

Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики
Кафедра «Прикладная математика»

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №1
по дисциплине «МАТЕМАТИЧЕСКАЯ СТАТИСТИКА»

Выполнил студент:

Трощенко Константин
группа: 3630102/80301

Проверил:

Баженов Александр Николаевич

Санкт-Петербург
2021

Содержание

1	Постановка задачи	2
2	Теория	2
2.1	Плотности	2
2.2	Гистограмма	3
3	Реализация	3
4	Результаты	4
5	Обсуждение	5
6	Приложение	5

Список иллюстраций

1	Нормальное распределение (1)	4
2	Распределение Коши (2)	4
3	Распределение Лапласа (3)	4
4	Распределение Пуассона (4)	5
5	Равномерное распределение (5)	5

1 Постановка задачи

Для 5 распределений:

- Нормальное распределение $N(x, 0, 1)$
- Распределение Коши $C(x, 0, 1)$
- Распределение Лапласа $L\left(x, 0, \frac{1}{\sqrt{2}}\right)$
- Распределение Пуассона $P(k, 10)$
- Равномерное распределение $U(x, -\sqrt{3}, \sqrt{3})$

Сгенерировать выборки размером 10,50 и 1000 элементов. Построить на одном рисунке гистограмму и график плотности распределения.

2 Теория

2.1 Плотности

- Нормальное распределение

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

- Распределение Коши

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

- Распределение Лапласа

$$L\left(x, 0, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} \quad (3)$$

- Распределение Пуассона

$$P(k, 10) = \frac{10^k}{k!} e^{-10} \quad (4)$$

- Равномерное распределение

$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{при } |x| \leq \sqrt{3} \\ 0 & \text{при } |x| > \sqrt{3} \end{cases} \quad (5)$$

2.2 Гистограмма

Гистограмма - это инструмент, позволяющий зрительно оценить распределение статистических данных, сгруппированных по частоте попадания данных в определенный (заранее заданный) интервал.

Как строится гистограмма: множество значений, которое может принимать элемент выборки, разбивается на несколько интервалов. Чаще всего эти интервалы берут одинаковыми, но это не является строгим требованием. Эти интервалы откладываются на горизонтальной оси, затем над каждым рисуется прямоугольник. Если все интервалы были одинаковыми, то высота каждого прямоугольника пропорциональна числу элементов выборки, попадающих в соответствующий интервал. Если интервалы разные, то высота прямоугольника выбирается таким образом, чтобы его площадь была пропорциональна числу элементов выборки, которые попали в этот интервал.

Гистограммы применяются в основном для визуализации данных на начальном этапе статистической обработки. Построение гистограмм используется для получения эмпирической оценки плотности распределения случайной величины. Для построения гистограммы наблюдаемый диапазон изменения случайной величины разбивается на несколько интервалов и подсчитывается доля от всех измерений, попавшая в каждый из интервалов. Величина каждой доли, отнесенная к величине интервала, принимается в качестве оценки значения плотности распределения на соответствующем интервале

3 Реализация

Лабораторная работа выполнена с помощью средств языка программирования Python 3 в среде разработки PyCharm. Исходный код лабораторной работы приведен в приложении.

4 Результаты

Гистограмма и график плотности распределения

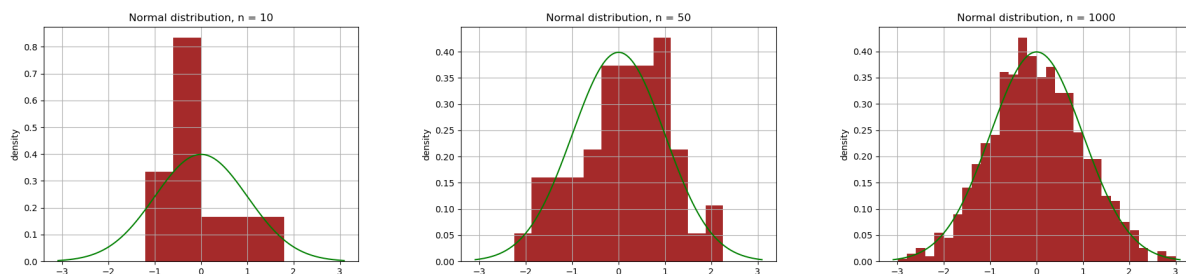


Рис. 1: Нормальное распределение (1)

