

Санкт-Петербургский политехнический университет Петра Великого

Институт прикладной математики и механики  
Высшая школа прикладной математики и вычислительной физики

Отчёт по лабораторной работе №1  
по дисциплине Математическая статистика

Выполнил студент:  
Курова Анна Николаевна  
группа: 3630102/70401

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

Санкт-Петербург 2020

# Содержание

<b>1</b>	<b>Постановка задачи</b>	<b>4</b>
<b>2</b>	<b>Теория</b>	<b>4</b>
2.1	Плотности распределения . . . . .	4
2.2	Гистограмма . . . . .	4
2.2.1	Определение . . . . .	4
2.2.2	Графическое описание . . . . .	5
2.2.3	Использование . . . . .	5
<b>3</b>	<b>Реализация</b>	<b>5</b>
<b>4</b>	<b>Результаты</b>	<b>5</b>
4.1	График плотности и гистограмма распределений . . . . .	5
<b>5</b>	<b>Обсуждения</b>	<b>7</b>
<b>6</b>	<b>Приложения</b>	<b>8</b>

## Список иллюстраций

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

# 1 Постановка задачи

Для 5 видов различных распределений сгенерировать выборки, состоящие из 10, 50 и 1000 элементов.

На одном графике построить плотность и гистограмму распределения для каждой из выборок

## 2 Теория

### 2.1 Плотности распределения

Формулы вычисления плотности для каждого распределения:

1. Нормальное распределение:

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

2. Распределение Коши:

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

3. Распределение Лапласа:

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

4. Распределение Пуассона:

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

5. Равномерное распределение:

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

### 2.2 Гистограмма

#### 2.2.1 Определение

Гистограмма - способ графического представления табличных данных.

Количественные соотношения некоторого показателя представлены в виде прямоугольников, площади которых пропорциональны. Чаще всего для удобства восприятия ширину прямоугольников берут одинаковую, при этом их высота определяет соотношения отображаемого параметра.

## 2.2.2 Графическое описание

Гистограмма строится следующим образом. Сначала множество значений, которое может принимать элемент выборки, разбивается на несколько интервалов (bins). 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 Реализация

Для реализации использовался высокоуровневый язык python.

Для наглядного представления графиков использовались библиотеки matplotlib и seaborn

Так же использовалась библиотека math

Для реализации нормального распределения использовался метод Бокса-Мюллера.

Для реализации распределения Коши модифицированный метод Бокса-Мюллера, только в данном случае мы получали тангенс и возвращали результат деления двух нормальных случайных величин

Для распределения Пуассона использовался метод на основе инверсии путем последовательного поиска (основан на алгоритме Кнута)

Для равномерного распределения одна из простейших схем:  $x_m = ax_{m-1} + b \pmod m$

# 4 Результаты

## 4.1 График плотности и гистограмма распределений

Красным обозначен график плотности распределения. Синяя линия является вспомогательной

