

Санкт-Петербургский политехнический университет Петра Великого

Институт прикладной математики и механики

Кафедра «Прикладная математика»

Отчёт

по лабораторной работе №4

по дисциплине

«Математическая статистика»

Выполнил студент

В. А. Рыженко

Проверил:

к.ф.-м.н., доцент

Баженов Александр Николаевич

Санкт-Петербург, 2020 г.

Содержание

1. Постановка задачи	4
2. Теория	4
2.1. Распределения	4
2.2. Эмпирическая функция распределения	5
2.2.1. Определение	5
2.2.2. Описание	5
2.3. Оценки плотности вероятности	5
2.3.1. Определение	5
2.3.2. Ядерные оценки	5
3. Реализация	6
4. Результаты	6
4.1. Эмпирическая функция распределения	6
4.2. Ядерные оценки плотности распределения	8
5. Обсуждение	15
6. Приложения	16

Список иллюстраций

1	Нормальное распределение (1)	6
2	Распределение Коши (2)	6
3	Распределение Лапласа (3)	7
4	Распределение Пуассона (4)	7
5	Равномерное распределение (5)	8
6	Нормальное распределение, $n = 20$ (1)	8
7	Нормальное распределение, $n = 60$	9
8	Нормальное распределение, $n = 100$	9
9	Распределение Коши, $n = 20$ (2)	10
10	Распределение Коши, $n = 60$	10
11	Распределение Коши, $n = 100$	11
12	Распределение Лапласа, $n = 20$ (3)	11
13	Распределение Лапласа, $n = 60$	12
14	Распределение Лапласа, $n = 100$	12
15	Распределение Пуассона, $n = 20$ (4)	13
16	Распределение Пуассона, $n = 60$	13
17	Распределение Пуассона, $n = 100$	14
18	Равномерное распределение, $n = 20$ (5)	14
19	Равномерное распределение, $n = 60$	15

20	Равномерное распределение, $n = 100$	15
----	--	----

1. Постановка задачи

Для 5 распределений:

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

Сгенерировать выборки размером 20, 60 и 100 элементов. Построить на них эмпирические функции распределения и ядерные оценки плотности распределения на отрезке $[-4; 4]$ для непрерывных распределений и на отрезке $[6; 14]$ для распределения Пуассона.

2. Теория

2.1. Распределения

- 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. Эмпирическая функция распределения

2.2.1. Определение

Эмпирической (выборочной) функцией распределения (э. ф. р.) называется относительная частота события $X < x$, полученная по данной выборке:

$$F_n^*(x) = P^*(X < x). \quad (6)$$

2.2.2. Описание

Для получения относительной частоты $P^*(X < x)$ просуммируем в статистическом ряде, построенном по данной выборке, все частоты n_i , для которых элементы z_i статистического ряда меньше x . Тогда $P^*(X < x) = \frac{1}{n} \sum_{z_i < x} n_i$.

2.3. Оценки плотности вероятности

2.3.1. Определение

Оценкой плотности вероятности $f(x)$ называется функция $\hat{f}(x)$, построенная на основе выборки, приближённо равная $f(x)$

2.3.2. Ядерные оценки

Представим оценку в виде суммы с числом слагаемых, равным объёму выборки:

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - x_i}{h_n}\right). \quad (7)$$

Здесь функция $K(u)$, называемая ядерной (ядром), непрерывна и является плотностью вероятности, x_1, \dots, x_n — элементы выборки, h_n — любая последовательность положительных чисел, обладающая свойствами

$$h_n \xrightarrow{n \rightarrow \infty} 0; \quad \frac{h_n}{n^{-1}} \xrightarrow{n \rightarrow \infty} \infty; \quad (8)$$

Гауссово (нормальное) ядро

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad (9)$$

Правило Сильвермана

$$h_n = 1.06\hat{\sigma}n^{-1/5}, \quad (10)$$

где $\hat{\sigma}$ - выборочное стандартное отклонение.

3. Реализация

Лабораторная работа выполнена с помощью встроенных средств языка программирования Python в среде разработки Visual Code. Исходный код лабораторной работы приведён в приложении.

