

Санкт-Петербургский политехнический университет Петра Великого
Институт прикладной математики и механики
Кафедра «Прикладная математика»

**ОТЧЁТ ПО ЛАБОРАТОРНЫМ РАБОТАМ
ПО ДИСЦИПЛИНЕ «МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА»**

Выполнил
студент группы 3630102/70401

Мельникова Анна Николаевна

Проверил
к. ф.-м. н., доцент

Баженов Александр Николаевич

Санкт-Петербург
2020

Содержание

1	Постановка задачи	4
1.1	Задание 1	4
1.2	Задание 2	4
1.3	Задание 3	4
1.4	Задание 4	4
2	Теория	5
2.1	Распределения	5
2.2	Гистограмма	5
2.2.1	Определение	5
2.2.2	Графическое описание	5
2.2.3	Использование	6
2.3	Вариационный ряд	6
2.4	Выборочные числовые характеристики	6
2.4.1	Характеристики положения	6
2.4.2	Характеристики рассеяния	7
2.5	Боксплот Тьюки	7
2.5.1	Определение	7
2.5.2	Описание	7
2.5.3	Построение	7
2.6	Теоретическая вероятность выбросов	7
2.7	Эмпирическая функция распределения	8
2.7.1	Статистический ряд	8
2.7.2	Определение	8
2.7.3	Описание	8
2.8	Оценки плотности вероятности	9
2.8.1	Определение	9
2.8.2	Ядерные оценки	9
3	Реализация	10
4	Результаты	10
4.1	Гистограммы и графики плотности распределения	10
4.2	Характеристики положения и рассеяния	12
4.3	Боксплот Тьюки	13
4.4	Доля выбросов	16
4.5	Теоретическая вероятность выбросов	17
4.6	Эмпирическая функция распределения	17
4.7	Ядерные оценки плотности распределения	20
5	Обсуждение	28
5.1	Гистограмма и график плотности распределения	28
5.2	Характеристики положения и рассеяния	28
5.3	Доля и теоретическая вероятность выбросов	28

5.4	Эмпирическая функция и ядерные оценки плотности распределения	28
6	Приложения	29
	Литература	30

Список иллюстраций

1	Нормальное распределение	10
2	Распределение Коши	10
3	Распределение Лапласа	11
4	Распределение Пуассона	11
5	Равномерное распределение	12
6	Нормальное распределение	14
7	Распределение Коши	14
8	Распределение Лапласа	15
9	Распределение Пуассона	15
10	Равномерное распределение	16
11	Нормальное распределение	17
12	Распределение Коши	18
13	Распределение Лапласа	18
14	Распределение Пуассона	19
15	Равномерное распределение	19
16	Нормальное распределение, $n = 20$	20
17	Нормальное распределение, $n = 60$	21
18	Нормальное распределение, $n = 100$	21
19	Распределение Коши, $n = 20$	22
20	Распределение Коши, $n = 60$	22
21	Распределение Коши, $n = 100$	23
22	Распределение Лапласа, $n = 20$	23
23	Распределение Лапласа, $n = 60$	24
24	Распределение Лапласа, $n = 100$	24
25	Распределение Пуассона, $n = 20$	25
26	Распределение Пуассона, $n = 60$	25
27	Распределение Пуассона, $n = 100$	26
28	Равномерное распределение, $n = 20$	26
29	Равномерное распределение, $n = 60$	27
30	Равномерное распределение, $n = 100$	27

Список таблиц

1	Статистический ряд	8
2	Статистический ряд	9
3	Нормальное распределение	12

4	Распределение Коши	12
5	Распределение Лапласа	13
6	Распределение Пуассона	13
7	Равномерное распределение	13
8	Доля выбросов	16
9	Теоретическая вероятность выбросов	17

1 Постановка задачи

Для 5 распределений:

1. $N(x, 0, 1)$ – нормальное распределение
2. $C(x, 0, 1)$ – распределение Коши
3. $L(x, 0, \frac{1}{\sqrt{2}})$ – распределение Лапласа
4. $P(k, 10)$ – распределение Пуассона
5. $U(x, -\sqrt{3}, \sqrt{3})$ – равномерное распределение

1.1 Задание 1

Сгенерировать выборки размером 10, 50 и 1000 элементов.

Построить на одном рисунке гистограмму и график плотности распределения.

1.2 Задание 2

Сгенерировать выборки размером 10, 100 и 1000 элементов. Для каждой выборки вычислить следующие статистические характеристики положения данных: \bar{x} , $medx$, z_R , z_Q , z_{tr} . Повторить такие вычисления 1000 раз для каждой выборки и найти среднее характеристическое положение и их квадратов:

$$E(z) = \bar{z} \quad (1)$$

Вычислить оценку дисперсии по формуле:

$$D(z) = \overline{z^2} - \bar{z}^2 \quad (2)$$

Представить полученные данные в виде таблиц.

1.3 Задание 3

Сгенерировать выборки размером 20 и 100 элементов. Построить для них боксплот Тьюки. Для каждого распределения определить долю выбросов экспериментально (сгенерировав выборку, соответствующую распределению 1000 раз, и вычислив среднюю долю выбросов) и сравнить с результатами, полученными теоретически.

1.4 Задание 4

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

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

2.2 Гистограмма

2.2.1 Определение

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

2.2.2 Графическое описание

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 [1].

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 [1].

