Санкт-Петербургский политехнический университет
Петра Великого

Институт прикладной математики и механики

Кафедра «Прикладная математика»

Отчет по лабораторной работе №1

по дисциплине
"Математическая статистика"

Выполнила студентка
группы 3630102/80201 Деркаченко Анна Олеговна

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

Санкт-Петербург
2021 г.

Содержание

Список иллюстраций

1. Постановка задачи

Даны распределения:

- нормальное распределение $N(x, 0, 1)$
- распределение Коши $C(x, 0, 1)$
- распределение Лапласа $L(x, 0, \frac{1}{\sqrt{2}})$
- распределение Пуассона $P(k, 10)$
- равномерное распределение $U(x, -\sqrt{3}, \sqrt{3})$

Необходимо:

1) Сгенерировать выборки размером 10, 50 и 1000 элементов
2) остроить на одном рисунке гистограмму и график плотности распределения

2. Теория

2.1. Распределения

Плотности рассматриваемых распределений:

- нормальное распределение
$$N(x, 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \tag{1}$$
- распределение Коши
$$C(x, 0, 1) = \frac{1}{\pi} \frac{1}{x^2 + 1} \tag{2}$$
- распределение Лапласа
$$L(x, 0, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|} \tag{3}$$
- распределение Пуассона
$$P(k, 10) = \frac{10^k}{k!} e^{-10} \tag{4}$$
- равномерное распределение
$$U(x, -\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}}, |x| \le \sqrt{3} \\ 0, |x| > \sqrt{3} \end{cases} \tag{5}$$

2.2. Гистограмма

Гистограмма - функция, построенная на основе выборки из некоторого распределения и приближающая его плотность вероятности.

Гистограммы используются для визуализации данных в начале статистической обработки.

Графическое построение гистограммы:

1) Множество значений, которые может принимать элемент выборки, разбивается на несколько интервалов

2) Данные интервалы откладываются на горизонтальной оси

3) Рисуется прямоугольник:

   - с высотой, пропорциональной числу элементов выборки, попадающих в данный интервал, если интервалы одинаковы
   - с площадью, пропорциональной числу элементов выборки, попадающих в данный интервал, если интервалы различны

3. Реализация

Реализация лабораторной работы проводилась на языке Python в среде разработки PyCharm с использованием дополнительных библиотек:

- scipy
- numpy
- matplotlib
- math

Исходный код лабораторной работы размещен в GitHub-репозитории.

URL: https://github.com/derkanw/Mathstat/tree/main/lab1

4. Результаты

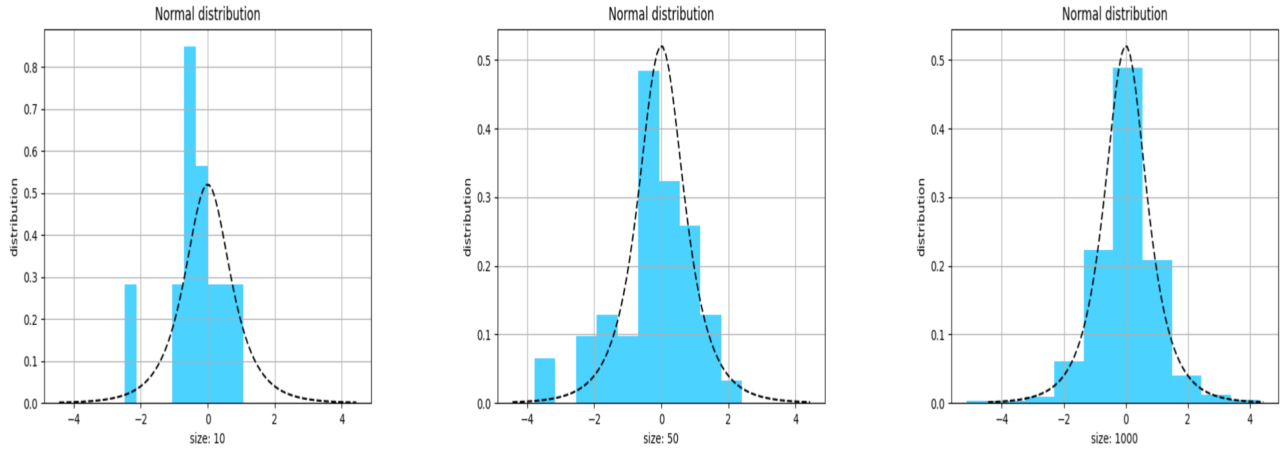

Рис. 1. Нормальное распределение (1)

