

Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

Отчет по лабораторной работе № 3
по дисциплине: Математическая статика.

Выполнила студентка:
Заболотских Екатерина Дмитриевна
группа: 3630102/70301

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

Санкт-Петербург
2020 г.

Оглавление

Постановка задачи.....	2
Теория.....	3
1. Диаграмма размахов (« ящик с усами »)	3
Определение.....	3
Построение	3
2. Выбросы	4
Определение.....	4
Теоретическая вероятность выбросов.....	4
Реализация	5
Результаты.....	6
Теоретическая доля выбросов.....	9
Доля выбросов	9
Список литературы.....	10

Список иллюстраций

Рисунок 1: Нормальное распределение.....	6
Рисунок 2: Распределение Коши.....	6
Рисунок 3: Распределение Лапласа	7
Рисунок 4: Распределение Пуассона	7
Рисунок 5: Равномерное распределение	8

Список таблиц

Таблица 1: теоретическая вероятность выбросов.....	9
Таблица 2: доля выбросов	9

Постановка задачи

Для каждого из 5 распределений:

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

Сгенерировать выборки размеров: 10, 50, 1000; и построить графики плотности распределения вероятности и гистограмму на одном рисунке.

Сгенерировать выборки размером 20, 100 элементов и построить для них боксплот Тьюки.

Для каждого распределения экспериментально определить долю выбросов (сгенерировав выборку, соответствующую распределению, 1000 раз, и вычислить среднюю долю выбросов) и сравнить с результатами, полученными теоретически.

Теория

1. Диаграмма размахов (« ящик с усами »)

Определение

— это график, позволяющий дать статистическую характеристику анализируемой совокупности. Графики этого типа очень популярны, поскольку позволяют дать очень полную статистическую характеристику анализируемой совокупности.

Построение

Чтобы нарисовать ящик для одной группы про исходные данные необходимо знать всего три характеристики:

- Первый квартиль: $Q_{25} = X_{[1/4]}$
- Медиану: $Q_{50} = X_{[1/2]}$
- Третий квартиль $Q_{75} = X_{[3/4]}$

Диаграммы размахов, или "ящички с усами", получили свое название за характерный вид: границами ящика служат первый и третий квартили, линия в середине ящика – медиана. Концы усов – края статистически значимой выборки (без выбросов). Длину «усов» определяют разность первого квартиля и полутора межквартильных расстояний и сумма третьего и полутора межквартильных расстояний. Формула имеет вид:

$$X_1 = Q_{25} - \frac{3}{2}(Q_{75} - Q_{25}), \quad (1)$$

$$X_2 = Q_{75} + \frac{3}{2}(Q_{75} - Q_{25}), \quad (2)$$

где X_1 – нижняя граница уса, X_2 – верхняя граница уса.

Данные, выходящие за границы усов (выбросы), отображаются на графике в виде маленьких кружков.

2. Выбросы

Определение

Выброс – результат измерения, выделяющийся из выборки. Если элемент выборки не лежит в диапазоне $[X_1, X_2]$, то это и есть выброс.

Теоретическая вероятность выбросов

Для непрерывных распределений:

$$P_B^T = P(x < X_1^T) + P(x > X_2^T) = F(X_1^T) + (1 - F(X_2^T)) \quad (3)$$

Для дискретных с учетом возможного скачка:

$$P_B^T = P(x < X_1^T) + P(x > X_2^T) = (F(X_1^T) - P(x = X_1^T)) + (1 - F(X_2^T)) \quad (4)$$

Где $F(X) = P(x \leq X)$ – функция распределения.

Реализация

Код программы, реализующий данную задачу, был написан на языке Python в интегрированной среде разработки PyCharm.

Были использованы библиотеки:

- **Numpy** – библиотека для работы с данными.
- **Matplotlib** – вывод графиков.

