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
В. А. Рыженко

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

Санкт-Петербург, 2020 г.

Содержание

Список иллюстраций

1. Постановка задачи

Для 5 распределений:

- Нормальное распределение N(x, 0, 1)
- Распределение Коши C(x, 0, 1)
- Распределение Лапласа L(x, 0, $\frac{1}{\sqrt{2}}$)
- Постановка задач исследованияРаспределение 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: $\overline{x}$ (8), $med\, x$ (9), z_R (10), z_Q (12), z_tr (13). 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

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

- Распределение Лапласа

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

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

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

2.2. Характеристики положения

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

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

- Полусумма экстремальных выборочных элементов
$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \tag{10}$$

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

2.3. Характеристики рассеяния

Выборочная дисперсия

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

3. Реализация

Лабораторная работа выполнена с помощью встроенных средств языка программирования Python в среде разработки Jupyter Notebook и Visual Code. Исходный код лабораторной работы приведён в приложении.

4. Результаты

4.1. Таблицы значений

Normal n = 10					
	$\overline{x}$ (8)	$medx$ (8)	z_R (10)	z_Q (12)	z_{tr} (13)
$E(z)$ (1)	-0.01	0.0	0.0	0.0	-0.4
$D(z)$ (2)	0.09935	0.13874	0.18181	0.11633	0.19166
Normal n = 100					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.01	-0.01	0.0	0.00	-0.53
$D(z)$	0.05471	0.07758	0.13405	0.06471	0.11674
Normal n = 1000					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.00	-0.01	0.0	0.01	-0.57
$D(z)$	0.03683	0.05231	0.11162	0.04357	0.08085

Таблица 1. Нормальное распределение

Cauchy n = 10					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	4	0	23	0	-4
$D(z)$	15908.09147	0.43148	397531.77972	1.26249	523.04326
Cauchy n = 100					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	1	0.0	-18	0.0	-7
$D(z)$	8637.35792	0.22896	1880864.30928	0.65602	2509.25432
Cauchy n = 1000					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	1	0.0	190	0.0	-6
$D(z)$	6106.00336	0.15343	87773120.69259	0.43918	1712.09914

Таблица 2. Распределение Коши

Laplace n = 10					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.0	0.00	0.0	0.0	-0.4
$D(z)$	0.10735	0.08220	0.41080	0.11063	0.18386
Laplace n = 100					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.00	0.00	0.0	0.00	-0.5
$D(z)$	0.05890	0.04395	0.40008	0.06062	0.11308
Laplace n = 1000					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.0	0.0	0.0	0.00	-0.53
$D(z)$	0.03961	0.02948	0.40734	0.04075	0.07833

Таблица 3. Распределение Лапласа

Poisson n = 10					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	10.0	10	10	10	12
$D(z)$	0.94092	1.30249	1.90001	1.19398	1.61149
Poisson n = 100					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	10.0	9.9	11	10.0	12
$D(z)$	0.51966	0.74990	1.51329	0.67369	1.14114
Poisson n = 1000					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	10.0	9.9	11	10.0	12.6
$D(z)$	0.34978	0.50642	1.43907	0.45051	0.81966

Таблица 4. Распределение Пуассона

Uniform n = 10					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.0	0.0	0.01	0.0	-0.4
$D(z)$	0.10099	0.23114	0.04599	0.13631	0.22124
Uniform n = 100					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.00	0.0	0.00	0.01	-0.5
$D(z)$	0.05538	0.13012	0.02329	0.07517	0.13984
Uniform n = 1000					
	$\overline{x}$	$medx$	z_R	z_Q	z_{tr}
$E(z)$	0.00	0.00	0.00	0.01	-0.57
$D(z)$	0.03726	0.08778	0.01553	0.05062	0.09711

Таблица 5. Равномерное распределение

4.2. Упорядоченные характеристики положения

1) Нормальное распределение:

$$z_{tr} < med\,x < \overline{x} = z_R < z_Q \tag{15}$$

2) Распределение Коши:

$$z_{tr} < med\,x < \overline{x} = z_Q < z_R \tag{16}$$

3) Распределение Лапласа:

$$z_{tr} < \overline{x} = med\,x = z_R = z_Q \tag{17}$$

4) Распределение Пуассона:

$$med\,x < \overline{x} = z_Q < z_R < z_{tr} \tag{18}$$

5) Равномерное распределение:

$$z_{tr} < \overline{x} = med\,x = z_R < z_Q \tag{19}$$

5. Обсуждение

Из полученных данных видно, что среднее (1) всех характеристик стремится к теоретическому, а оценка дисперсии (2) к нулю при увеличении размера выборки. В случае распределения Коши (4) это верно только для медианы (9).

6. Приложения

Репозиторий на GitHub с релизацией: github.com.