4. Результаты

4.1. Эмпирическая функция распределения

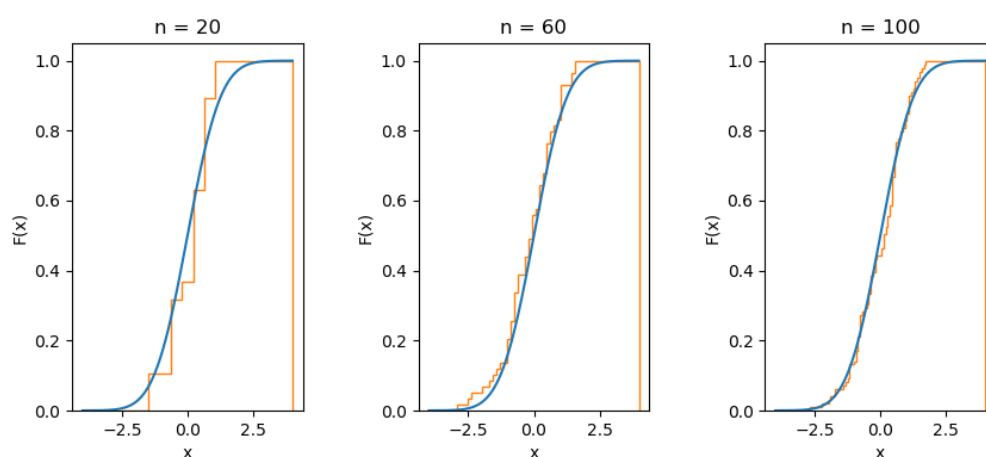


Рис. 1. Нормальное распределение (1)

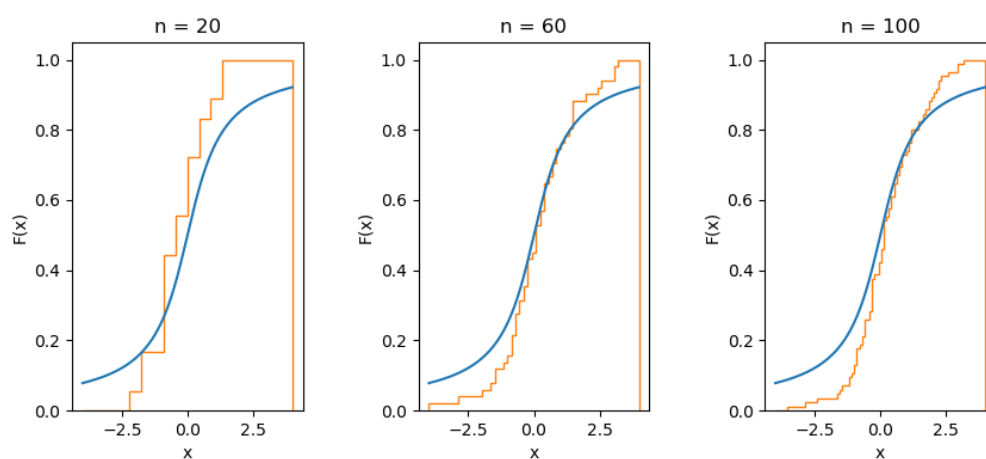


Рис. 2. Распределение Коши (2)

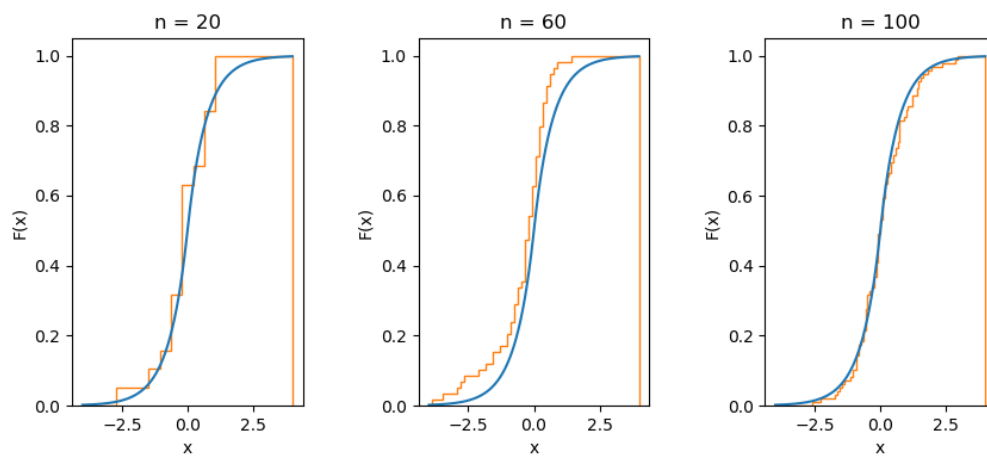


Рис. 3. Распределение Лапласа (3)

