Санкт-Петербургский политехнический университет
Петра Великого

Физико-механический институт

Кафедра «Прикладная математика»

Отчет по лабораторной работе №2

по дисциплине
"Математическая статистика"

Выполнил студент
группы 5030102/00101 Нгуен Хоанг Линь

Проверил
доцент, к.ф.-м.н. Баженов Александр Николаевич

Санкт-Петербург
2023 г.

Содержание

1. Постановка задачи **4**

2. Теория **4**

2.1. Рассматриваемые распределения 4

2.2. Выборочные числовые характеристики 5

2.2.1. Характеристики положения 5

2.2.2. Характеристики рассеивания 5

3. Реализация **5**

4. Результаты **6**

5. Обсуждение **8**

Список таблиц

1. Нормальное распределение (1) 6
2. Распределение Коши(2) 6
3. Распределение Лапласа(3) 7
4. Распределение Пуассона(4) 7
5. Равномерное распределение(5) 8

1. Постановка задачи

Для 5 распределений:

- Нормальное распределение $N(x, 0, 1)$
- Распределение Коши $C(x, 0, 1)$
- Распределение Лапласа $L(x, 0, \frac{1}{\sqrt{2}})$
- Распределение Пуассона $P(k, 10)$
- Равномерное распределение $U(x, -\sqrt{3}, \sqrt{3})$

Необходимо:

1) Сгенерировать выборки размером 10, 100 и 1000 элементов
2) Вычислить для каждой них статистические характеристики положения данных: $\overline{x}, medx, z_R, z_Q, z_tr$
3) Повторить данные вычисления 1000 раз для каждой выборки и найти среднее характеристик положения $E(z) = \overline{z}$ и вычислить оценку дисперсии $D(z) = \overline{z^2} - \overline{z}^2$
4) Представить полученные результаты в виде таблиц

2. Теория

2.1. Рассматриваемые распределения

Плотности:

- Нормальное распределение
$$N(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{1}$$
- Распределение Коши
$$C(x, 0, 1) = \frac{1}{\pi} \frac{1}{x^2 + 1} \tag{2}$$
- Распределение Лапласа
$$L(x, 0, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} \tag{3}$$
- Распределение Пуассона
$$P(k, 10) = \frac{10^k}{k!} e^{-10} \tag{4}$$
- Равномерное распределение
$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}}, |x| \le \sqrt{3} \\ 0, |x| > \sqrt{3} \end{cases} \tag{5}$$

2.2. Выборочные числовые характеристики

Вариационный ряд - последовательность элементов выборки, расположенных в неубывающем порядке.

2.2.1. Характеристики положения

- Выборочное среднее

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{6}$$

- Выборочная медиана

$$medx = \begin{cases} x_{l+1}, n = 2l+1 \\ \frac{x_l + x_{l+1}}{2}, n = 2l \end{cases} \tag{7}$$

- Полусумма экстремальных выборочных элементов

$$z_R = \frac{x_1 + x_n}{2} \tag{8}$$

- Полусумма квартилей
  Выборочная квартиль z_p порядка p определяется формулой

$$z_p = \begin{cases} x_{[np]+1}, np - \text{дробное} \\ x_{np}, np - \text{целое} \end{cases} \tag{9}$$

  Полусумма квартилей

$$z_Q = \frac{z_{1/4} + z_{3/4}}{2} \tag{10}$$

- Усечённое среднее

$$z_{tr} = \frac{1}{n-2r}\sum_{i=r+1}^{n-r} x_i, r \approx \frac{n}{4} \tag{11}$$

2.2.2. Характеристики рассеивания

Выборочная дисперсия определяется по формуле:

$$D = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})^2 \tag{12}$$

3. Реализация

Лабораторная работа выполнена на языке Python в виртуальной среде Anaconda с интерпретатором версии 3.9 в среде разработки Visual Studio Code. Дополнительные зависимости:

