Санкт-Петербургский политехнический университет Петра Великого

Институт прикладной математики и механики Кафедра «Прикладная математика»

> Отчёт по лабораторной работе №1 по дисциплине «Математическая статистика»

> > Выполнил студент: Шагвалиев Михаил Александрович группа: 3630102/80201

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

Санкт-Петербург 2020 г.

Содержание

1	Постановка задачи	3
2	Теория 2.1 Рассматриваемые распределения 2.2 Гистограмма 2.2.1 Построение гистограммы	4 4 4
3	Реализация	5
4	Результаты 4.1 Гистограммы и графики	6
5	Обсуждение	9
6	Приложения	10
Cı	писок литературы	10

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:

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

Требуется сгенерировать выборки размером 10, 50 и 1000 элементов, построить на одном рисунке гистограмму и график плотности распределения.

2 Теория

2.1 Рассматриваемые распределения

Плотности:

• Нормальное распределение

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

• Распределение Коши

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

• 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

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

2.2 Гистограмма

В математической статистике гистрограмма — это функция, приближающая плотность вероятности некоторого распределения, построенная на основе выборки из него. [1]

2.2.1 Построение гистограммы

Пусть $X = \{x_1, ..., x_n\}$ — выборка, [a, b] — отрезок, на котором строится гистрограмма

1. [a, b] разбивается на m равных интервалов:

$$\Delta_i = (a + \frac{i-1}{m}(b-a), a + \frac{i}{m}(b-a)), i \in \{1, ..., m\}$$
(6)

2. Подсчитаем количество элементов выборки, попавших в интервал Δ_i :

$$d_i = |\{k \mid x_k \in \Delta_i\}|, i \in \{1, ..., m\}$$
(7)

3. Построим прямоугольники с основанием Δ_i и высотой $\frac{d_i}{c}$, где c — общее число элементов выборки, попавших в [a,b] (таким образом, гистрограмма "отнормирована": сумма площадей всех прямоугольников равняется единице).

В качестве (a,b,m) возьмем, соответственно, $(\min_{x\in X} x, \max_{x\in X} x, \lfloor 3\sqrt[3]{n}\rfloor)$

3 Реализация

Работа выполнена с помощью языка **Python** в IDE **PyCharm**, также были использованы библиотеки:

- scipy генерация данных
- numpy работа с массивами
- matplotlib отрисовка графиков

Исходный код работы приведен в приложении.

4 Результаты

4.1 Гистограммы и графики

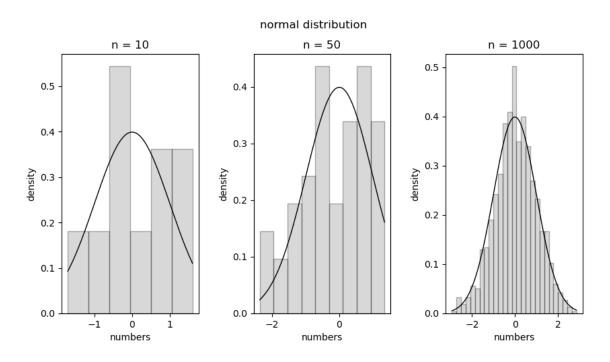


Рис. 1: Нормальное распределение

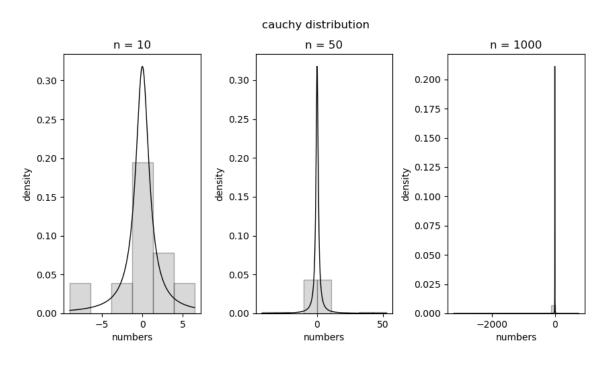


Рис. 2: Распределение Коши

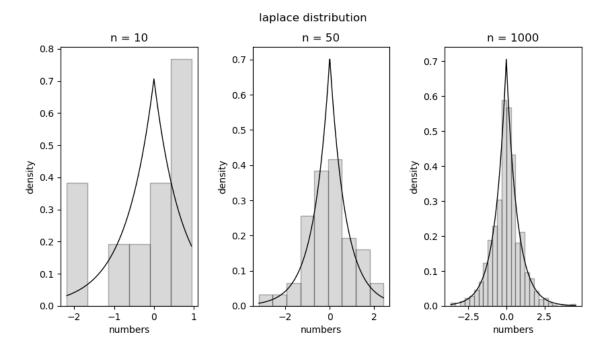


Рис. 3: Распределение Лапласа

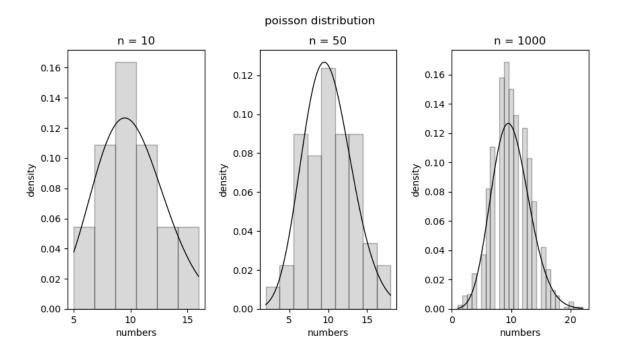


Рис. 4: Распределение Пуассона

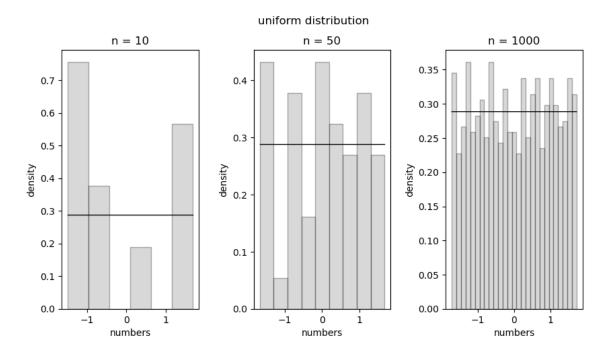


Рис. 5: Равномерное распределение

5 Обсуждение

По результатам проведенной работы можно сделать вывод о том, что чем больше выборка, тем лучше гистрограмма приближает функцию плотности распределения случайной величины. Чем меньше выборка — тем хуже по гистограмме определяется распределение случайной величины.

6 Приложения

1. Код лабораторной: https://github.com/MekhailS/math-statistics-labs/tree/master/lab1_histogram

Список литературы

 $[1] \ Histogram. \ URL: \ \verb|https://en.wikipedia.org/wiki/Histogram|$