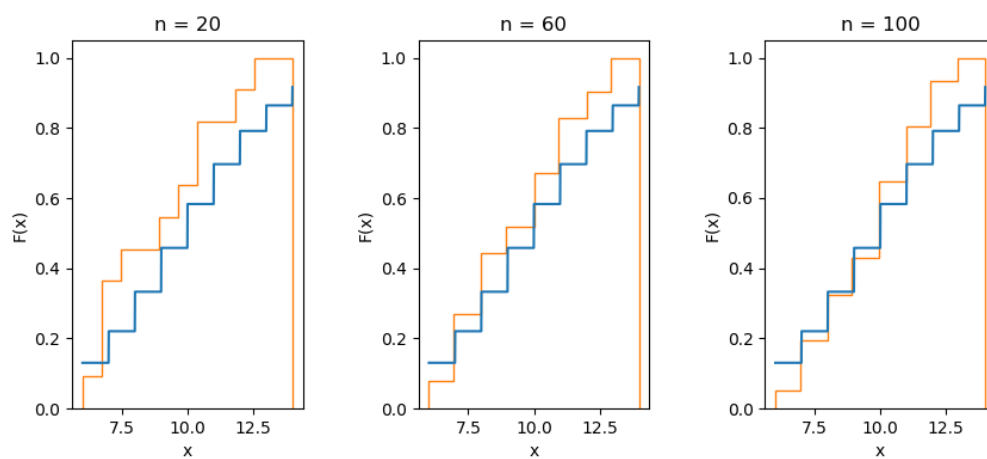


Рис. 4. Распределение Пуассона (4)

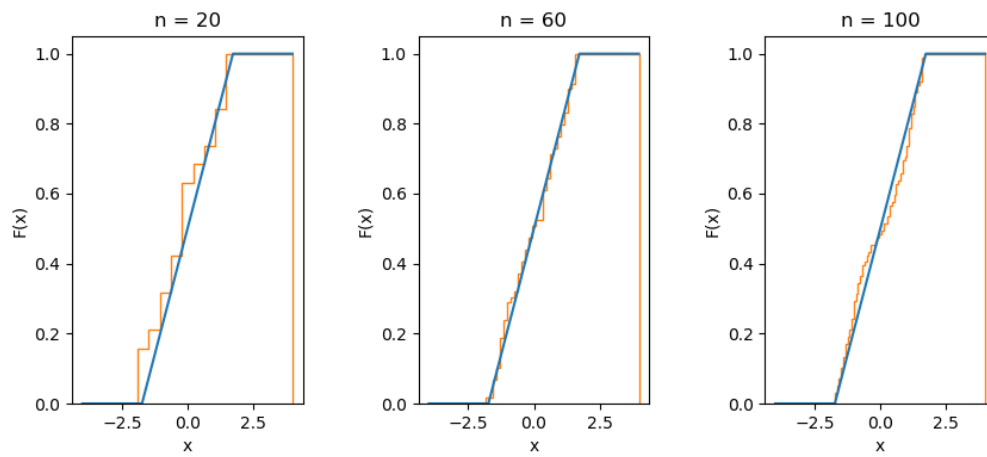


Рис. 5. Равномерное распределение (5)

4.2. Ядерные оценки плотности распределения

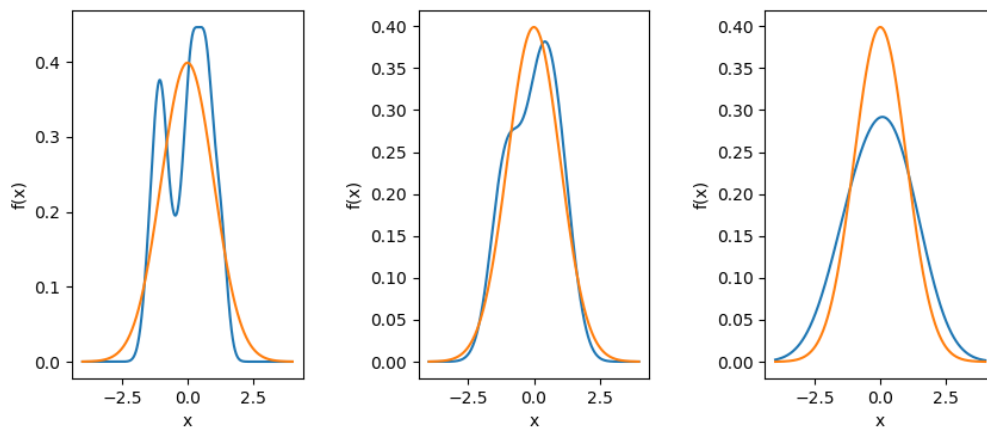


Рис. 6. Нормальное распределение, $n = 20$ (1)

