

Санкт-Петербургский политехнический университет Петра Великого  
Институт прикладной математики и механики  
**Высшая школа прикладной математики и вычислительной физики**

**ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №2**

по дисциплине  
«Математическая статистика»

Выполнила студентка  
группы 3630102/80401

Мамаева Анастасия Сергеевна

Проверил  
Доцент, к.ф.-м.н.

Баженов Александр Николаевич

Санкт-Петербург  
2021

# СОДЕРЖАНИЕ

<b>СПИСОК ТАБЛИЦ . . . . .</b>	<b>3</b>
<b>1 Постановка задачи . . . . .</b>	<b>4</b>
<b>2 Теория . . . . .</b>	<b>4</b>
2.1 Распределения . . . . .	4
2.2 Вариационный ряд . . . . .	4
2.3 Выборочные числовые характеристики . . . . .	5
2.3.1 Характеристики положения . . . . .	5
2.3.2 Характеристики рассеяния . . . . .	5
<b>3 Программная реализация . . . . .</b>	<b>5</b>
<b>4 Результаты . . . . .</b>	<b>6</b>
4.1 Характеристики положения и рассеяния . . . . .	6
<b>5 Обсуждение . . . . .</b>	<b>8</b>
<b>6 Приложение . . . . .</b>	<b>8</b>

## СПИСОК ТАБЛИЦ

1	Нормальное распределение (3) . . . . .	6
2	Распределение Коши (4) . . . . .	6
3	Распределение Лапласа (5) . . . . .	7
4	Распределение Пуассона (6) . . . . .	7
5	Равномерное распределение (7) . . . . .	7

# 1 Постановка задачи

Сгенерировать выборки размером 10, 100 и 1000 элементов. Для каждой выборки вычислить следующие статистические характеристики положения данных:  $\bar{x}$ ,  $medx$ ,  $z_R$ ,  $z_Q$ ,  $z_{tr}$ . Повторить такие вычисления 1000 раз для каждой выборки и найти среднее характеристик положения и их квадратов:

$$E(z) = \bar{z} \quad (1)$$

Вычислить оценку дисперсии по формуле:

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

Представить полученные данные в виде таблиц.

## 2 Теория

### 2.1 Распределения

- Нормальное распределение

$$N(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \quad (3)$$

- Распределение Коши

$$C(x, 0, 1) = \frac{1}{\pi} \frac{1}{x^2 + 1} \quad (4)$$

- Распределение Лапласа

$$L(x, 0, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} \quad (5)$$

- Распределение Пуассона

$$P(k, 10) = \frac{10^k}{k!} e^{-10} \quad (6)$$

- Равномерное распределение

$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{при } |x| \leq \sqrt{3} \\ 0 & \text{при } |x| > \sqrt{3} \end{cases} \quad (7)$$

### 2.2 Вариационный ряд

Вариационным рядом называется последовательность элементов выборки, расположенных в неубывающем порядке. Одинаковые элементы повторяются. Запись вариационного ряда:  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ . Элементы вариационного ряда  $x_{(i)}$  ( $i = 1, 2, \dots, n$ ) называются порядковыми статистиками.

## 2.3 Выборочные числовые характеристики

С помощью выборки образуются её числовые характеристики. Это числовые характеристики дискретной случайной величины  $X^*$ , принимающей выборочные значения  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ .

### 2.3.1 Характеристики положения

- Выборочное среднее

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (8)$$

- Выборочная медиана

$$medx = \begin{cases} x_{(l+1)} & n = 2l + 1 \\ \frac{x_{(l)} + x_{(l+1)}}{2} & n = 2l \end{cases} \quad (9)$$

- Полусумма экстремальных выборочных элементов

$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \quad (10)$$

- Полусумма квартилей

Выборочная квартиль  $z_p$  порядка  $p$  определяется формулой

$$z_p = \begin{cases} x_{([np]+1)} & np\text{—дробное} \\ x_{(np)} & np\text{—целое} \end{cases} \quad (11)$$

Полусумма квартилей

$$z_Q = \frac{z_{1/4} + z_{3/4}}{2} \quad (12)$$

- Усечённое среднее

$$z_{tr} = \frac{1}{n - 2r} \sum_{i=r+1}^{n-r} x_{(i)}, r \approx \frac{n}{4} \quad (13)$$

### 2.3.2 Характеристики рассеяния

Выборочная дисперсия

$$D = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (14)$$

## 3 Программная реализация

Лабораторная работа выполнена на языке Python версии 3.7 в среде разработки JupyterLab. Использовались дополнительные библиотеки:

1. `scipy`

2. numpy

В приложении находится ссылка на GitHub репозиторий с исходным кодом.