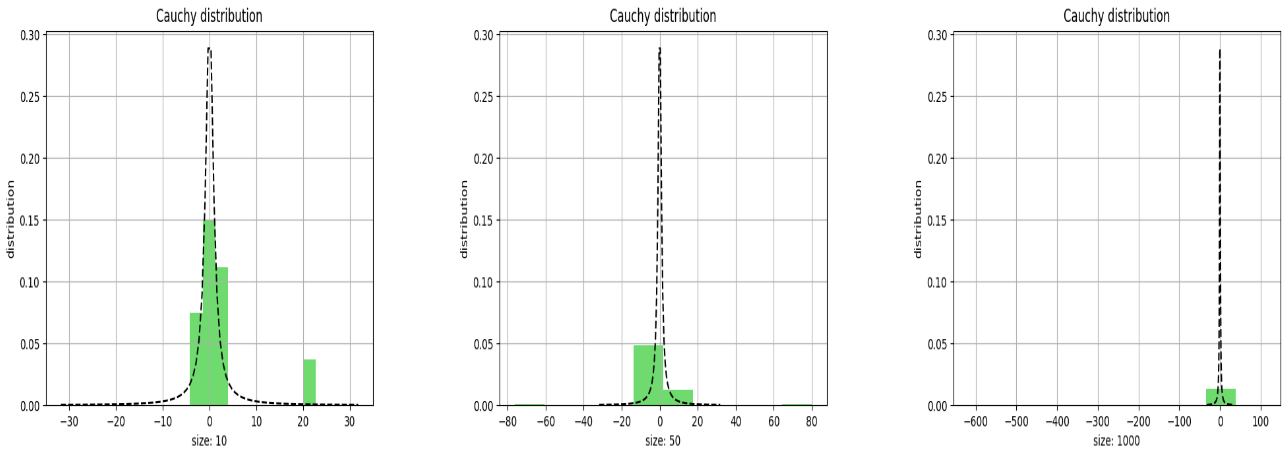

Рис. 2. Распределение Коши(2)

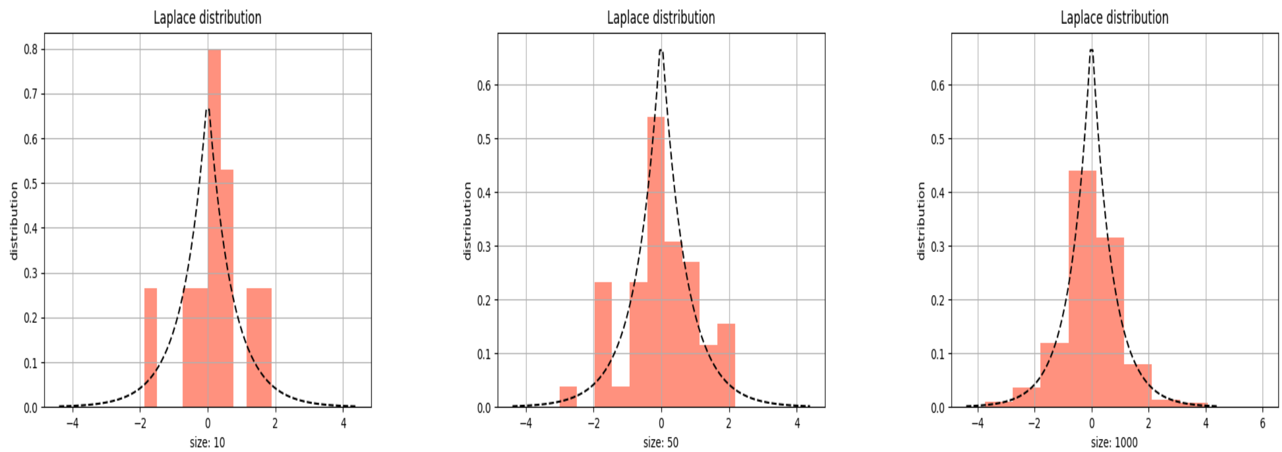

Рис. 3. Распределение Лапласа (3)