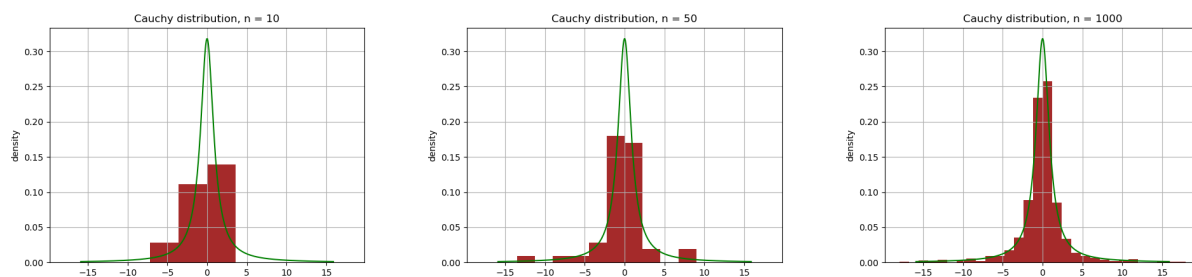


Рис. 2: Распределение Коши (2)

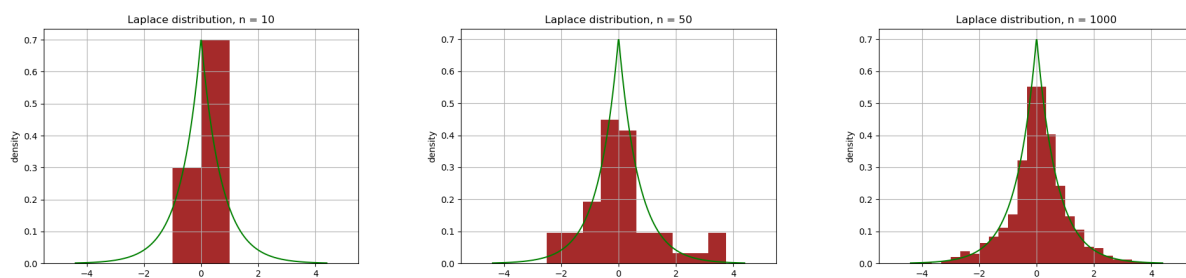


Рис. 3: Распределение Лапласа (3)

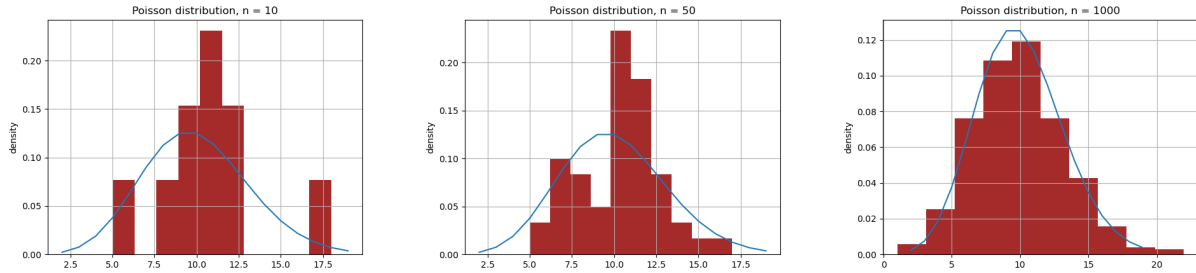


Рис. 4: Распределение Пуассона (4)

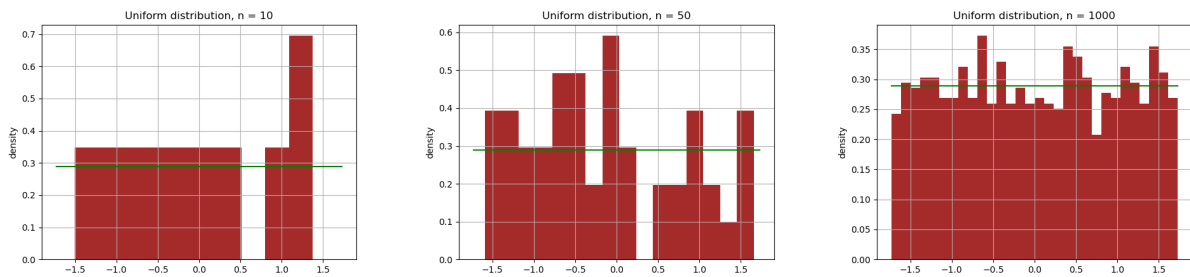


Рис. 5: Равномерное распределение (5)

5 Обсуждение

По результатам проведенных исследований можно сделать вывод, что чем больше выборка для каждого из распределений, тем ближе гистограмма к графику плотности распределения. И наоборот: чем меньше выборка, тем сильнее видна разница гистогаммы и плотности распределения.

При выборки из 10 элементов практически невозможно отличить различные распределения друг от друга. При 50 элементах гистограммы все так же далеки по точности от графиков распределения, однако некоторые из них (например, Коши и Лапласа) уже обретают похожие очертания. При 1000 элементах можно отличить каждое распределение, кроме нормального и Пуассона, которые визуальнo все равно остаются довольно схожими.

6 Приложение

Код программы: <https://github.com/FaceHunterr/MatStat/tree/main/lab1>.