Санкт-Петербургский Политехнический университет

Петра Великого

Институт Прикладной Математики и Механики

Кафедра «Прикладная Математика и Информатика»

Отчет

По лабораторной работе № 2

По Дисциплине «Математическая статистика»

Выполнил:

Студент Селянкин Федор

Группа 3630102/70301

Проверил:

к.ф. – м.н., доцент

Баженов Александр Николаевич

Санкт-Петербург
2020 г.

Содержание

Список таблиц:

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

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:

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

Вычислить оценку дисперсии по формуле:

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

Представить полученные данные в виде таблиц.

Теория

Распределения

- Нормальное распределение

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

- Распределение Коши

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

- Распределение Лапласа

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

- Распределение Пуассона

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

- Равномерное распределение

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

Статистические характеристики

- Выборочное среднее

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

- Выборочная медиана

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

- Полусумма экстремальных выборочных элементов

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

- Полусумма квартилей

  Выборочная квартиль z_p порядка p определяется формулой.

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

Полусумма квартилей

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

- Усеченное среднее

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

Характеристики рассеяния

- Выборочная дисперсия

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

Реализация

Лабораторная работа выполнена с помощью встроенных средств языка программирования Python в среде разработки PyCharm, с использованием дополнительных библиотек для отображения и расчетов. Исходный код лабораторной выложен на веб-сервисе GitHub [2].

Результаты

Нормальное распределение

n		$\overline{x}$	$med\ x$	z_R	z_Q	z_{tr}
10	E	0.01	0	0	-0.3	0
	D	0.096971	0.142454	0.193487	0.113430	0.166238
100	E	0	0	0	-0.01	0
	D	0.010519	0.015980	0.103325	0.012640	0.011316
1000	E	0	0	0	0	0
	D	0.000981	0.001644	0.062102	0.001258	0.001219
$\overline{x} < z_{tr} < z_Q < med\ x < z_R$						

Таблица 1 Нормальное распределение

Распределение Коши

n		$\overline{x}$	$med\ x$	z_R	z_Q	z_{tr}
10	E	0	0	0	-1	0
	D	70974.612820	0.306827	31123.008220	5.088886	0.833265
100	E	0	0	0	-0.04	0
	D	340057.722086	0.025110	4790055.844103	0.050678	0.025347
1000	E	0	0	0	0	0
	D	3152.073514	0.002294	36577755.775665	0.005139	0.002566
$med\ x < z_{tr} < z_Q < \overline{x} < z_R$						

Таблица 2 Распределение Коши

Распределение Лапласа

n		$\overline{x}$	$med\ x$	z_R	z_Q	z_{tr}
10	E	0	0	0	-0.3	0
	D	0.101021	0.076025	0.446803	0.123058	0.098940
100	E	0	0	0	-0.02	0
	D	0.010233	0.005696	0.363117	0.010345	0.005857
1000	E	0	0	0	-0.002	-0.0007
	D	0.001074	0.000531	0.390686	0.001000	0.000597
$med\ x < z_{tr} < z_Q < \overline{x} < z_R$						

Таблица 3Распределение Лапласа

Распределение Пуассона

n		$\overline{x}$	$med\ x$	z_R	z_Q	z_{tr}
10	E	10.0	10.0	10.0	10.0	11.0
	D	1.008408	1.484790	1.740384	1.231344	1.469999
100	E	10.0	9.9	10.0	10.0	9.9
	D	0.099818	0.224111	0.955558	0.156679	0.118266
1000	E	10.0	10.000000	11.0	9.9	9.9
	D	0.009958	0.000000	0.663388	0.004436	0.011769
$med\ x < z_Q < \overline{x} < z_{tr} < z_R$						

Таблица 4Распределение Пуассона

Равномерное распределение

n		$\overline{x}$	$med\ x$	z_R	z_Q	z_{tr}
10	E	0	0	0	-0.31	0
	D	0.099079	0.227962	0.043533	0.146461	0.228256
100	E	0	0	0	-0.015	0
	D	0.010254	0.028950	0.000637	0.013887	0.019951
1000	E	0.0011	0	-0.000076	-0.002	0
	D	0.000970	0.002742	0.000006	0.001472	0.001991
$z_R < \overline{x} < z_Q < z_{tr} < med\ x$						

Таблица 5 Равномерное распределение

Литература

1. Вероятностные разделы математики. Учебник для бакалавров технических направлений.//Под ред. Максимова Ю.Д. — Спб.: .Иван Федоров.,2001. — 592 с., илл.
2. Ссылка на репозиторий GitHub https://github.com/SelyankinFyodor/math-statistics/tree/master/Lab2
3. Распределение Коши https://en.wikipedia.org/wiki/Cauchy_distribution

Обсуждения

С увеличением мощности выборки величины среднее и усеченное среднее имеют все меньшее рассеяние (с увеличением мощности в 10 раз, дисперсия уменьшается в 10 раз)[3]. Для Распределения Коши в качестве характеристики положения не разумно брать среднее или полусумму экстремалей.