	0.028717	0.036225
n = 1000					
	\bar{x}	$med\ x$	z_R	z_Q	z_{tr}
$E(z)$	0.000262	0.001315	0.000049	-0.001617	0.000761
$D(z)$	0.001046	0.003068	0.000005	0.001600	0.002090

Таблица 5: Равномерное распределение

5 Выводы

В ходе лабораторной работы были рассмотрены пять различных распределений: нормальное, Коши, Стьюдента, Пуассона и равномерное. Для каждого из них были сгенерированы выборки разного размера (10, 50, 1000 элементов).

Были построены гистограммы для каждого распределения и наложены на них графики плотности соответствующих распределений. Это позволило визуально сравнить форму распределения выборок с теоретическими моделями.

Также были вычислены различные характеристики положения и рассеяния для каждой выборки: выборочное среднее, медиана, полусумма экстремальных выборочных элементов, полусумма квартилей и усечённое среднее. Для оценки дисперсии использовалась соответствующая формула.

Полученные результаты позволяют сделать следующие выводы:

В случае нормального распределения, при увеличении размера выборки, оценки характеристик положения и рассеяния приближаются к их теоретическим значениям.

Для распределения Коши характеристики положения и рассеяния менее стабильны и могут значительно отличаться от теоретических значений даже при больших размерах выборки.

Распределение Стьюдента при малых размерах выборки также демонстрирует некоторую неустойчивость оценок, но при увеличении размера выборки оценки становятся более точными.

Для распределения Пуассона и равномерного распределения наблюдается стабильность оценок характеристик положения и рассеяния при любом размере выборки.

Таким образом, выполненная работа позволила ознакомиться с различными статистическими характеристиками и методами их оценки, а также провести практическое исследование на примере различных распределений.