## 4 Результаты

### 4.1 Характеристики положения и рассеяния

Как было проведено округление:

В оценке  $x = E \pm D$  вариации подлежит первая цифра после точки. В данном случае  $x = 0.0 \pm 0.1k$ ,  $k$  - зависит от доверительной вероятности и вида распределения (рассматривается в дальнейшем цикле лабораторных работ). Округление сделано для  $k = 1$ .

Characteristic	Mean	Median	$z_R$	$z_Q$	$z_{tr}$
Normal E(z) 10	-0.000003	-0.008031	0.036451	-0.021004	-0.012014
Normal D(z) 10	0.101239	0.091199	0.480519	0.475769	0.172801
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Normal E(z) 100	-0.002052	-0.004684	0.012205	-0.011495	-0.008917
Normal D(z) 100	0.05601	0.050246	0.478019	0.50419	0.097187
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Normal E(z) 1000	-0.001219	-0.003105	0.006667	-0.010076	-0.006475
Normal D(z) 1000	0.037681	0.033809	0.475955	0.497166	0.065463
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0

Таблица 1: Нормальное распределение (3)

Characteristic	Mean	Median	$z_R$	$z_Q$	$z_{tr}$
Cauchy E(z) 10	-1.787828	-0.025951	-5.130831	-3.948137	-2.645886
Cauchy D(z) 10	4390.971	0.403047	71558.563	21478.02	2724.385
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Cauchy E(z) 100	-2.946291	-0.011733	-2.919987	7.274195	-5.064231
Cauchy D(z) 100	11499.485	0.214055	37287.433	194370.44	37068.767
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Cauchy E(z) 1000	-2.272925	-0.007192	-1.740728	4.70099	-4.116761
Cauchy D(z) 1000	7794.082	0.143556	25043.95	129732.73	25183.86
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0

Таблица 2: Распределение Коши (4)

Characteristic	Mean	Median	$z_R$	$z_Q$	$z_{tr}$
Laplace E(z) 10	0.022501	0.011001	0.011483	0.023003	0.027397
Laplace D(z) 10	0.103582	0.077859	0.490929	0.47279	0.169248
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Laplace E(z) 100	0.012271	0.005379	0.021852	0.021828	0.015041
Laplace D(z) 100	0.057048	0.041891	0.492694	0.492542	0.095815
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Laplace E(z) 1000	0.008403	0.003924	0.022277	0.022505	0.010596
Laplace D(z) 1000	0.038374	0.028103	0.487406	0.505289	0.064555
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0

Таблица 3: Распределение Лапласа (5)

Characteristic	Mean	Median	$z_R$	$z_Q$	$z_{tr}$
Poisson E(z) 10	9.979	9.8525	9.9655	10.044	9.997167
Poisson D(z) 10	0.975799	1.373994	4.68356	5.018564	1.779575
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Poisson E(z) 100	9.98473	9.8535	9.9615	10.0425	9.986043
Poisson D(z) 100	0.534271	0.786788	4.841018	5.073944	0.982735
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Poisson E(z) 1000	9.990781	9.902	9.955	10.057	9.99035
Poisson D(z) 1000	0.359408	0.529563	4.769308	5.025751	0.661672
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0

Таблица 4: Распределение Пуассона (6)

Characteristic	Mean	Median	$z_R$	$z_Q$	$z_{tr}$
Uniform E(z) 10	-0.003058	-0.008189	-0.008308	-0.058391	-0.000635
Uniform D(z) 10	0.10312	0.240584	0.509498	0.516179	0.1717448
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Uniform E(z) 100	-0.001261	-0.005273	-0.013241	-0.026719	-0.001152
Uniform D(z) 100	0.056234	0.134553	0.506669	0.513444	0.095636
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0
Uniform E(z) 1000	-0.000621	-0.003033	-0.005643	-0.017729	-0.000709
Uniform D(z) 1000	0.037829	0.090714	0.499291	0.501763	0.064392
$E(z) \pm \sqrt{D(z)}$	0.0	0.0	0.0	0.0	0.0

Таблица 5: Равномерное распределение (7)

## 5 Обсуждение

Исходя из данных, приведенных в таблицах, можно судить о том, что дисперсия характеристик рассеяния для распределения Коши является некой аномалией: значения слишком большие даже при увеличении размера выборки - понятно, что это результат выбросов, которые мы могли наблюдать в результатах предыдущего задания.

## 6 Приложение

Код программы GitHub URL:

<https://github.com/Brightest-Sunshine/Math-Statistic-2021/blob/main/Lab2/Lab2.ipynb>