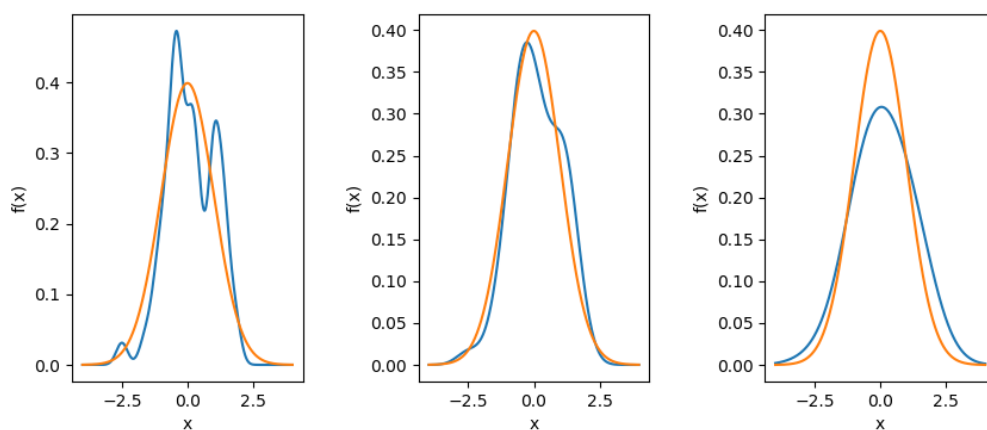


Рис. 7. Нормальное распределение, $n = 60$

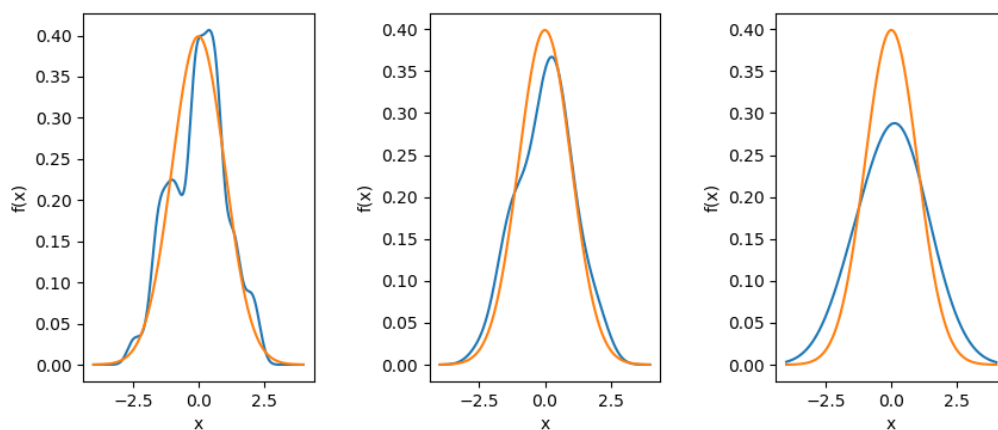


Рис. 8. Нормальное распределение, $n = 100$

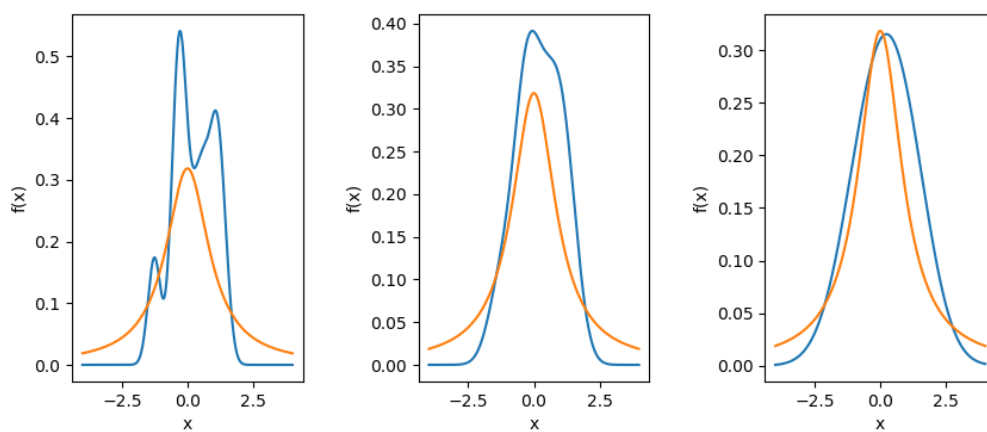


Рис. 9. Распределение Коши, $n = 20$ (2)