2.3 Вариационный ряд

Вариационным рядом называется последовательность элементов выборки, расположенных в неубывающем порядке. Одинаковые элементы повторяются [2, с. 409]. Запись вариационного ряда: $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. Элементы вариационного ряда $x_{(i)}$ ($i = 1, 2, \dots, n$) называются порядковыми статистиками.

2.4 Выборочные числовые характеристики

С помощью выборки образуются её числовые характеристики. Это числовые характеристики дискретной случайной величины X^* , принимающей выборочные значения $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ [2, с. 411].

2.4.1 Характеристики положения

- Выборочное среднее

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (8)$$

- Выборочная медиана

$$medx = \begin{cases} x_{(l+1)} & n = 2l + 1 \\ \frac{x_{(l)} + x_{(l+1)}}{2} & n = 2l \end{cases} \quad (9)$$

- Полусумма экстремальных выборочных элементов

$$z_R = \frac{x_{(1)} + x_{(n)}}{2} \quad (10)$$

- Полусумма квартилей

Выборочная квартиль z_p порядка p определяется формулой

$$z_p = \begin{cases} x_{([np]+1)} & np - \text{дробное} \\ x_{(np)} & np - \text{целое} \end{cases} \quad (11)$$

Полусумма квартилей

$$z_Q = \frac{z_{1/4} + z_{3/4}}{2} \quad (12)$$

- Усечённое среднее

$$z_{tr} = \frac{1}{n - 2r} \sum_{i=r+1}^{n-r} x_{(i)}, r \approx \frac{n}{4} \quad (13)$$

2.4.2 Характеристики рассеяния

Выборочная дисперсия

$$D = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (14)$$

2.5 Боксплот Тьюки

2.5.1 Определение

Боксплот (англ. box plot) — график, использующийся в описательной статистике, компактно изображающий одномерное распределение вероятностей.

2.5.2 Описание

Такой вид диаграммы в удобной форме показывает медиану, нижний и верхний квартили и выбросы. Несколько таких ящиков можно нарисовать бок о бок, чтобы визуально сравнивать одно распределение с другим; их можно располагать как горизонтально, так и вертикально. Расстояния между различными частями ящика позволяют определить степень разброса (дисперсии) и асимметрии данных и выявить выбросы [3].

2.5.3 Построение

Границами ящика служат первый и третий квартили, линия в середине ящика — медиана. Концы усов — края статистически значимой выборки (без выбросов). Длину «усов» определяют разность первого квартиля и полутора межквартильных расстояний и сумма третьего квартиля и полутора межквартильных расстояний. Формула имеет вид

$$X_1 = Q_1 + \frac{3}{2}(Q_3 - Q_1), X_2 = Q_3 + \frac{3}{2}(Q_3 - Q_1), \quad (15)$$

где X_1 - нижняя граница уса, X_2 - верхняя граница уса, Q_1 - первый квартиль, Q_3 - третий квартиль.

Данные, выходящие за границы усов (выбросы), отображаются на графике в виде маленьких кружков [3].

2.6 Теоретическая вероятность выбросов

По формуле (15) можно вычислить теоретические нижнюю и верхнюю границы уса (X_1^T , X_2^T соответственно). Выбросами считаются величины x , такие что:

$$\begin{cases} x < X_1^T \\ x > X_2^T \end{cases} \quad (16)$$

Теоретическая вероятность выбросов для непрерывных распределений

$$P_B^T = P(x < X_1^T) + P(x > X_2^T) = F(X_1^T) + (1 - F(X_2^T)), \quad (17)$$

где $F(X) = P(x \leq X)$ - функция распределения.

Теоретическая вероятность выбросов для дискретных распределений

$$P_B^T = P(x < X_1^T) + P(x > X_2^T) = (F(X_1^T) - P(x = X_1^T)) + (1 - F(X_2^T)), \quad (18)$$

где $F(X) = P(x \leq X)$ - функция распределения.

2.7 Эмпирическая функция распределения

2.7.1 Статистический ряд

Статистическим рядом называется последовательность различных элементов выборки z_1, z_2, \dots, z_k , расположенных в возрастающем порядке с указанием частот n_1, n_2, \dots, n_k , с которыми эти элементы содержатся в выборке. Статистический ряд обычно записывается в виде таблицы

z	z_1	z_1	...	z_k
n	n_1	n_2	...	n_k

Таблица 1: Статистический ряд

2.7.2 Определение

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

$$F_n^*(x) = P^*(X < x) \quad (19)$$

2.7.3 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 $P^*(X < x) = \frac{1}{n} \sum_{z_i < x} n_i$. Получаем

$$F^*(x) = \frac{1}{n} \sum_{z_i < x} n_i \quad (20)$$

$F^*(x)$ — функция распределения дискретной случайной величины X_* , заданной таблицей распределения

X^*	z_1	z_1	\dots	z_k
P	$\frac{n_1}{n}$	$\frac{n_2}{n}$	\dots	$\frac{n_k}{n}$

Таблица 2: Статистический ряд

Эмпирическая функция распределения является оценкой, т. е. приближённым значением, генеральной функции распределения

$$F_n^*(x) \approx F_X(x) \quad (21)$$

2.8 Оценки плотности вероятности

2.8.1 Определение

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

$$f(x) \approx \hat{f}(x) \quad (22)$$

2.8.2 Ядерные оценки

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

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - x_i}{h_n}\right) \quad (23)$$

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

$$h_n \xrightarrow{n \rightarrow \infty} 0, \quad \frac{h_n}{n^{-1}} \xrightarrow{n \rightarrow \infty} \infty \quad (24)$$

Такие оценки называются непрерывными ядерными [2, с. 421-423].

