

Санкт-Петербургский Политехнический Университет имени Петра Великого
Институт Прикладной Математики и Механики
Кафедра "Прикладная Математика"

Отчет по лабораторной работе №1
по дисциплине
"Математическая Статистика"

Выполнил студент:
Тырыкин Я. А.
группа 3630102/80401
Проверил:
к.ф.-м.н., доцент
Баженов А. Н.

Санкт-Петербург
2021

Содержание

1	Постановка задачи	4
2	Теория	4
2.1	Распределения	4
2.2	Гистограмма	4
2.2.1	Определение	4
2.2.2	Графическое описание	5
2.2.3	Использование	5
3	Модульная структура программы	5
4	Результаты	6
4.1	Гистограммы и графики распределений	6
5	Обсуждение	8
5.1	Гистограммы и графики распределений	8
6	Ресурсы	8

Список иллюстраций

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

1 Постановка задачи

Для 5 распределений:

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

Необходимо сгенерировать выборки размером 10, 50 и 100 элементов. Построить на одном рисунке гистограмму и график плотности распределения.

2 Теория

2.1 Распределения

- Нормальное распределение:

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

- Распределение Коши:

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

- Распределение Лапласа:

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

- 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

2.2.1 Определение

Гистограмма в математической статистике — это функция, приближающая плотность вероятности некоторого распределения, построенная на основе выборки из него

2.2.2 Графическое описание

Графически гистограмма строится следующим образом. Сначала 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

2.2.3 Использование

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 Модульная структура программы

Лабораторная работа выполнена с применением средств языка Python версии 3.7 в среде разработки PyCharm IDE (в частности, с применением встроенных методов библиотеки SciPy и Matplotlib). Исходный код лабораторной работы находится по ссылке в приложении к отчёту.