Результаты

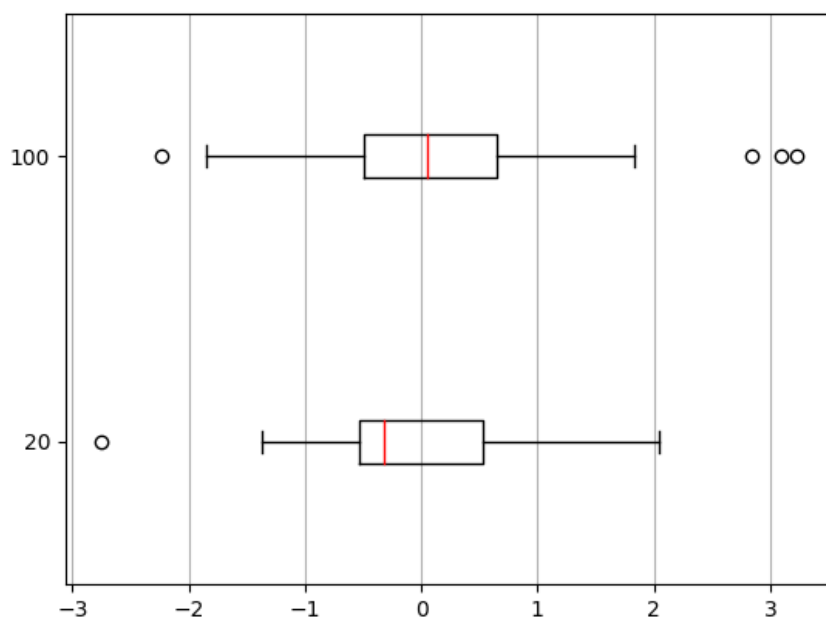


Рисунок 1: Нормальное распределение

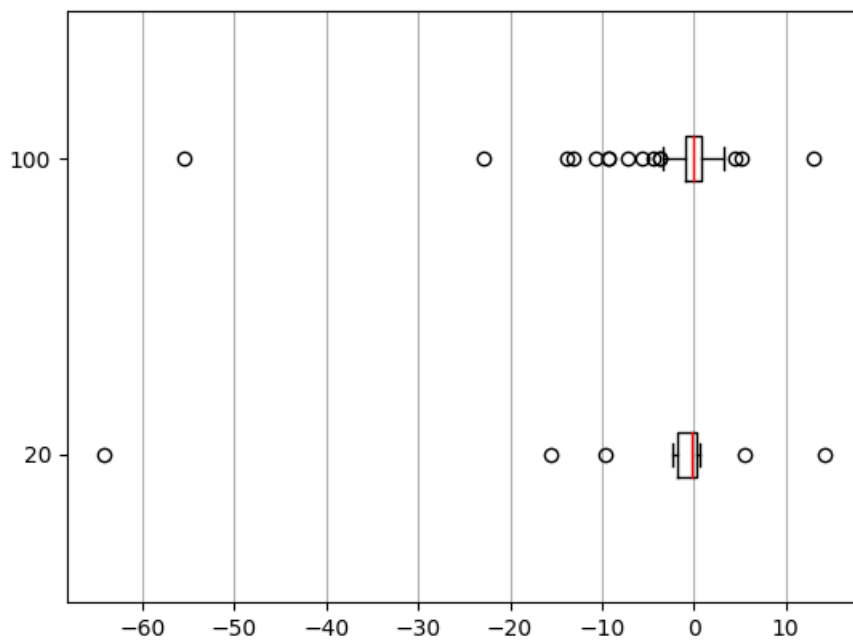


Рисунок 2: Распределение Коши

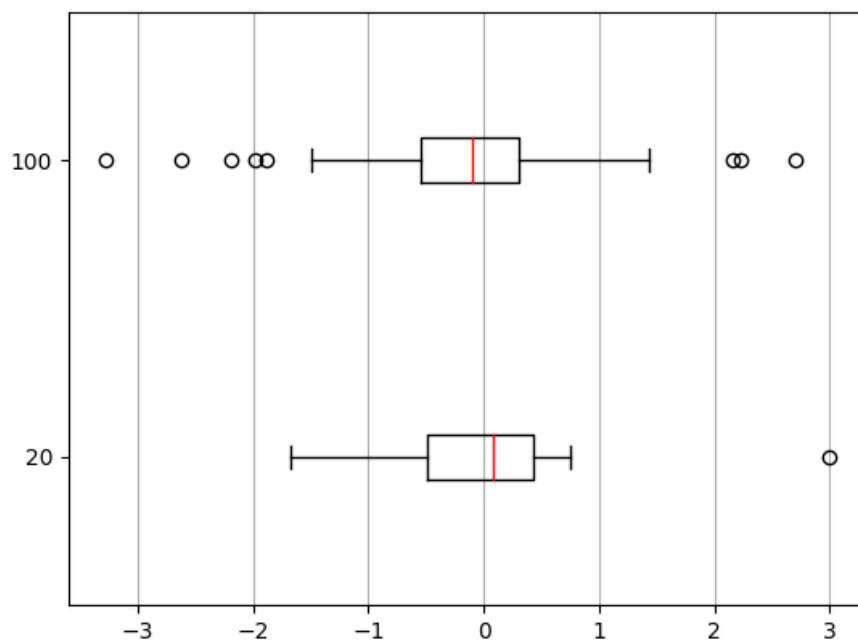


Рисунок 3: Распределение Лапласа

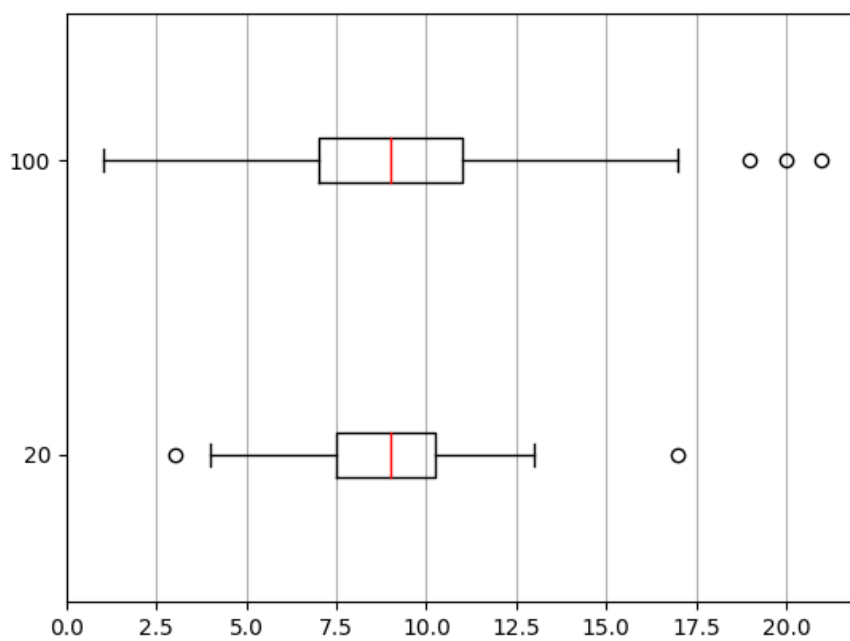


Рисунок 4: Распределение Пуассона

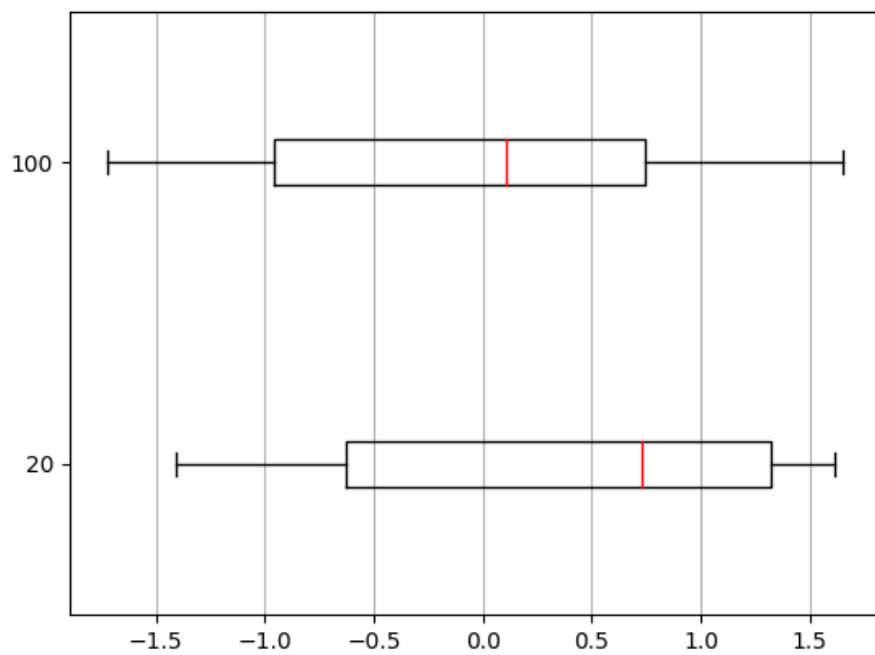


Рисунок 5: Равномерное распределение

Теоретическая доля выбросов

Распределение	Q_{25}^T	Q_{75}^T	$X_1^T(1)$	$X_2^T(2)$	$P_B^T(3)(4)$
Нормальное	-0.674	0.674	-2.698	2.698	0.007
Коши	-1	1	-4	4	0.156
Лапласа	-0.490	0.490	-1.961	1.961	0.063
Пуассона	8	12	2	18	0.008
Равномерное	-0.866	0.866	-3.464	3.464	0

Таблица 1: теоретическая вероятность выбросов

Доля выбросов

Выборка	Доля выбросов	Дисперсия
Нормальное n = 20	0.02285	0.001715
Нормальное n = 100	0.01041	0.00017
Коши n = 20	0.15105	0.004681
Коши n = 100	0.15442	0.0011
Лаплас n = 20	0.07565	0.004625
Лаплас n = 100	0.06503	0.00096
Пуассон n = 20	0.0234	0.001802
Пуассон n = 100	0.01031	0.000223
Равномерное n = 20	0.00335	0.000401
Равномерное n = 100	0.0	0.0

Таблица 2: доля выбросов

Список литературы

1. Конспекты лекции
2. Википедия: <https://ru.wikipedia.org/wiki>

Ссылка на github: <https://github.com/KateZabolotskih/MathStatLabs>