- scipy
- numpy

4. Результаты

	$\overline{x}$	$med(x)$	z_R	z_Q	z_{tr}
size 10					
$E(z)$	-0.008474	-0.011502	-0.004567	0.313433	0.26929
$D(z)$	0.092723	0.133581	0.185508	0.117345	0.105668
$E+\sqrt{D}$	0.296032	0.353986	0.426139	0.655989	0.594356
$E-\sqrt{D}$	-0.312979	-0.37699	-0.435273	-0.029124	-0.055777
Estimation	0	0	0	0	0
size 100					
$E(z)$	0.001189	0.001817	0.001457	0.016332	0.030168
$D(z)$	0.010864	0.016026	0.10168	0.013248	0.012352
$E+\sqrt{D}$	0.105419	0.12841	0.32033	0.131432	0.141308
$E-\sqrt{D}$	-0.10304	-0.124777	-0.317415	-0.098768	-0.080973
Estimation	0	0	0	0	0
size 1000					
$E(z)$	-0.000752	-0.000457	-0.008089	0.000512	0.00199
$D(z)$	0.001038	0.001648	0.060405	0.001275	0.001246
$E+\sqrt{D}$	0.031466	0.040137	0.237685	0.036224	0.037286
$E-\sqrt{D}$	-0.03297	-0.041052	-0.253863	-0.0352	-0.033306
Estimation	0.0	0.0	0	0.0	0.0

Таблица 1. Нормальное распределение (1)

	$\overline{x}$	$med(x)$	z_R	z_Q	z_{tr}
size 10					
$E(z)$	-0.79855	-0.00035	-4.101512	1.19138	0.698774
$D(z)$	1317.311589	0.350235	32666.76735	15.821066	2.479704
$E+\sqrt{D}$	35.496238	0.591457	176.637989	5.16895	2.273481
$E-\sqrt{D}$	-37.093337	-0.592156	-184.841013	-2.78619	-0.875934
Estimation	-	0	-	0	0
size 100					
$E(z)$	1.438919	0.007712	70.252783	0.049879	0.049742
$D(z)$	517.855301	0.026	1282426.145731	0.055577	0.027417
$E+\sqrt{D}$	24.195353	0.168957	1202.69534	0.285628	0.215322
$E-\sqrt{D}$	-21.317516	-0.153532	-1062.189775	-0.18587	-0.115837
Estimation	-	0	-	0	0
size 1000					
$E(z)$	17.893711	0.00024	8923.594175	0.005589	0.006013
$D(z)$	330963.684932	0.002206	82722241630.99397	0.004493	0.002395
$E+\sqrt{D}$	593.188144	0.047207	296538.33933	0.072622	0.054951
$E-\sqrt{D}$	-557.400722	-0.046726	-278691.150981	-0.061443	-0.042925
Estimation	-	0.0	-	0.0	0.0

Таблица 2. Распределение Коши(2)

	$\overline{x}$	$med(x)$	z_R	z_Q	z_{tr}
size 10					
$E(z)$	-0.007129	0.004394	-0.046904	0.29745	0.235897
$D(z)$	0.097945	0.071578	0.414221	0.113508	0.080113
$E+\sqrt{D}$	0.305832	0.271935	0.596696	0.63436	0.51894
$E-\sqrt{D}$	-0.32009	-0.263147	-0.690503	-0.039459	-0.047146
Estimation	0	0	0	0	0
size 100					
$E(z)$	-0.002896	-0.001324	0.003334	0.011611	0.017572
$D(z)$	0.008925	0.005687	0.419663	0.009482	0.005956
$E+\sqrt{D}$	0.091576	0.074088	0.651147	0.108986	0.094748
$E-\sqrt{D}$	-0.097368	-0.076737	-0.64448	-0.085765	-0.059604
Estimation	0.0	0.0	0	0	0.0
size 1000					
$E(z)$	-0.001022	-5.2e-05	0.008371	0.001062	0.002033
$D(z)$	0.000988	0.000519	0.386344	0.001015	0.000612
$E+\sqrt{D}$	0.030407	0.022729	0.629937	0.032921	0.026767
$E-\sqrt{D}$	-0.03245	-0.022833	-0.613195	-0.030797	-0.0227
Estimation	0.0	0.0	0	0.0	0.0