Замечание: Свойство, означающее сближение оценки с оцениваемой величиной при $n \rightarrow \infty$ в каком-либо смысле, называется состоятельностью оценки.

Если плотность $f(x)$ кусочно-непрерывная, то ядерная оценка плотности является состоятельной при соблюдении условий, накладываемых на параметр сглаживания h_n , а также на ядро $K(u)$.

Гауссово (нормальное) ядро [4, с. 38]

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad (25)$$

Правило Сильвермана [4, с. 44]

$$h_n = 1.06 \hat{\sigma} n^{-\frac{1}{5}}, \quad (26)$$

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

3 Реализация

Лабораторная работа выполнена с помощью встроенных средств языка программирования Python в среде разработки PyCharm (для визуализации использовался пакет matplotlib, для вычислений - numpy). Исходный код лабораторной работы и tex-файлы отчета приведены в приложении в виде ссылки на репозиторий GitHub.

4 Результаты

4.1 Гистограммы и графики плотности распределения

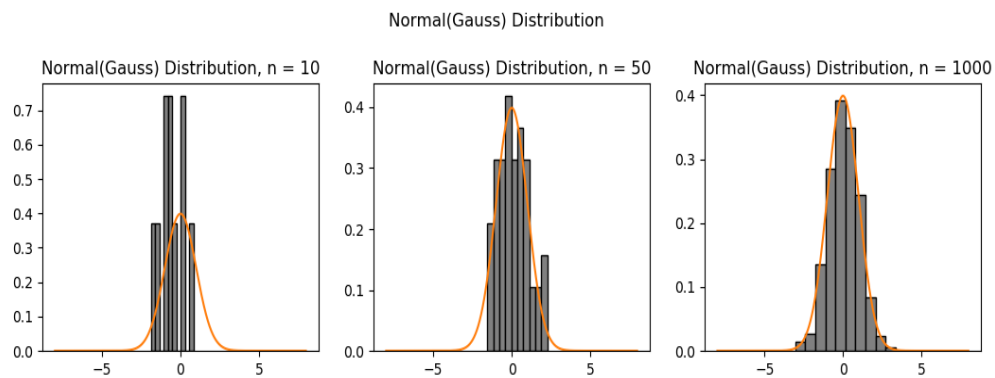


Рис. 1: Нормальное распределение

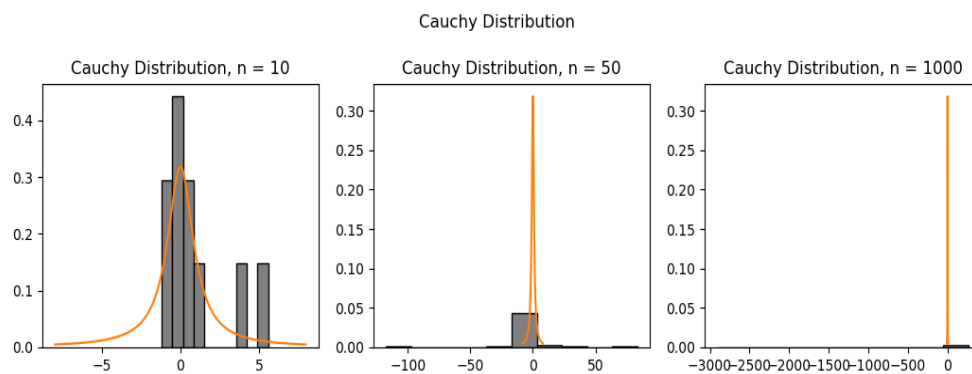


Рис. 2: Распределение Коши

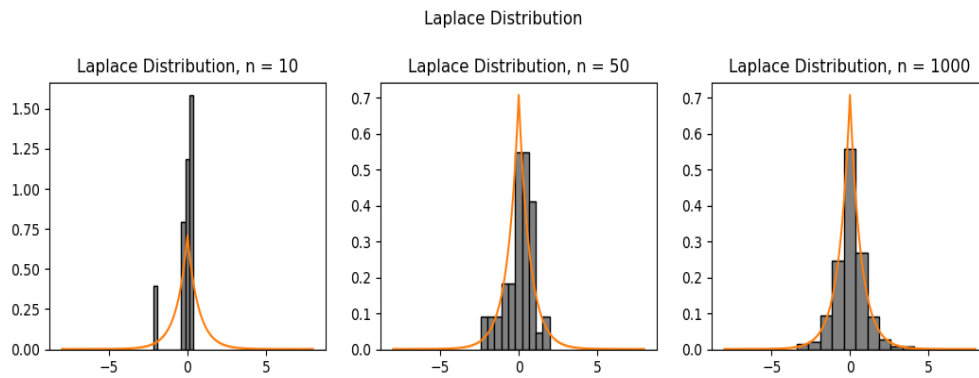


Рис. 3: Распределение Лапласа

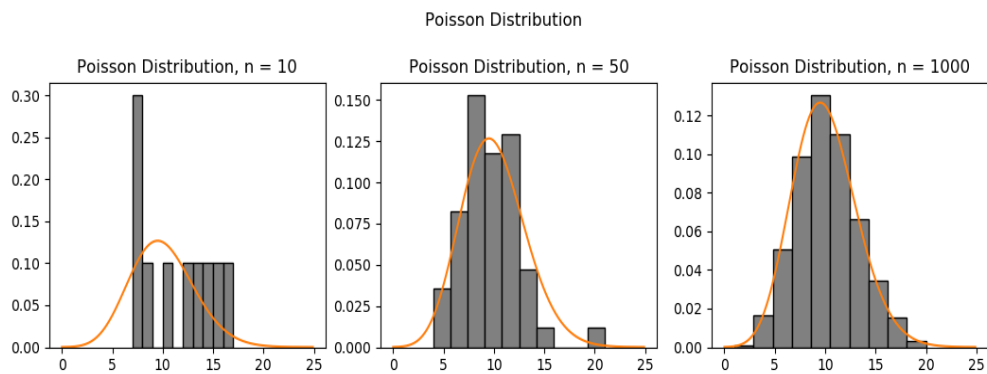


Рис. 4: Распределение Пуассона

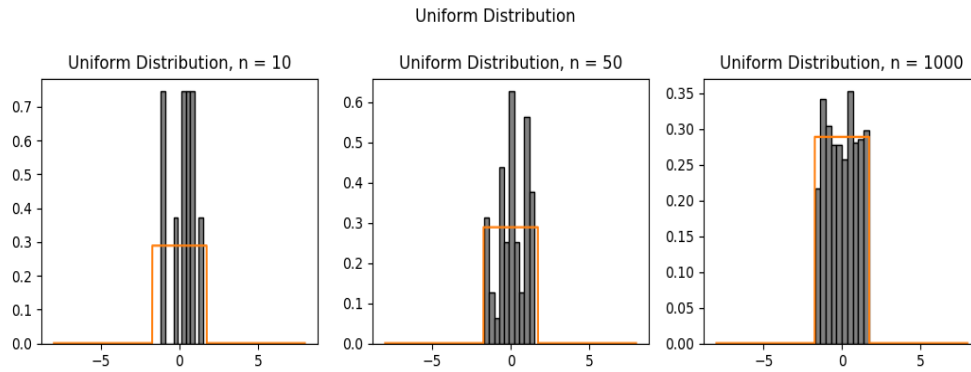


Рис. 5: Равномерное распределение

4.2 Характеристики положения и рассеяния

Мощность выборки указана в первом столбике справа.

Characteristic	Mean	Median	z_R	z_Q	z_tr
Normal E(z) n = 10	0.00	0.0	0.0	0.0	0.0
Normal D(z) n = 10	0.099688	0.136509	0.189635	0.111227	0.110212
Normal E(z) n = 100	-0.001	0.00	0.00	0.00	0.00
Normal D(z) n = 100	0.009908	0.01583	0.094601	0.012515	0.011928
Normal E(z) n = 1000	-0.0001	0.000	0.00	0.000	0.000
Normal D(z) n = 1000	0.000974	0.001515	0.05807	0.001175	0.001128