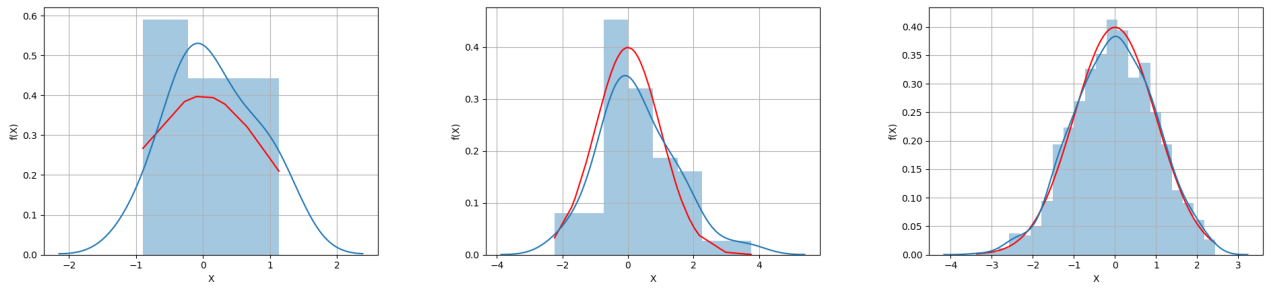


Рис. 1: Нормальное распределение

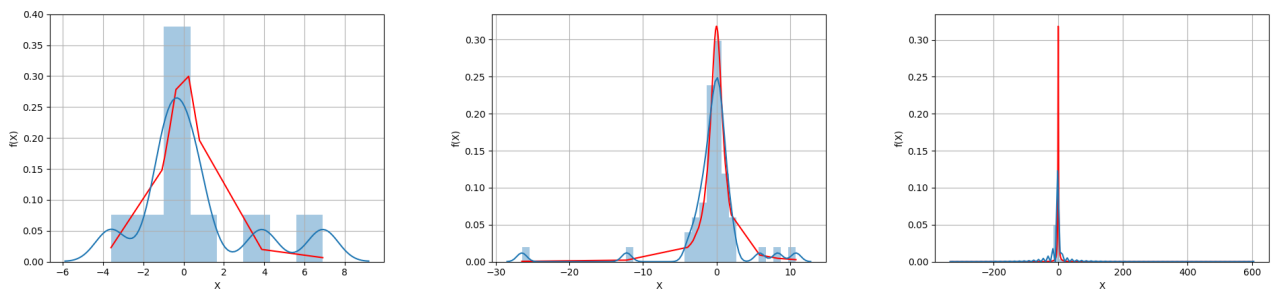


Рис. 2: Распределение Коши

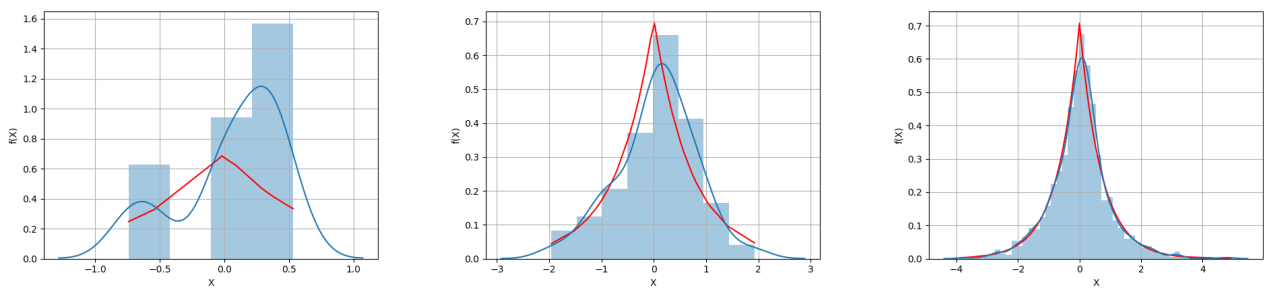


Рис. 3: Распределение Лапласа

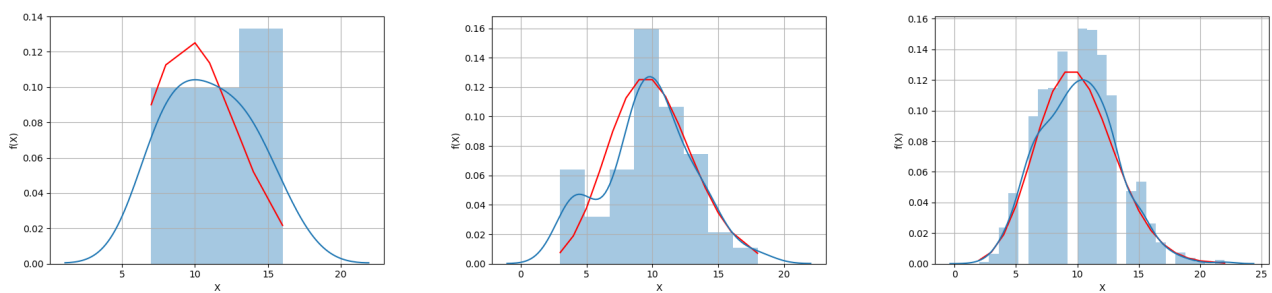


Рис. 4: Распределение Пуассона

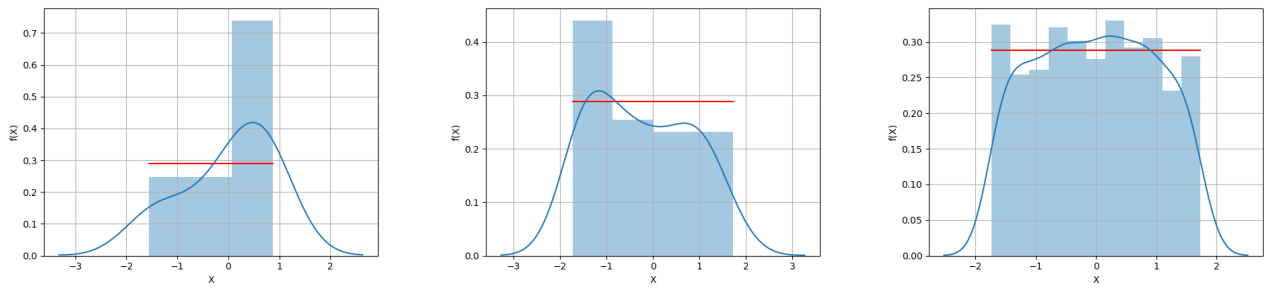


Рис. 5: Равномерное распределение

## 5 Обсуждения

Исходя из полученных графиков можем сделать некоторые выводы. Чем больше выборка, тем ближе гистограмма и график плотности. Можно заметить, что иногда наблюдаются некоторые "скачки" в гистограмме, особенно при малых выборках. Первые четыре распределения имеют один ярко выраженный пик - они унимодальны.

На рисунке номер пять большое количество пиков - распределение однородно. Все гистограммы имеют симметричную форму относительно центральной линии.

## 6 Приложения

Код программы: <https://github.com/katonapng/MathStat>

### Список литературы

- [1] Histogram. URL: <https://en.wikipedia.org/wiki/Histogram>–196.
- [2] Генераторы дискретно распределенных случайных величин:  
<https://m.habr.com/ru/post/265321/>
- [3] Генераторы непрерывно распределенных случайных величин:  
<https://m.habr.com/ru/post/263993/>
- [4] Generating Poisson random values:  
<https://www.johndcook.com/blog/2010/06/14/generating-poisson-random-values/>