Таблица 3. Распределение Лапласа(3)

	$\overline{x}$	$med(x)$	z_R	z_Q	z_{tr}
size 10					
$E(z)$	10.0024	9.8395	10.3185	10.937	10.766333
$D(z)$	1.083954	1.50099	1.868308	1.477031	1.348122
$E+\sqrt{D}$	11.043531	11.064649	11.685361	12.152332	11.92742
$E-\sqrt{D}$	8.961269	8.614351	8.951639	9.721668	9.605247
Estimation	10 ± 1	10 ± 1	10 ± 2	10 ± 2	10 ± 1
size 100					
$E(z)$	9.9982	9.8465	10.9	9.9605	9.94652
$D(z)$	0.100821	0.215688	1.0055	0.16469	0.124796
$E+\sqrt{D}$	10.315724	10.310922	11.902746	10.36632	10.299785
$E-\sqrt{D}$	9.680676	9.382078	9.897254	9.55468	9.593255
Estimation	10 ± 1	10 ± 1	10 ± 2	10 ± 2	10 ± 1
size 1000					
$E(z)$	10.001136	9.9965	11.6705	9.9945	9.86845
$D(z)$	0.009839	0.003238	0.69518	0.00422	0.011399
$E+\sqrt{D}$	10.100328	10.053401	12.504274	10.05946	9.975218
$E-\sqrt{D}$	9.901944	9.939599	10.836726	9.92954	9.761682
Estimation	10 ± 1	10 ± 1	10 ± 2	10 ± 2	10 ± 1

Таблица 4. Распределение Пуассона(4)

	$\overline{x}$	$med(x)$	z_R	z_Q	z_{tr}
size 10					
$E(z)$	-0.010868	-0.020561	-0.001797	0.309553	0.297099
$D(z)$	0.100284	0.229497	0.042455	0.11875	0.152572
$E+\sqrt{D}$	0.305809	0.458498	0.20425	0.654154	0.687703
$E-\sqrt{D}$	-0.327545	-0.49962	-0.207843	-0.035049	-0.093505
Estimation	0	0	0	0	0
size 100					
$E(z)$	-0.000405	0.001731	0.000333	0.016786	0.034421
$D(z)$	0.010312	0.029023	0.000565	0.015811	0.02071
$E+\sqrt{D}$	0.101145	0.172091	0.024098	0.142527	0.178331
$E-\sqrt{D}$	-0.101954	-0.16863	-0.023432	-0.108955	-0.109488
Estimation	0	0	0.0	0	0
size 1000					
$E(z)$	0.001586	0.002118	2.8e-05	0.003339	0.005729
$D(z)$	0.00095	0.002862	6e-06	0.001452	0.001921
$E+\sqrt{D}$	0.032404	0.055621	0.002391	0.041442	0.049564
$E-\sqrt{D}$	-0.029233	-0.051384	-0.002335	-0.034763	-0.038105
Estimation	0.0	0.0	0.00	0.0	0.0

Таблица 5. Равномерное распределение(5)

5. Обсуждение

Полученные данные показывают, что выборки из большего количества элементов лучше уточняют значение характеристик случайной величины.

Для нормального, равномерного распределения и распределения Лапласа эти значения схожи и близки к нулю. У распределение Пуассона среднее значение $E(z)$ во всех выборках близко к 10, это значение параметра задания данного распределения. В характеристиках распределения Коши появляются аномально большие значения, это может объясняться неопределенностью математического ожидания и бесконечностью дисперсии случайной величины, распределенной по закону Коши.