Таблица 3: Нормальное распределение

Characteristic	Mean	Median	z_R	z_Q	z_tr
Cauchy E(z) n = 10	-0.6	0.0	-3.0	0.0	0.0
Cauchy D(z) n = 10	350.215	0.335402	8468.97	0.852087	0.491287
Cauchy E(z) n = 100	-2.1	0.00	-104.000000	0.00	0.00
Cauchy D(z) n = 100	2224.86	0.023412	5.40494e+06	0.051584	0.02502
Cauchy E(z) n = 1000	1.0	0.000	511.00000000	0.000	0.000
Cauchy D(z) n = 1000	1195.76	0.002459	2.96612e+08	0.004937	0.002572

Таблица 4: Распределение Коши

Characteristic	Mean	Median	z_R	z_Q	z_tr
Laplace E(z) n = 10	0.0	0.00	0.0	0.00	0.00
Laplace D(z) n = 10	0.109844	0.071526	0.473888	0.094274	0.075035
Laplace E(z) n = 100	0.00	0.001	0.0	0.00	-0.002
Laplace D(z) n = 100	0.010475	0.005896	0.391706	0.010239	0.006378
Laplace E(z) n = 1000	0.0006	0.0001	0.0	0.0008	0.0006
Laplace D(z) n = 1000	0.000941	0.000479	0.386371	0.000996	0.000569

Таблица 5: Распределение Лапласа

Characteristic	Mean	Median	z_R	z_Q	z_tr
Poisson E(z) n = 10	10.00	9.8	10.30	9.9	9.8
Poisson D(z) n = 10	1.06412	1.44673	2.07481	1.23124	1.18912
Poisson E(z) n = 100	10.00	9.8	10.9	9.9	9.8
Poisson D(z) n = 100	0.098619	0.20429	0.940904	0.141224	0.119525
Poisson E(z) n = 1000	10.001	9.996	11.6	9.994	9.86
Poisson D(z) n = 1000	0.009728	0.003984	0.695475	0.002396	0.010941

Таблица 6: Распределение Пуассона

Characteristic	Mean	Median	z_R	z_Q	z_tr
Uniform E(z) n = 10	0.01	0.0	0.01	0.0	0.0
Uniform D(z) n = 10	0.098933	0.223483	0.045204	0.141623	0.160647
Uniform E(z) n = 100	0.003	0.00	0.0001	0.00	0.00
Uniform D(z) n = 100	0.009705	0.029092	0.000589	0.013734	0.019446
Uniform E(z) n = 1000	0.000	0.002	-5.1e-05	0.000	-0.001
Uniform D(z) n = 1000	0.001027	0.003036	6e-06	0.001515	0.002039

Таблица 7: Равномерное распределение

4.3 Боксплот Тьюки

Для каждого распределения представлен боксплот Тьюки для выборок размером 20 и 100 элементов.

