

Санкт-Петербургский политехнический университет Петра Великого

Институт прикладной математики и механики

Высшая школа прикладной математики

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №1

по дисциплине «Математическая статистика»

Выполнила
студентка гр.3630102/80101

А.А. Тимофеева

Руководитель доцент, к.ф.-м.н.

А.Н.Баженов

Санкт-Петербург 2021

СОДЕРЖАНИЕ

СПИСОК ИЛЛЮСТРАЦИЙ	3
1 ПОСТАНОВКА ЗАДАЧИ.....	4
2 ТЕОРИЯ.....	4
2.1 Плотности.....	4
2.2 Гистограмма.....	5
3 РЕАЛИЗАЦИЯ.....	5
4 РЕЗУЛЬТАТЫ.....	6
4.1 Гистограмма и график плотности распределения.....	6
5 ОБСУЖДЕНИЕ.....	7
6 ПРИЛОЖЕНИЕ.....	8

СПИСОК ИЛЛЮСТРАЦИЙ

Рис.1 Нормальное распределение (1).....	6
Рис.2 Распределение Коши (2)	6
Рис.3 Распределение Лапласа (3)	6
Рис.4 Распределение Пуассона (4)	7
Рис.5 Равномерное распределение (5)	7

1 ПОСТАНОВКА ЗАДАЧИ

Для 5 распределений:

- 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 $U(x, -\sqrt{3}, \sqrt{3})$

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

2 ТЕОРИЯ

2.1 Плотности

- Нормальное распределение

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

- 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

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

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.

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 встроенных средств языка программирования Python в среде разработки PyCharm. Исходный код лабораторной работы приведён в приложении.

4 РЕЗУЛЬТАТЫ

4.1 Гистограмма и график плотности распределения

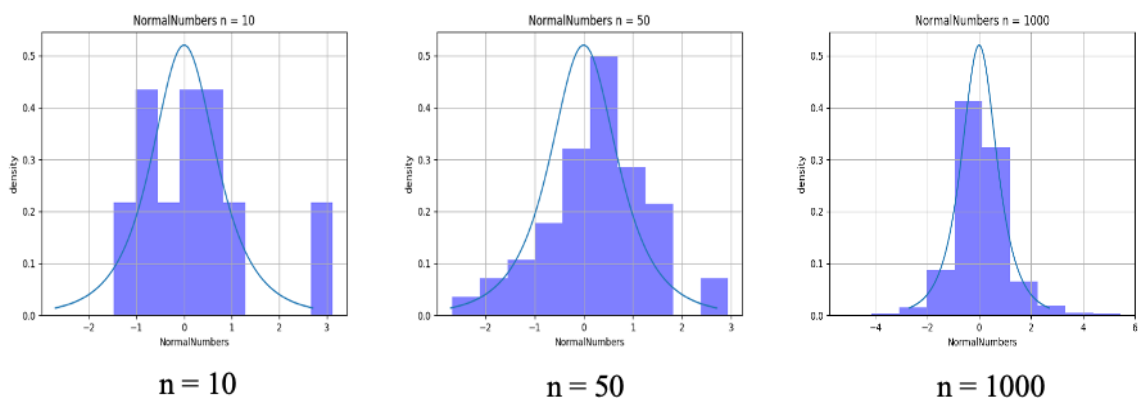


Рис.1 Нормальное распределение (1)

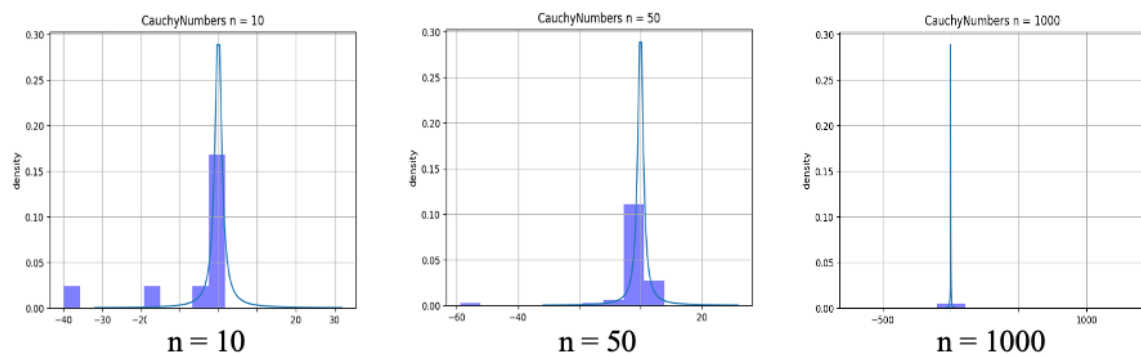


Рис.2 Распределение Коши (2)