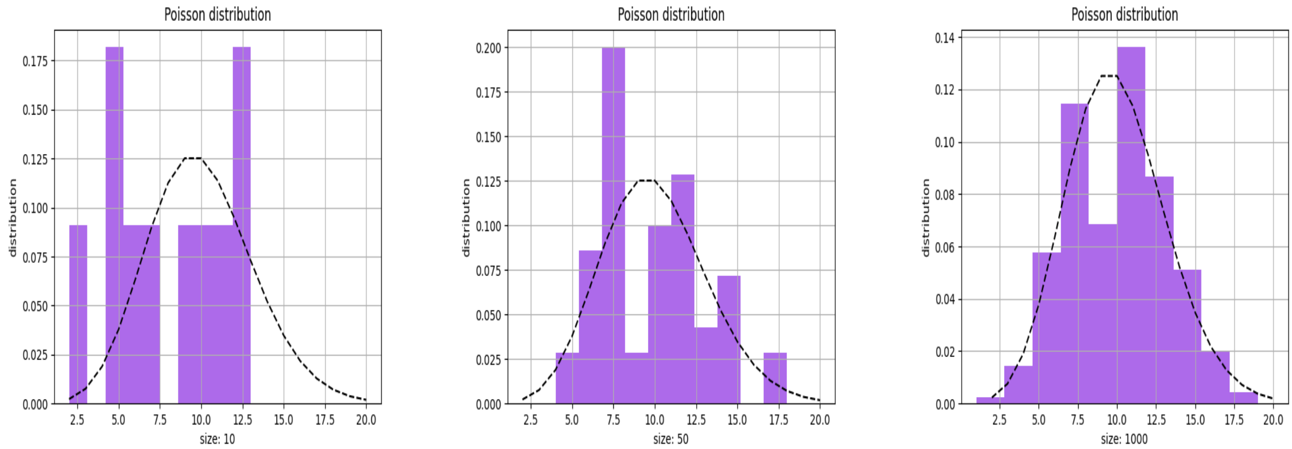

Рис. 4. Распределение Пуассона(4)

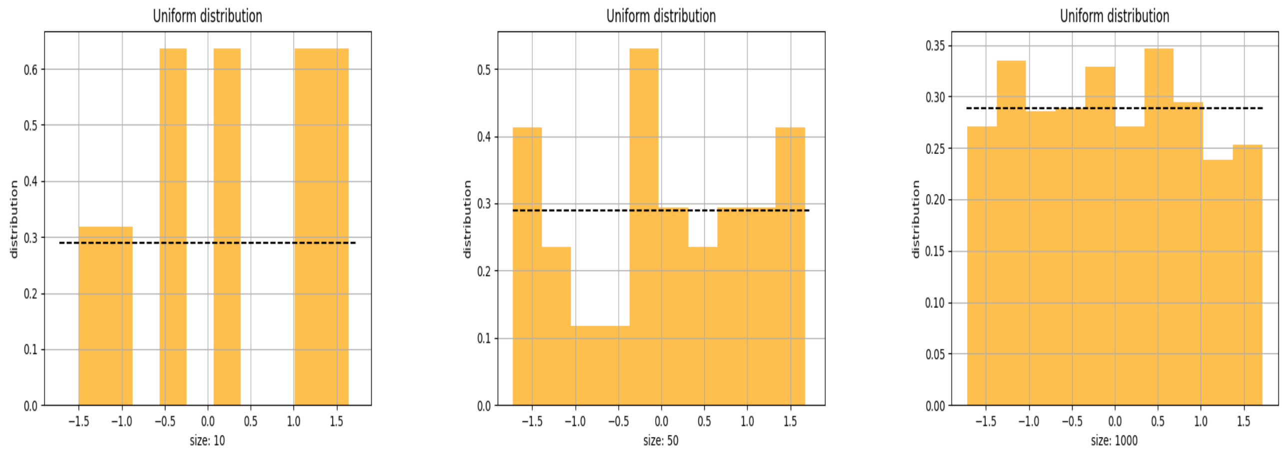

Рис. 5. Равномерное распределение (5)

5. Обсуждение

Анализ результатов демонстрирует следующую закономерность для всех распределений: чем больше выборка из распределения, тем больше соответсвует гистограмма распределения графику плотности его распределения. То есть для определения характера распределения величины стоит рассматривать выборки больших размеров.

При сравнении гистаграмм распределений между собой можно выделить особенность, что гистограммы нормального распределения, распределения Коши, Лапласа плохо отличимы между собой, тем более при маленьком размере выборки. Для этих распределений характерно относительное совпадение пика гистограммы с пиком графика плотности соответсвующего распределения. Также большая часть распределений различается по высоте пиковой точки.

Стоит отметить, что распределение Пуассона является более широким, что можно выделить, как отличительную черту распределения. Аналогично для равномерного распределения характерно наличие столбцов более или менее однородной высоты.

На гистограммах наблюдаются всплески, явно превыщающие исходное значение плотности распределения в заданной точке. Но они стихают по мере увеличения размеров выборки.