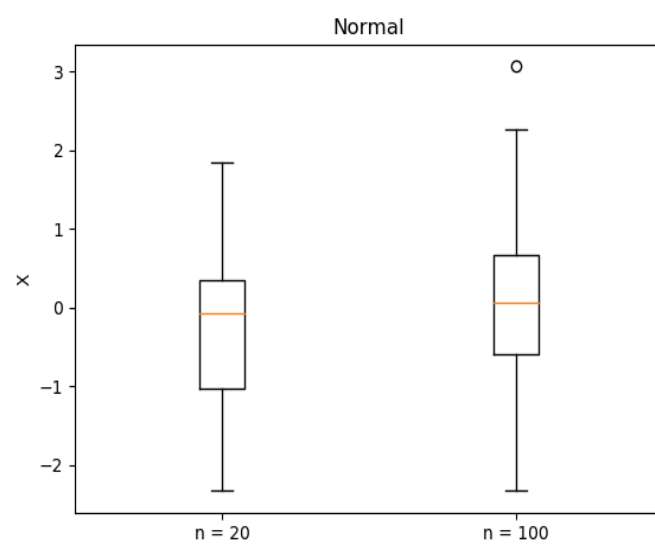


Рис. 6: Нормальное распределение

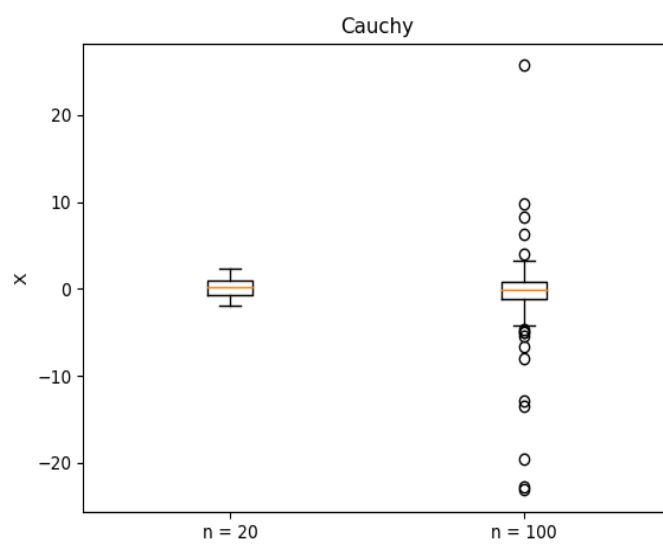


Рис. 7: Распределение Коши

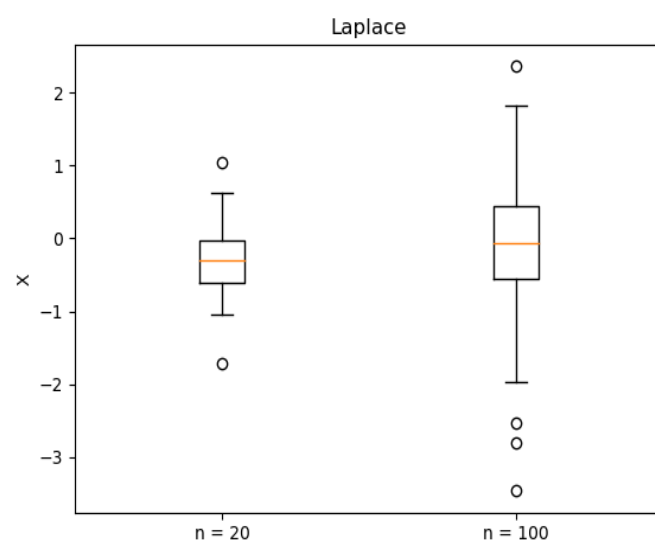


Рис. 8: Распределение Лапласа

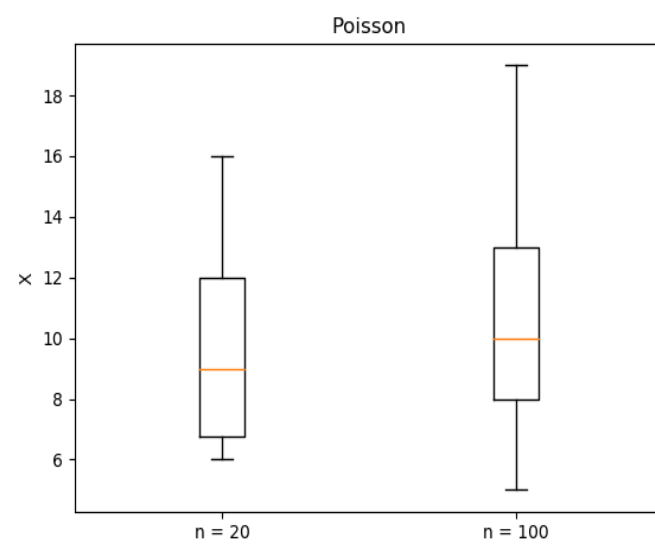


Рис. 9: Распределение Пуассона

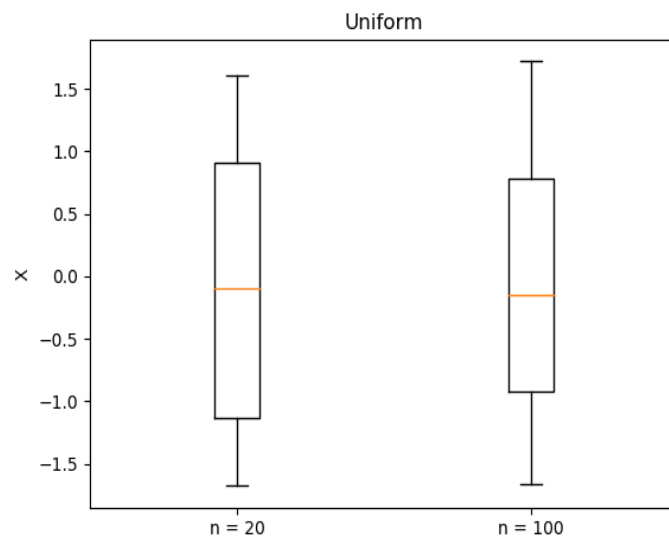


Рис. 10: Равномерное распределение

4.4 Доля выбросов

Выборка	Доля выбросов
Normal, n = 20	0.02
Normal, n = 100	0.01
Cauchy, n = 20	0.15
Cauchy, n = 100	0.16
Laplace, n = 20	0.08
Laplace, n = 100	0.06
Poisson, n = 20	0.02
Poisson, n = 100	0.01
Uniform, n = 20	0
Uniform, n = 100	0

Таблица 8: Доля выбросов

4.5 Теоретическая вероятность выбросов

Распределение	Q_1^T	Q_2^T	X_1^T	X_2^T	P_B^T
Нормальное распределение	-0.674	0.674	-2.698	2.698	0.007
Распределение Коши	-1	1	-4	4	0.156
Распределение Лапласа	-0.490	0.490	-1.961	1.961	0.063
Распределение Пуассона	8	12	2	18	0.008
Равномерное распределение	-0.866	0.866	-3.464	3.464	0