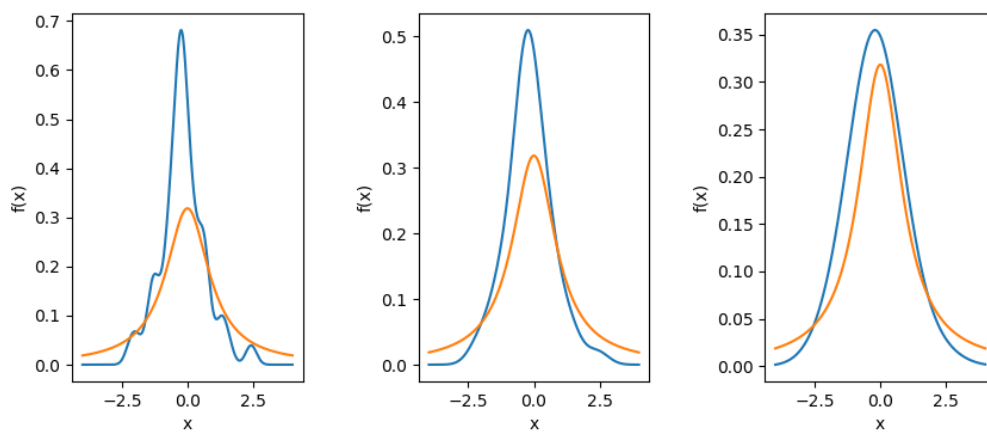


Рис. 10. Распределение Коши, $n = 60$

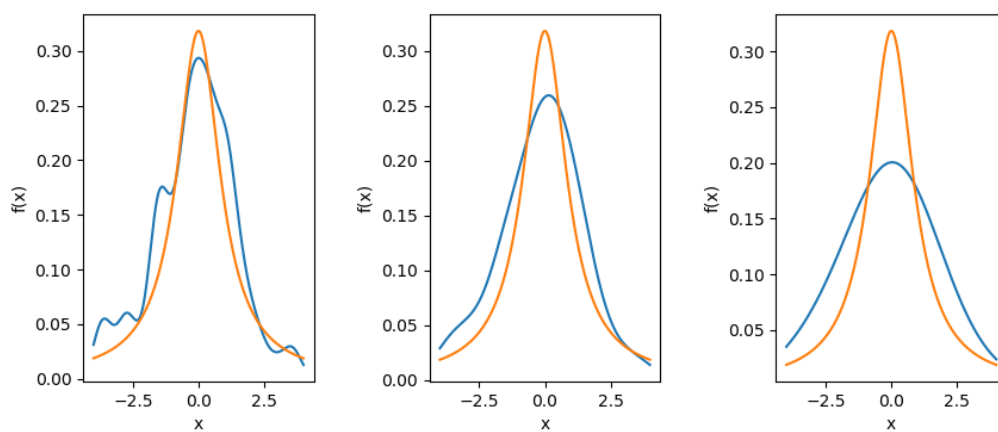


Рис. 11. Распределение Коши, $n = 100$

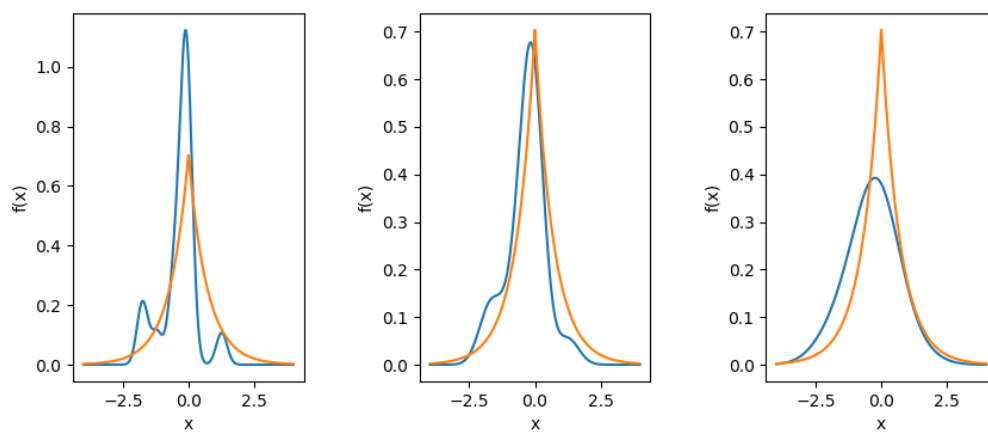


Рис. 12. Распределение Лапласа, $n = 20$ (3)

