

Санкт-Петербургский политехнический университет Петра Великого

Институт прикладной математики и механики

Кафедра «Телематика (при ЦНИИ РТК)»

Отчет по лабораторной работе

Построение гистограмм по выборкам с заданным распределением

По дисциплине «Теория вероятностей и Математическая статистика»

Выполнил

Студент гр. 3630201/80101

М.Д. Маляренко

Руководитель

к.ф.-м.н., доцент

А.Н. Баженов

« ____ » _____ 2020г.

Санкт-Петербург
2020

Содержание

1	Постановка задачи	4
2	Теория	5
2.1	Рассматриваемые распределения	5
2.2	Гистограмма	5
3	Реализация	6
4	Результаты	7
	Заключение	9
	Список литературы	10
	Приложение А. Репозиторий с исходным кодом	11

Список иллюстраций

1	Нормальное распределение $N(x, 0, 1)$	7
2	Распределение Лапласа $L(x, 0, 1/\sqrt{2})$	7
3	Распределение Коши $C(x, 0, 1)$	7
4	Дискретное распределение Пуассона $P(k, 10)$	8
5	Нормальное распределение $U(x, -\sqrt{3}, \sqrt{3})$	8

1 Постановка задачи

Для заданных распределений случайных величин необходимо сгенерировать выборки из 10, 50, 1000 элементов. Для каждой выборки построить гистограмму и функцию плотности распределения вероятностей на одном графике.

Заданные распределения:

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

2 Теория

2.1 Рассматриваемые распределения

Плотности заданных распределений вероятностей [2]:

1. Нормальное распределение

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

2. Распределение Лапласа

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

3. Распределение Коши

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

4. Дискретное распределение Пуассона

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

5. Равномерное распределение

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

2.2 Гистограмма

Гистограмма распределения вероятностей – наглядное представление функции плотности вероятности случайной величины, построенное по выборке из распределения [1].

Для построения гистограммы множество значений элементов выборки разбивается на интервалы, которые чаще выбираются равными, но это не обязательное условие.

Гистограмма представляет собой столбчатый график, где над каждым интервалом значений элементов выборки строится столбец-прямоугольник. При равных интервалах высота прямоугольника строится пропорционально числу элементов выборки соответствующего интервала, при разных интервалах площадь прямоугольника пропорциональна числу элементов выборки которые попали в этот интервал.

При выборе разбиения множества значений стоит учитывать, что существует зависимость между детализацией оценки плотности распределения, точностью её значений и количеством интервалов разбиения. В данной работе определение оптимального количества n интервалов разбиения при $N = 10$ и $N = 50$ производилось эвристически, при $N = 1000$ применялась формула

$$n = \lfloor \sqrt{N} \rfloor \quad (6)$$

При построении гистограммы распределения Пуассона каждый интервал соответствует значению дискретной случайной величины.

3 Реализация

Для построения гистограмм и графиков теоретической плотности распределения применялась система аналитических вычислений *Math*. Скрипт лабораторной работы расположен в репозитории GitHub, ссылка на который приведена в соответствующем разделе.

4 Результаты

По сгенерированным выборкам с $N = 10, 50, 1000$ случайных чисел с заданными распределениями были построены следующие графики:

На Рис. 1 представлены графики нормального распределения. Теоретическая плотность распределения вероятностей построена по формуле (1).



Рис. 1: Нормальное распределение $N(x, 0, 1)$

На Рис. 2 представлены графики распределения Лапласа. Теоретическая плотность распределения вероятностей построена по формуле (2).



Рис. 2: Распределение Лапласа $L(x, 0, 1/\sqrt{2})$

На Рис. 3 представлены графики распределения Коши. Теоретическая плотность распределения вероятностей построена по формуле (3).