Таблица 9: Теоретическая вероятность выбросов

4.6 Эмпирическая функция распределения

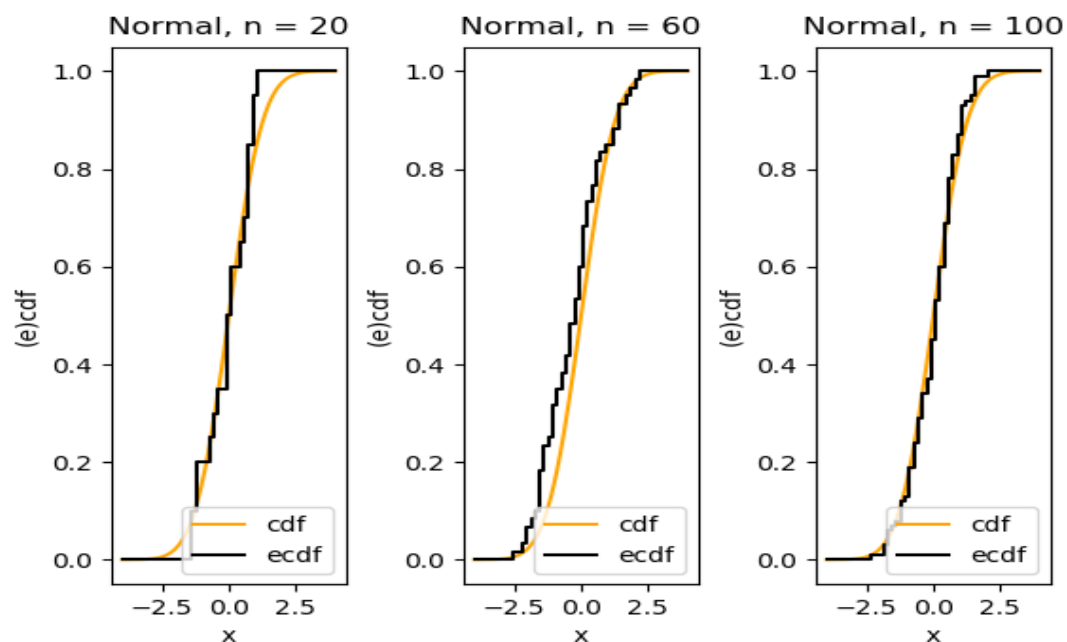


Рис. 11: Нормальное распределение

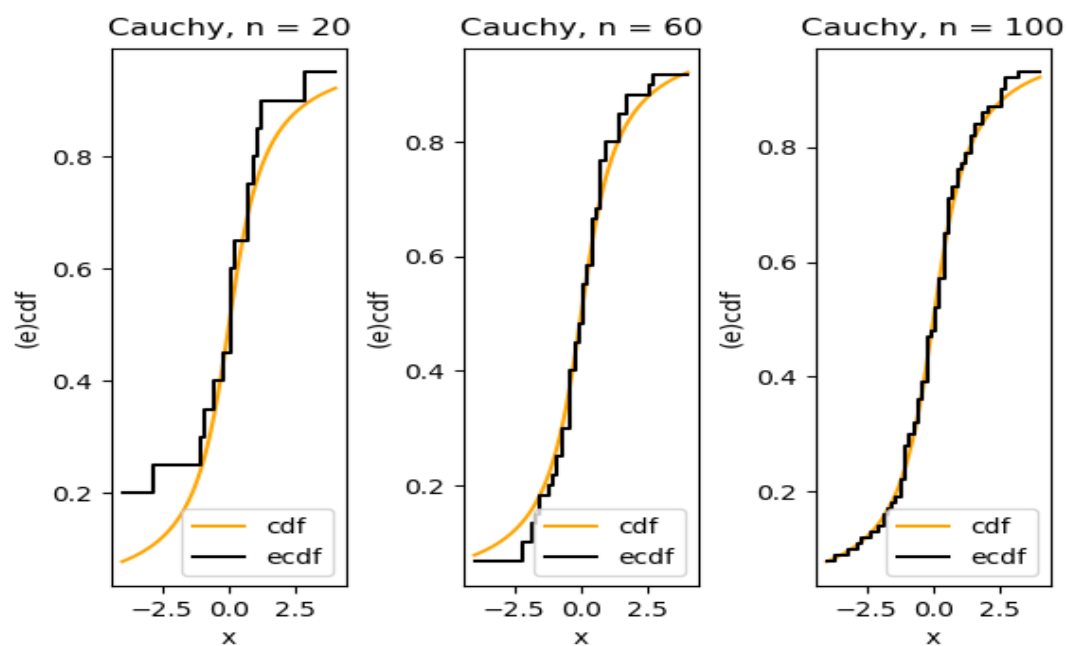