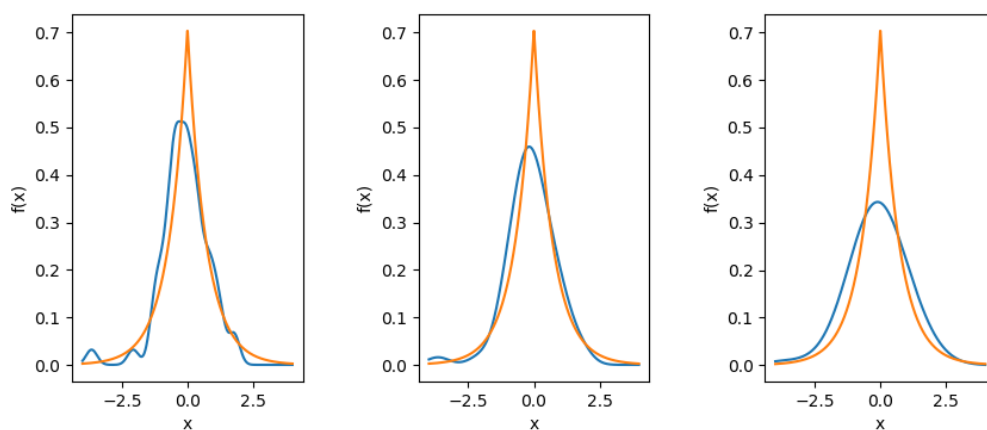


Рис. 13. Распределение Лапласа, $n = 60$

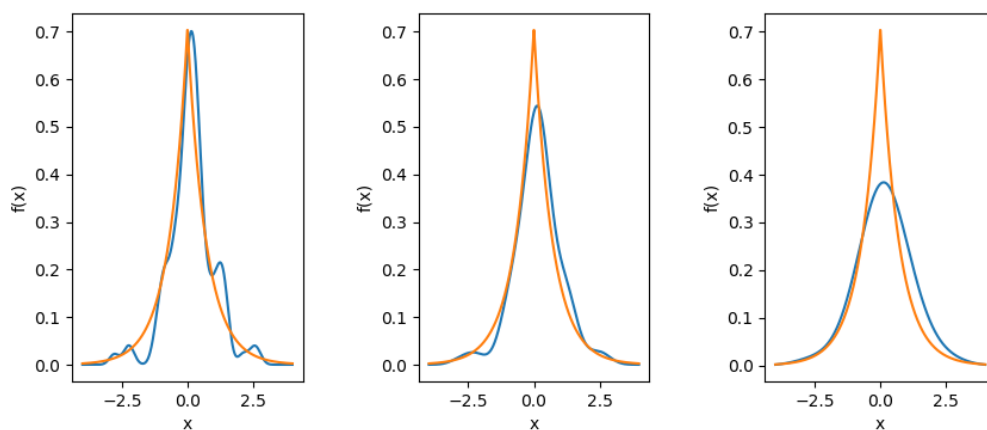


Рис. 14. Распределение Лапласа, $n = 100$

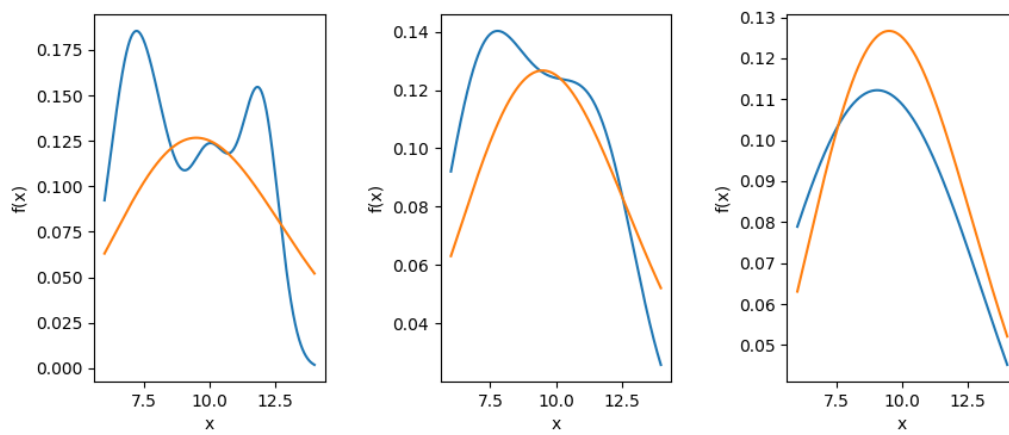


Рис. 15. Распределение Пуассона, $n = 20$ (4)

