Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

ОТЧЁТ ПО ЛАБОРАТОРНЫМ РАБОТАМ ПО ДИСЦИПЛИНЕ **«МАТЕМАТИЧЕСКАЯ СТАТИСТИКА»**

Выполнил студент
Войнова Алёна Игоревна
группы 3630102/80201

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

Санкт-Петербург
2021

Содержание

1 Постановка задачи

Для 5 распределений:

1. $N(x, 0, 1)$ – нормальное распределение
2. $C(x, 0, 1)$ – распределение Коши
3. $L(x, 0, \frac{1}{\sqrt{2}})$ – распределение Лапласа
4. $P(k, 10)$ – распределение Пуассона
5. $U(x, -\sqrt{3}, \sqrt{3})$ – равномерное распределение

1.1 Задание 1

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

$$E(z) = \overline{z}(1) \tag{1}$$

Вычислить оценку дисперсии по формуле:

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

Представить полученные данные в виде таблиц.

2 Теория

2.1 Распределения

- Нормальное распределение

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

- Распределение Коши

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

- Распределение Лапласа

$$L(x, 0, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|}(3) \tag{5}$$

- Распределение Пуассона

$$P(k, 10) = \frac{10^k}{k!} e^{-10}(4) \tag{6}$$

- Равномерное распределение

$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}} & |x| \le \sqrt{3} \\ 0 & |x| > \sqrt{3} \end{cases} (5) \tag{7}$$

2.1.1 Выборочные числовые характеристики

С помощью выборки образуются её числовые характеристики. Это числовые характеристики дискретной случайной величины $x_1, x_2, ..., x_n$ [1, с. 411].

2.1.2 Характеристики положения

- Выборочное среднее

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{8}$$

- Выборочная медиана

$$medx = \begin{cases} x_{(l+1)} & \text{при } n = 2l+1, \\ \frac{x_{(l)}+x_{(l+1)}}{2} & \text{при } n = 2l. \end{cases} \tag{9}$$

- Полусумма экстремальных выборочных элементов

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

- Полусумма квартилей

  Выборочная квартиль z_p порядка p определяется формулой

$$z_p = \begin{cases} x_{([np]+1)} & \text{при } np \text{ дробном,} \\ x_{(np)} & \text{при } np \text{ целом.} \end{cases} \tag{11}$$

  Полусумма квартилей

$$z_Q = \frac{z_{1/4} + z_{3/4}}{2} \tag{12}$$

- Усечённое среднее

$$z_{tr} = \frac{1}{n-2r}\sum_{i=r+1}^{n-r} x_i, \text{ где } r \approx \frac{n}{4} \tag{13}$$

2.1.3 Характеристики рассеяния

Выборочная дисперсия

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

3 Реализация

Лабораторная работа выполнена с помощью средств языка программирования **Python** в среде разработки **Jupyter**. Исходный код лабораторной работы приведён в приложении.

4 Результаты

4.1 Характеристики положения и рассеяния

Мощность выборки указана в первом столбике справа.

Normal	$\bar{x}$	$medx$	z_R	z_Q	z_{tr}
E(z) = 10	0.0016	-0.0011	0.0085	-0.0049	0.0055
D(z) = 10	0.0997	0.1425	0.1875	0.1109	0.1647
E(z) = 100	-0.0053	-0.0018	-0.0083	-0.0045	-0.0068
D(z) = 100	0.0093	0.0147	0.0903	0.0114	0.0191
E(z) = 1000	0.0002	0.0004	0.0027	0.0009	0.0007
D(z) = 1000	0.001	0.0016	0.0581	0.0012	0.002

Таблица 1: Нормальное распределение

Cauchy	$\bar{x}$	$medx$	z_R	z_Q	z_{tr}
E(z) = 10	1.2235	0.0016	6.1494	-0.0196	2.5435
D(z) = 10	907.4583	0.3132	22469.7157	0.8644	1527.8722
E(z) = 100	1.9565	0.0119	97.4769	0.0166	4.4856
D(z) = 100	6060.7936	0.0251	15128110.8008	0.0555	23401.4874
E(z) = 1000	0.1619	0.0017	75.7176	0.0024	0.055
D(z) = 1000	74.9162	0.0026	16842786.6595	0.0046	107.6337

Таблица 2: Распределение Коши

laplace	x_	med(x)	z_R	z_Q	z_tr
E(z) = 10	-0.0068	0.0002	-0.019	-0.0001	-0.0063
D(z) = 10	0.1053	0.0748	0.4226	0.0959	0.1765
E(z) = 100	-0.0031	-0.0024	-0.0262	-0.0022	-0.005
D(z) = 100	0.0098	0.0059	0.4181	0.0097	0.0189
E(z) = 1000	-0.001	-0.0008	0.0144	-0.0005	-0.0036
D(z) = 1000	0.0009	0.0005	0.4368	0.0009	0.0018

Таблица 3: Распределение Лапласа

poisson	x_	med(x)	z_R	z_Q	z_tr
E(z) = 10	10.0278	9.869	10.304	9.9448	10.017
D(z) = 10	0.9922	1.3688	1.9806	1.1249	1.5921
E(z) = 100	9.9914	9.839	10.9265	9.9001	9.9863
D(z) = 100	0.1014	0.2176	0.9883	0.1489	0.2109
E(z) = 1000	9.9977	9.9915	11.703	9.9914	9.9939
D(z) = 1000	0.0102	0.0077	0.6508	0.0042	0.0211

Таблица 4: Распределение Пуассона

uniform	x_	med(x)	z_R	z_Q	z_tr
E(z) = 10	0.0128	0.0213	0.0061	0.0171	0.0108
D(z) = 10	0.101	0.2308	0.0475	0.1384	0.1747
E(z) = 100	-0.0049	-0.005	-0.0016	-0.0063	-0.0079
D(z) = 100	0.0097	0.029	0.0006	0.0145	0.021
E(z) = 1000	-0.0002	-0.0006	-0.0	-0.0007	-0.0001
D(z) = 1000	0.001	0.0029	0.0	0.0015	0.002

Таблица 5: Равномерное распределение

5 Обсуждение

Проанализировав полученные результаты, можно заметить, что для нормального распределения, распределения Лапласа и равномерного распределения (z) и $D(z)$ для всех характеристик уменьшаются с ростом выборки.
В распеределении Пуассона значения $E(z)$ для всех характеристик колеблется в районе 10, но в $D(z)$, аналогично рассмотренным выше распределениям , наблюдается уменьшение значений при росте выборки.
Особо выделяется распределение Коши. Значения $D(z)$ для $\bar{x}$ и z_R, $E(z)$ для z_R достигают больших порядков. Такое поведение дисперсии 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 Приложения

URL: Выполненная лабораторная работа на GitHub
https://github.com/pikabol88/Math-Statistics/blob/main/Lab2.ipyn