Рис. 12: Распределение Коши

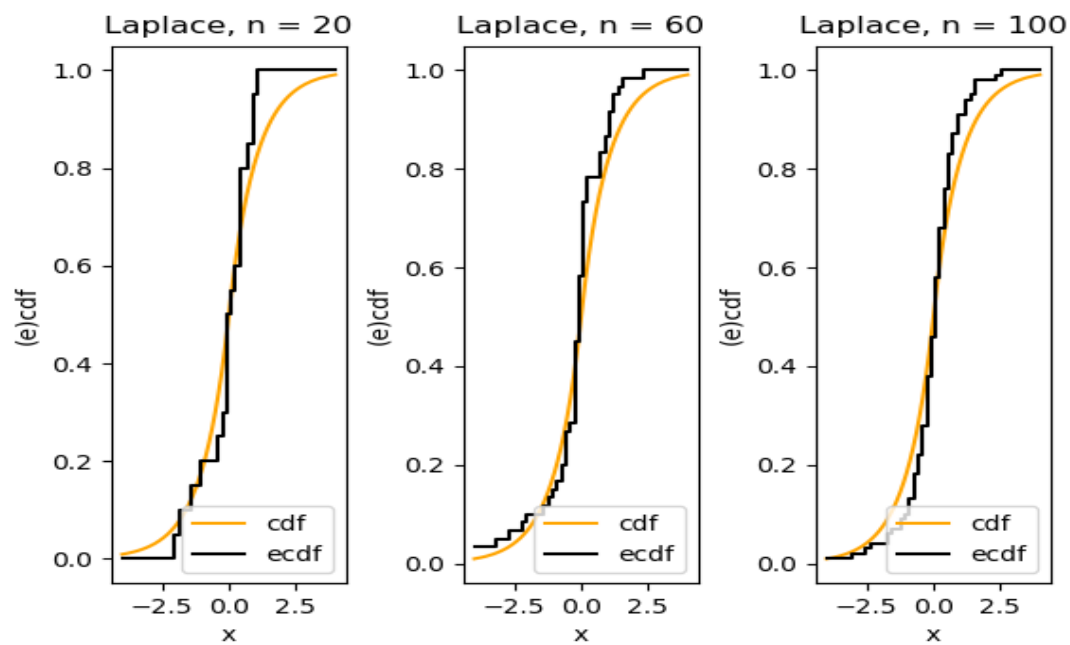


Рис. 13: Распределение Лапласа

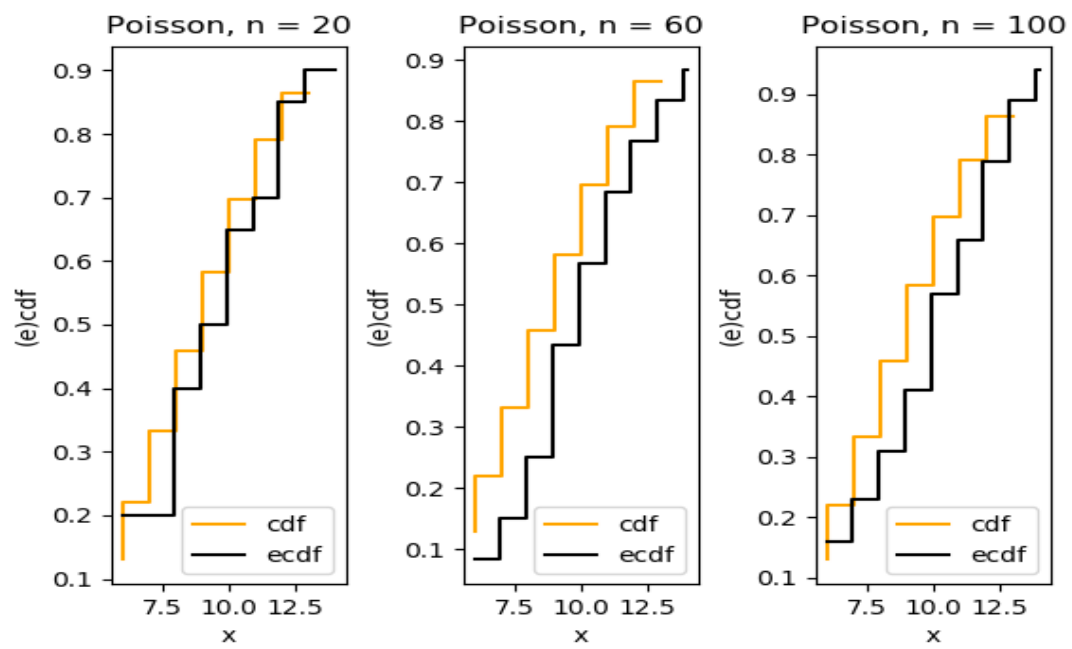


Рис. 14: Распределение Пуассона