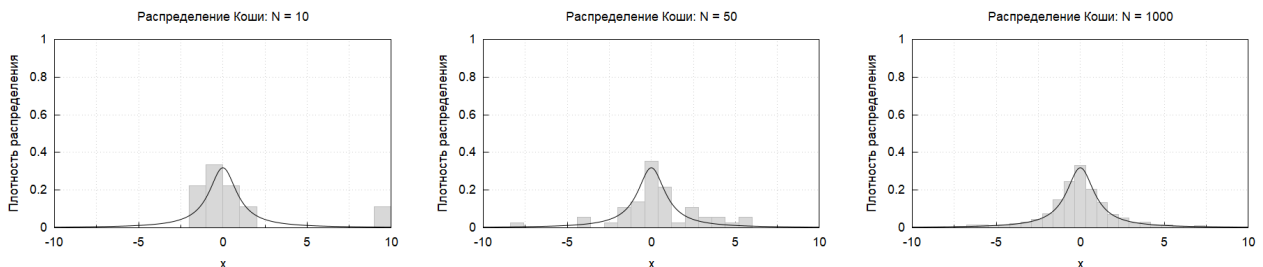


Рис. 3: Распределение Коши $C(x, 0, 1)$

На Рис. 4 представлены графики дискретного распределения Пуассона. Теоретическая плотность распределения вероятностей построена по формуле (4)

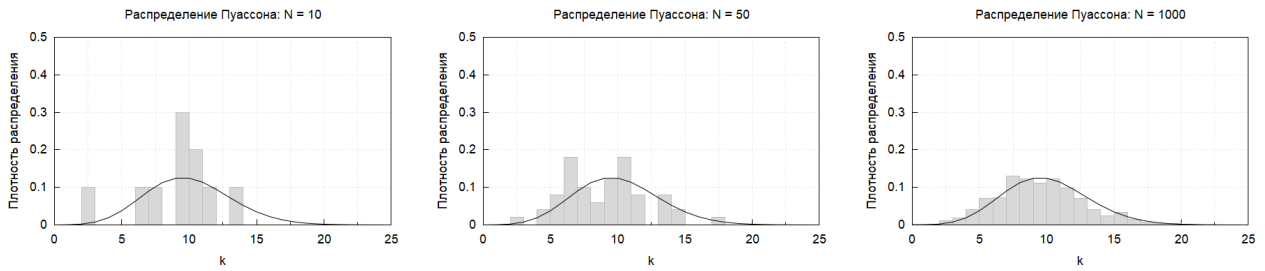


Рис. 4: Дискретное распределение Пуассона $P(k, 10)$

На Рис. 5 представлены графики равномерного распределения. Теоретическая плотность распределения вероятностей построена по формуле (5)

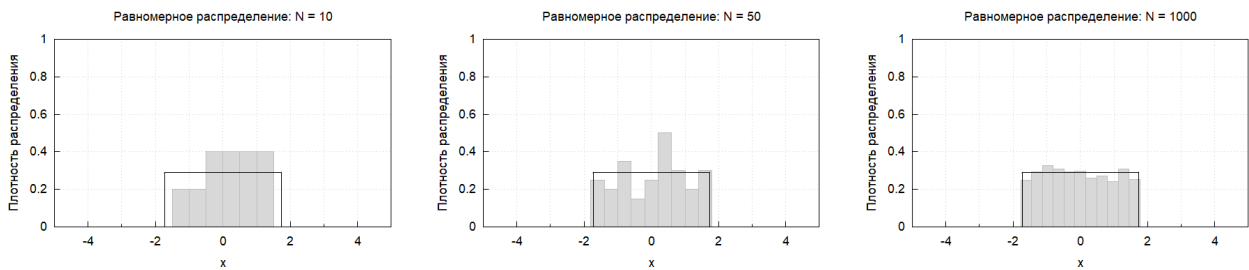


Рис. 5: Нормальное распределение $U(x, -\sqrt{3}, \sqrt{3})$

Заключение

В ходе данной лабораторной работы были сгенерированы выборки трёх размеров для пяти заданных распределений. По данным выборкам были сформированы гистограммы и они были визуально сопоставлены с теоретической плотностью распределения.

Как видно из графиков, с увеличением мощности выборки плотность распределения вероятностей приближается к своему теоретическому значению.

Для построения выборок и графиков была использована среда аналитических вычислений *Mathima* в связке с утилитой *gnuplot*.

Список литературы

- [1] Histogram <https://en.wikipedia.org/wiki/Histogram> Дата обращения 9.11.2020
- [2] Теоретическое приложение к лабораторным работам №1-4 по дисциплине «Математическая статистика». – СПб.: СПбПУ, 2020. – 12 с

Приложение А. Репозиторий с исходным кодом

Исходный код скрипта для среды *Maxima* находится в репозитории GitHub – URL <https://github.com/malyarenko-md/TeorVer>