4 Результаты

4.1 Гистограммы и графики распределений

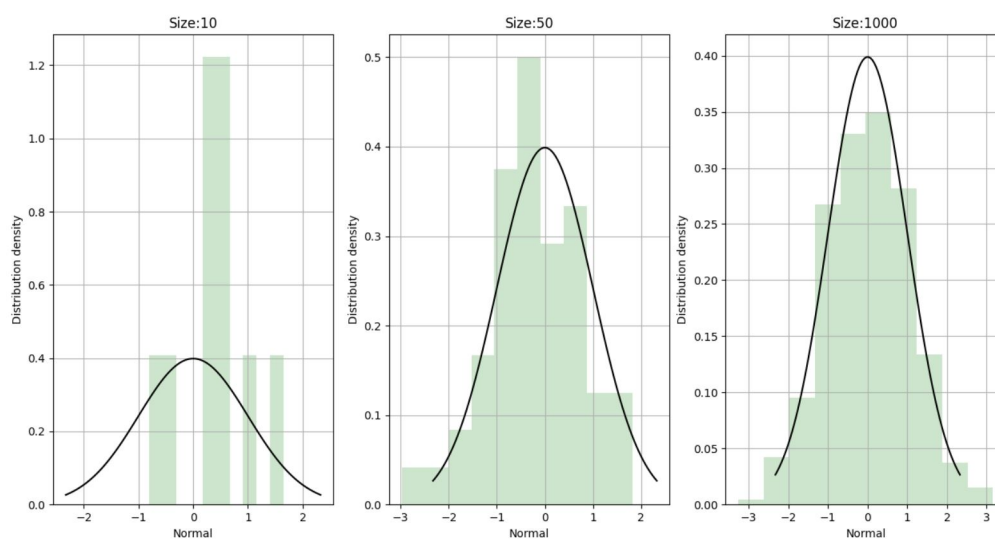


Рис. 1: Нормальное распределение

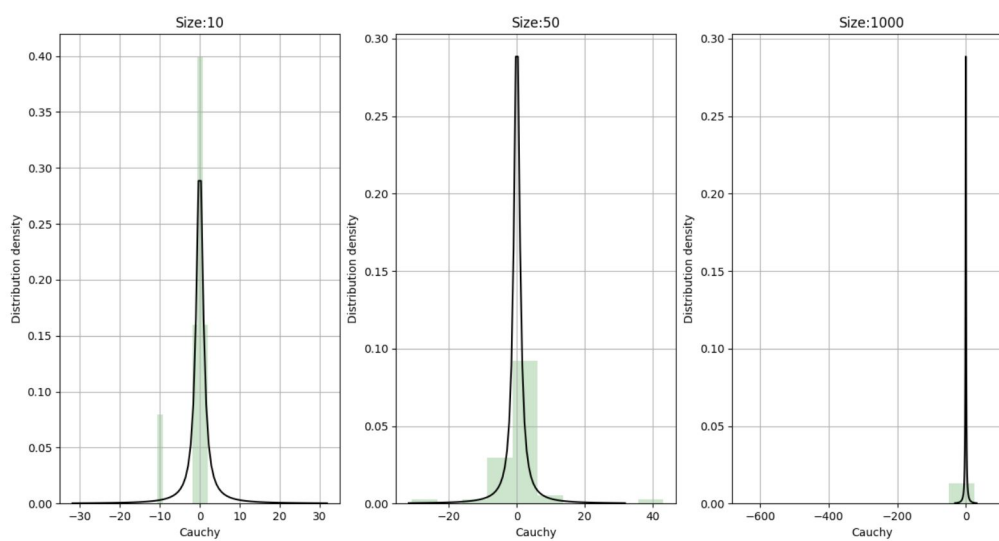


Рис. 2: Распределение Коши

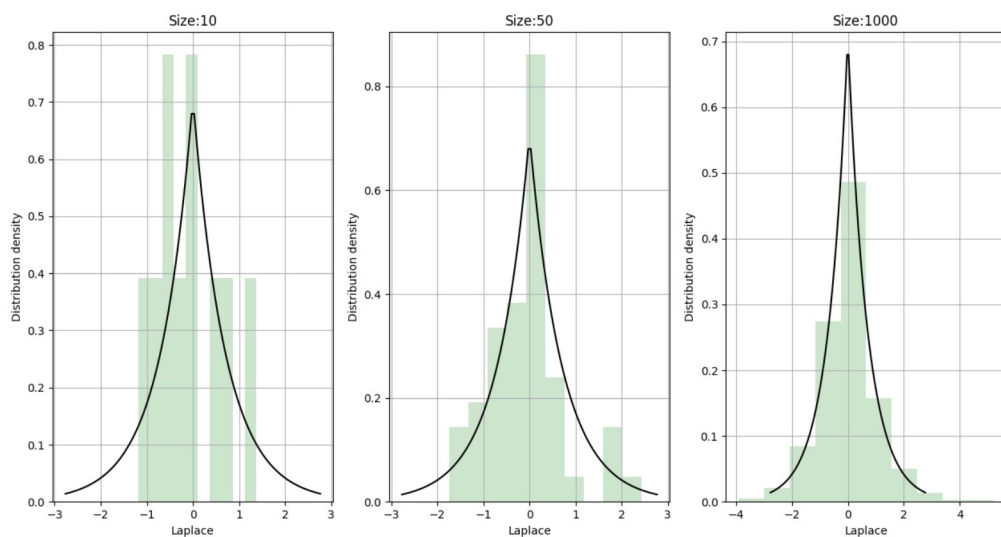


Рис. 3: Распределение Лапласа

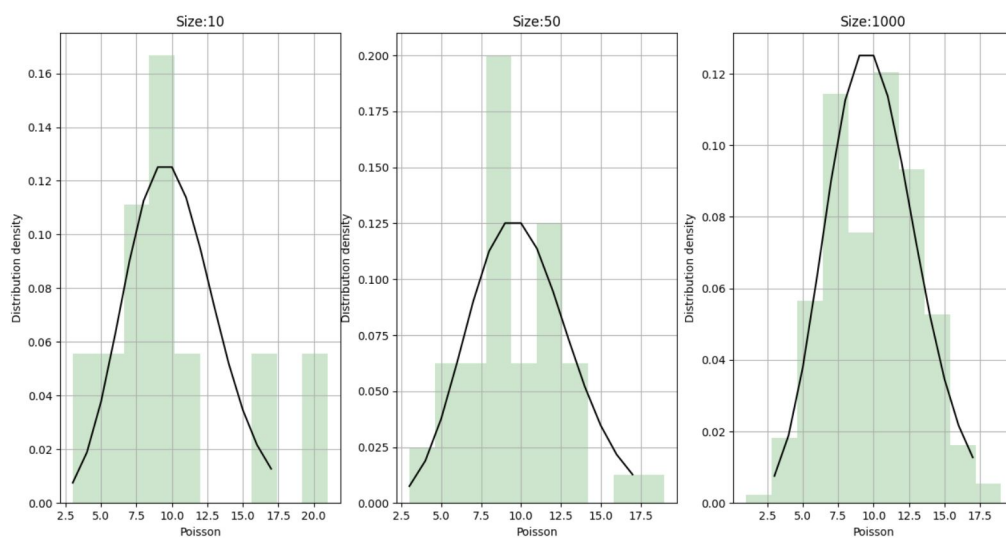


Рис. 4: Распределение Пуассона

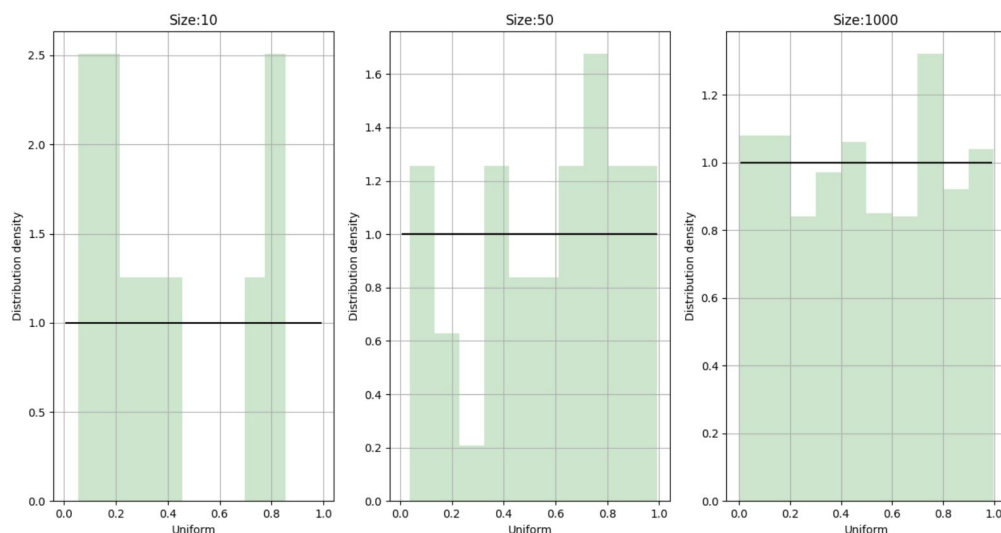


Рис. 5: Равномерное распределение

5 Обсуждение

5.1 Гистограммы и графики распределений

По результатам проделанной работы можем сделать вывод о том, что чем больше выборка для каждого из распределений, тем ближе ее гистограмма к графику плотности вероятности того закона, по которому распределены величины сгенерированной выборки. Чем меньше выборка, тем менее она показательна - тем хуже по ней определяется характер распределения величины.

Также можно заметить, что максимумы гистограмм и плотностей распределения почти нигде не совпали. Наблюдаются всплески гистограмм, что наиболее хорошо прослеживается на распределении Коши.

6 Ресурсы

Код программы, реализующей отрисовку обозначенных распределений:

<https://github.com/YaroslavAggressive/Mathematical-statistics-lab-works>