Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

Отчёт
по лабораторной работе №1
по дисциплине
«Математическая статистика»

Выполнил студент:
Самутичев Евгений Романович
группа: 3630102/70201

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

Санкт-Петербург
2020 г.

Содержание

Список иллюстраций

1 Постановка задачи

Для каждого из 5 распределений:

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

выборку размера: 10, 100, 1000 - сгенерировать 1000 раз, для каждой генерации произвести вычисления выборочных характеристик $\bar{x}, \mathrm{med}\ x, z_R, z_Q, z_{tr}$ для всех генераций в рамках одного размера выборки получить значения среднего характеристик положения:

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

и оценку дисперсии:

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

Представить полученные данные в виде таблиц.

2 Теория

2.1 Вариационный ряд

Если элементы выборки $x_1, ..., x_n$ упорядочить по возрастанию на каждом элементарном исходе (рассматриваем их как случайные величины), получится новый набор случайный величин, называемый *вариационным рядом*:

$$x_{(1)} \leq ... \leq x_{(n)}$$

Элемент $x_{(k)}$ называется *k-ой порядковой статистикой.*

2.2 Выборочные характеристики

При работе с выборкой нам неизвестно распределение по которому она получена, а значит и соответствующие характеристики распределения. Однако, существуют оценки - т.н. *выборочные характеристики*:

- Выборочное среднее
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3}$$

- Выборочная медиана
$$\text{med } x = \begin{cases} x_{(k+1)} & \text{при } n = 2k+1 \\ \frac{x_{(k)}+x_{(k+1)}}{2} & \text{при } n = 2k \end{cases} \tag{4}$$

- Полусумма экстремальных выборочных элементов
$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \tag{5}$$

- Выборочный квантиль уровня α
$$z_\alpha = \frac{x_{(\lfloor q \rfloor + 1)} + x_{(\lceil q \rceil + 1)}}{2}, \text{где } q = (n-1)\alpha \tag{6}$$

  формула, используемая в **NumPy**, в этом случае $z_0 = \min\limits_{i=1,...,n} x_{(i)}, z_1 = \max\limits_{i=1,...,n} x_{(i)}$, $z_{0.5} = \text{med}\, x$

- Полусумма квантилей
$$z_Q = \frac{z_{0.25} + z_{0.75}}{2} \tag{7}$$

- Усеченное среднее
$$z_{tr} = \frac{1}{n-2r}\sum_{i=r+1}^{n-r} x_{(i)} \tag{8}$$

Выборочные характеристики являются случайными величинами, поэтому в работе и производится усреднение их значений для 1000 генераций и вычисление среднеквадратичного отклонения.

3 Реализация

Работа выполнена с использованием языка **Python** в интегрированной среде разработки **PyCharm**, были задействованы библиотеки:

- **NumPy** - построение вариационного ряда и вычисления
- **SciPy** - модуль **stats** для генерации данных по распределениям

Исходный код работы приведен в приложении.

4 Результаты

5 Обсуждение

6 Приложения

1. Исходный код лабораторной `https://github.com/zhenyatos/statlabs/tree/master/Lab2`

Список литературы

[1] А. Н. Ширяев, *Вероятность-1*. Изд. МЦНМО, Москва, 2017. 551 стр.

[2] Н. И. Чернова, *Математическая статистика: Учеб. пособие*. Новосиб. гос. ун-т. Новосибирск, 2007. 148 стр.