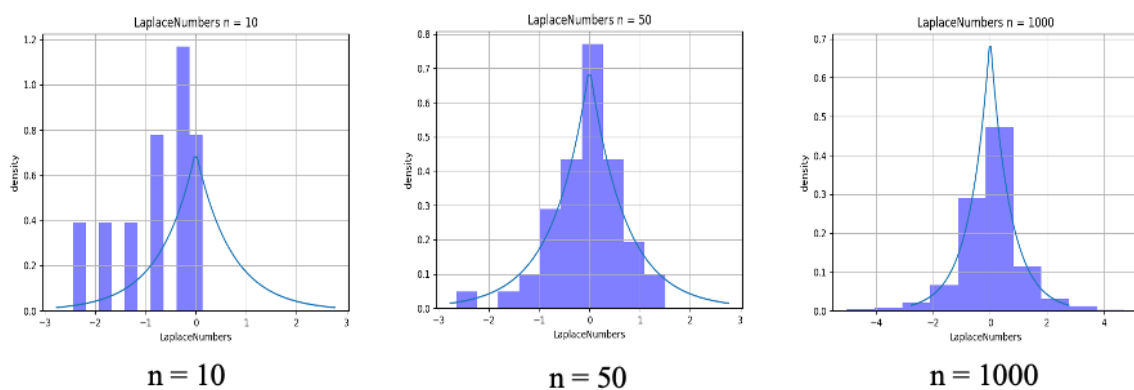


Рис.3 Распределение Лапласа (3)

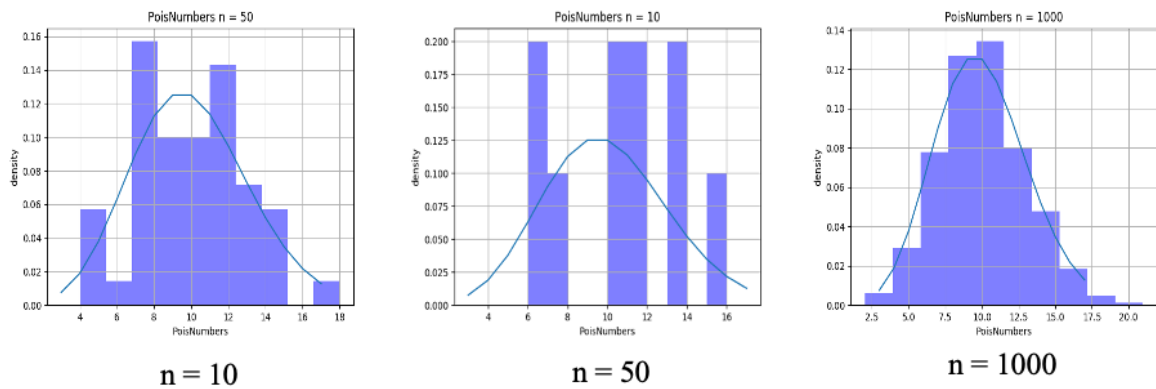


Рис.4 Распределение Пуассона (4)

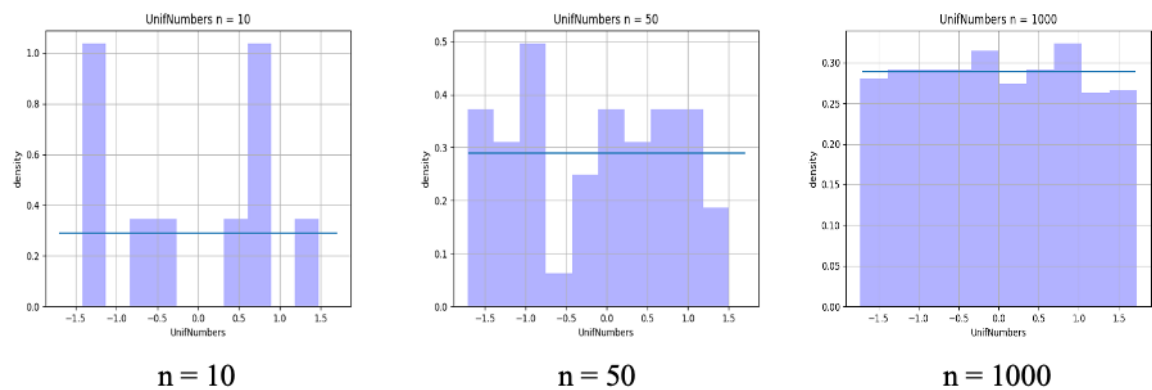


Рис.5 Равномерное распределение (5)

5 ОБСУЖДЕНИЕ

По результатам проделанной работы можем сделать вывод, 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.

Различить распределения по виду гистограммы немного трудно, особенно если говорить о маленьких выборках, например $n = 10$. При выборке из 50 элементов гистограммы нормального распределения, распределения Лапласа, распределения Пуассона начинают принимать виды очень похожие на описываемые ими распределения. Рассматривая выборку размером $n = 1000$, можно вполне точно отличить распределение Пуассона от распределения Лапласа. Также вполне удастся различить равномерное распределение. А вот гистограммы нормального распределения и распределения Лапласа визуально кажутся одинаковыми.

6 ПРИЛОЖЕНИЕ

Код программы URL: <https://github.com/tmffv/MathStat/tree/master/src>