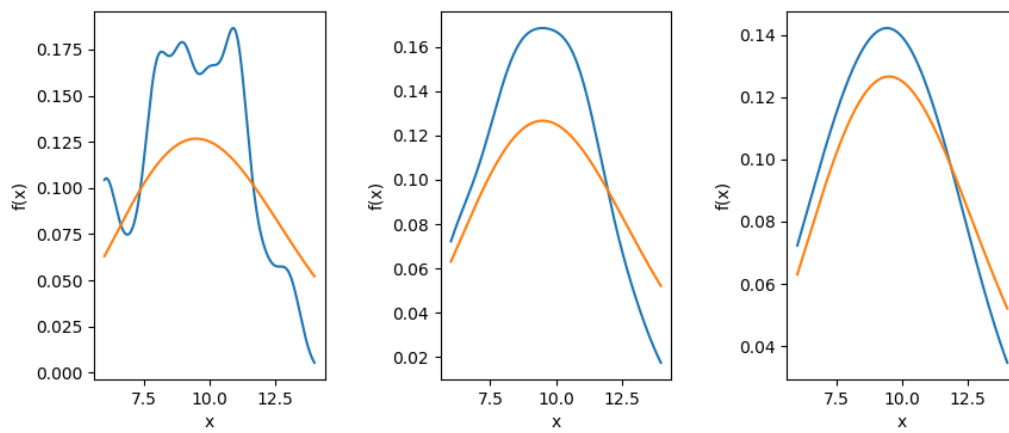


Рис. 16. Распределение Пуассона, $n = 60$

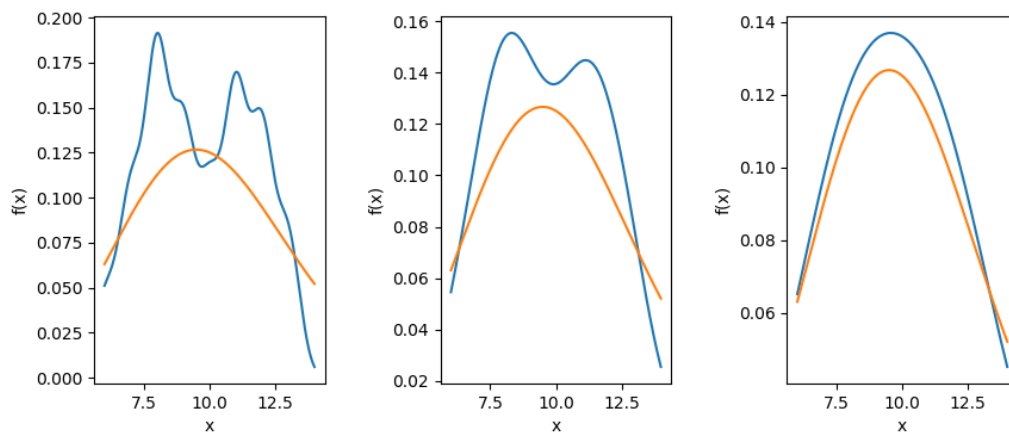


Рис. 17. Распределение Пуассона, $n = 100$

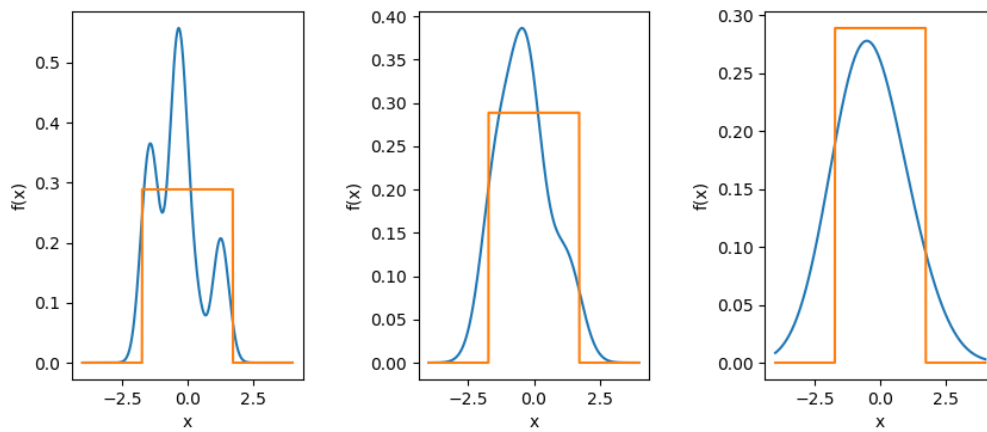


Рис. 18. Равномерное распределение, $n = 20$ (5)

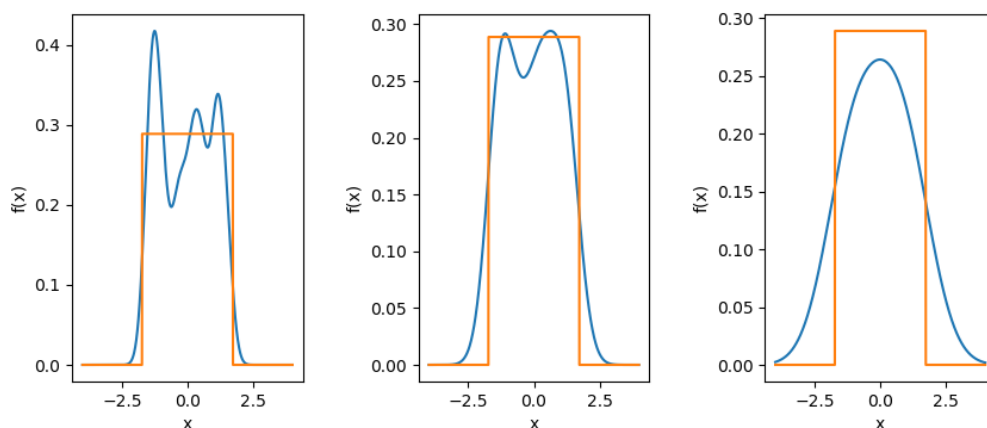


Рис. 19. Равномерное распределение, $n = 60$

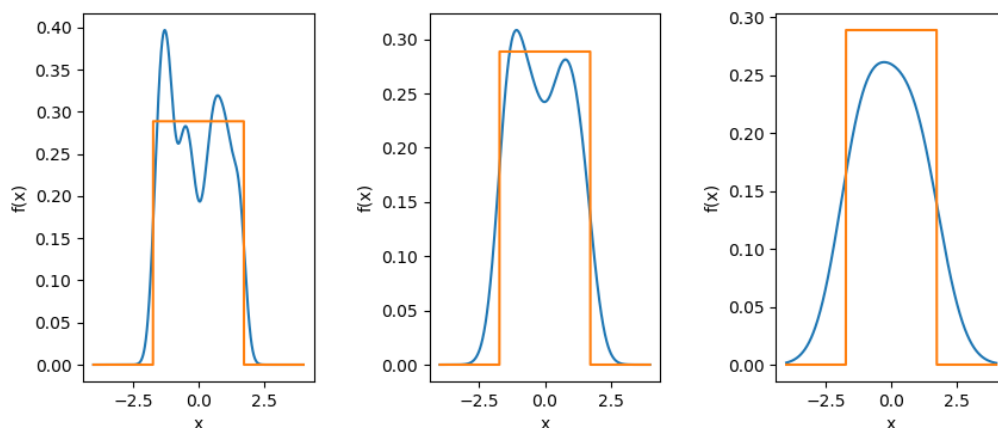


Рис. 20. Равномерное распределение, $n = 100$

5. Обсуждение

Из полученных данных видно, что точность аппроксимации эмпирической функцией (6) распределения увеличивается с увеличением выборки. Для распределения Пуассона (4) точность аппроксимации наименьшая. При увеличении размера выборки увеличивается точность аппроксимации плотности распределения для всех распределений кроме распределения Пуассона. Для нормального и равномерного распределения и распределения Лапласа лучше подходит параметр $h = h_n$. Для распределения коши и Пуассона лучше подходит параметр $h = \frac{h_n}{2}$.

6. Приложения

Репозиторий на GitHub с релизацией: github.com.