Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

Отчёт
по лабораторной работе №2
по дисциплине
«Математическая статистика»

Выполнил студент:

Кондратьев Д. А.
группа: 3630102/70301

Проверил:

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

Санкт-Петербург
2020 г.

Содержание

Список таблиц

1. Постановка задачи

Для 5-ти рапределений:

- 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 $U(x, -\sqrt{3}, \sqrt{3})$;

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

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

Вычислить оценку дисперсии по формуле:

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

Представить полученные данные в виде таблиц.

2. Теория

2.1. Распределения

- Нормальное распределение

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

- Распределение Коши

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

- Распределение Лапласа

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

- Распределение Пуассона

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

- Равномерное Распределение

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

2.2. Вариационный ряд

Вариационным рядом называется последовательность элементов выборки, расположенных в неубывающем порядке. Одинаковые элементы повторяются [1, с. 409].

Запись вариационного ряда: $x_1, x_2, ..., x_n$.

Элементы вариационного ряда x_i $(i = 1, 2, ..., n)$ 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 $x_1, x_2, ..., x_n$ [1, с. 411].

2.3.1. Характеристики положения

- Выборочное среднее:

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

- Выборочная медиана:

$$med\ x = \begin{cases} x_{k+1}, & n = 2k+1 \\ \frac{1}{2}\left(x_k + x_{k+1}\right), & n = 2k \end{cases} \tag{9}$$

- Полусумма экстремальных выборочных элементов:

$$z_R = \frac{1}{2}\left(x_1 + x_n\right) \tag{10}$$

- Полусумма квартилей:

$$z_Q = \frac{1}{2}\left(z_{\frac{1}{4}} + z_{\frac{3}{4}}\right) \tag{11}$$

- Усечённое среднее:

$$z_{tr} = \frac{1}{n-2r}\sum_{i=r+1}^{n-r} x_i,\ r \approx \frac{n}{4} \tag{12}$$

2.3.2. Характеристики рассеяния

Выборочная дисперсия:

$$D = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{13}$$

3. Реализация

Лабораторная работа выполнена на программном языке *Python 3.8* в среде разработки *Jupyter Notebook 6.0.3*. В работе использовались следующие пакеты языка *Python*:

- *numpy* — для генерации выборки и работы с массивами;
- *scipy.stats* — содержит все необходимые распределения.

Ссылка на исходный код лабораторной работы приведена в приложении.

4. Результаты

4.1. Характеристики положения и рассеяния

Таблица 1. Нормальное распределение

	$n = 10$		$n = 100$		$n = 1000$	
	$E(x)$ (1)	$D(x)$ (2)	$E(x)$	$D(x)$	$E(x)$	$D(x)$
$\overline{x}$ (8)	0.00	0.0985	0.00	0.0107	-0.002	0.0010
$med\ x$ (9)	0.0	0.1337	0.00	0.0149	-0.001	0.0016
z_R (10)	0.0	0.1847	0.00	0.0928	-0.01	0.0630
z_Q (11)	0.0	0.1177	-0.01	0.0132	-0.004	0.0012
z_{tr}(12)	0.0	0.1113	0.00	0.0123	-0.002	0.0012

Таблица 2. Распределение Коши

	$n = 10$		$n = 100$		$n = 1000$	
	$E(x)$	$D(x)$	$E(x)$	$D(x)$	$E(x)$	$D(x)$
$\overline{x}$	2.9	8261.4971	-0.3	282.5623	-0.7	1680.6291
$med\ x$	0.00	0.3821	-0.01	0.0260	-0.002	0.0024
z_R	14.2	206063	-10.5	693687	-358.5	417271337
z_Q	0.0	1.2177	-0.03	0.0493	-0.003	0.0049
z_{tr}	0.0	0.5952	0.00	0.0261	-0.001	0.0026

Таблица 3. Распределение Лапласа

	$n = 10$		$n = 100$		$n = 1000$	
	$E(x)$	$D(x)$	$E(x)$	$D(x)$	$E(x)$	$D(x)$
$\overline{x}$	-0.02	0.0897	-0.002	0.0099	0.002	0.001
$med\ x$	-0.02	0.0648	0.003	0.0058	0.0003	0.0005
z_R	0.0	0.3887	0.0	0.409	0.0	0.4084
z_Q	-0.01	0.0943	-0.01	0.01	-0.001	0.001
z_{tr}	-0.01	0.0662	0.001	0.0063	0.0007	0.0006

Таблица 4. Распределение Пуассона

	$n = 10$		$n = 100$		$n = 1000$	
	$E(x)$	$D(x)$	$E(x)$	$D(x)$	$E(x)$	$D(x)$
$\overline{x}$	10.02	1.0132	10.00	0.0976	9.994	0.0098
$med\ x$	9.9	1.4617	9.9	0.1944	9.995	0.005
z_R	10.4	1.9290	10.95	1.0299	11.8	0.6673
z_Q	9.9	1.2086	9.9	0.1590	9.995	0.0037
z_{tr}	9.9	1.1459	9.9	0.1166	9.85	0.0107

Таблица 5. Равномерное распределение

	$n = 10$		$n = 100$		$n = 1000$	
	$E(x)$	$D(x)$	$E(x)$	$D(x)$	$E(x)$	$D(x)$
$\overline{x}$	-0.02	0.0893	0.000	0.0099	0.001	0.001
$med\ x$	0.0	0.2088	0.01	0.0294	0.001	0.003
z_R	-0.01	0.0392	0.0002	0.0006	0.0001	0.0000
z_Q	0.0	0.1282	-0.02	0.0146	-0.001	0.0014
z_{tr}	0.0	0.1479	0.00	0.0200	0.001	0.0019

5. Обсуждение

Исходя из полученных результатов можно сделать следующие выводы:

- Дисперсия может гарантировать порядок точности среднего значения только до первой значащей цифры после запятой в дисперсии включительно. Поэтому есть необходимость проверки данных результатов на совпадение этой точности.

- При увеличении количества выборки гарантируемая точность будет возрастать.

- Предыдущие выводы касаются всех распределений, кроме распределения Коши. Так как оно имеет бесконечную дисперсию и, следовательно, никакой точности гарантировать не может.

- Можно вывести следующее отношение для характеристик положения при $n = 1000$:

    1) для нормального распределения: $z_R < z_Q < z_{tr} \leq \overline{x} < med\ x$;
    2) для распределения Коши: $z_R < \overline{x} < z_Q < med\ x < z_{tr}$;
    3) для распределения Лапласа: $z_Q < z_R < med\ x < z_{tr} < \overline{x}$;
    4) для распределения Пуассона: $z_{tr} < \overline{x} < z_Q \leq med\ x < z_R$;
    5) для равномерного распределения: $z_Q < z_R < z_{tr} \leq med\ x \leq \overline{x}$.

6. Литература

1) **Вероятностные разделы математики.** Учебник для бакалавров технических направлений.//Под ред. Максимова Ю.Д. — Спб.: «Иван Федоров», 2001. — 592 с., илл.

7. Приложение

1) Код лабораторной. URL: `https://github.com/DmitriiKondratev/MatStat/blob/master/Lab_2/Lab_2.ipynb`

2) Код отчёта. URL: `https://github.com/DmitriiKondratev/MatStat/blob/master/Lab_2/Lab_report_2.tex`