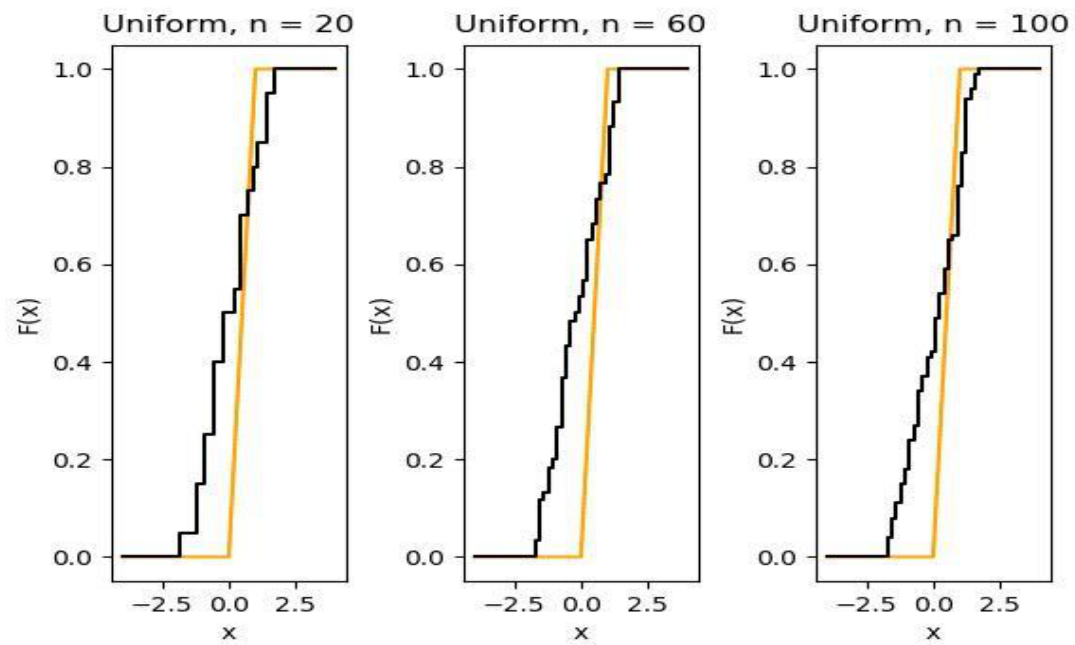


Рис. 15: Равномерное распределение

4.7 Ядерные оценки плотности распределения

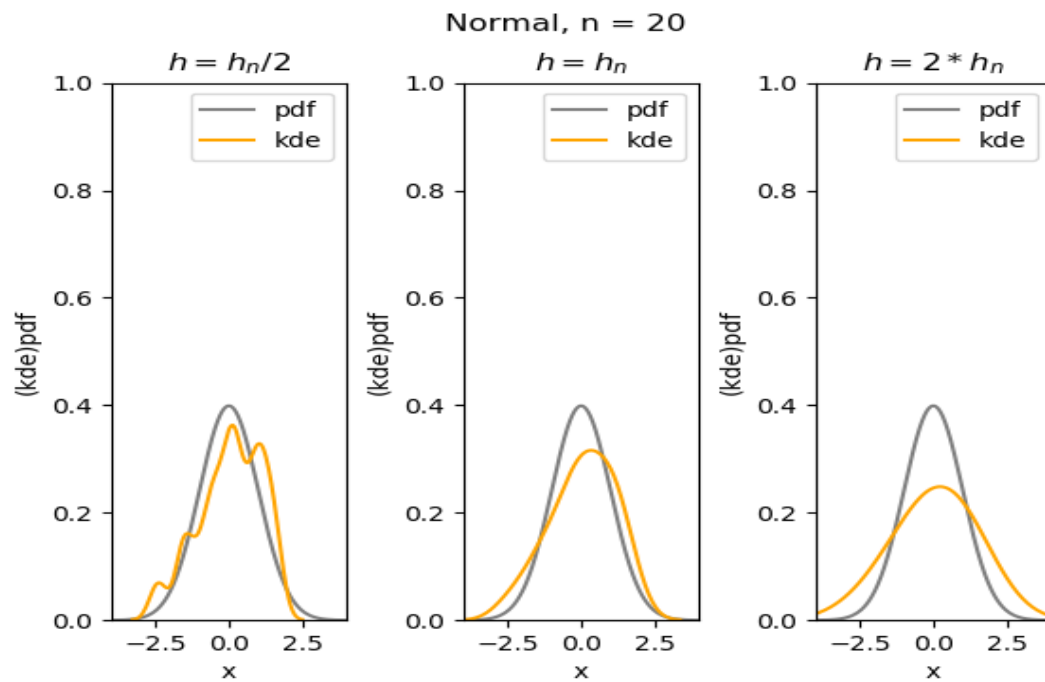


Рис. 16: Нормальное распределение, $n = 20$

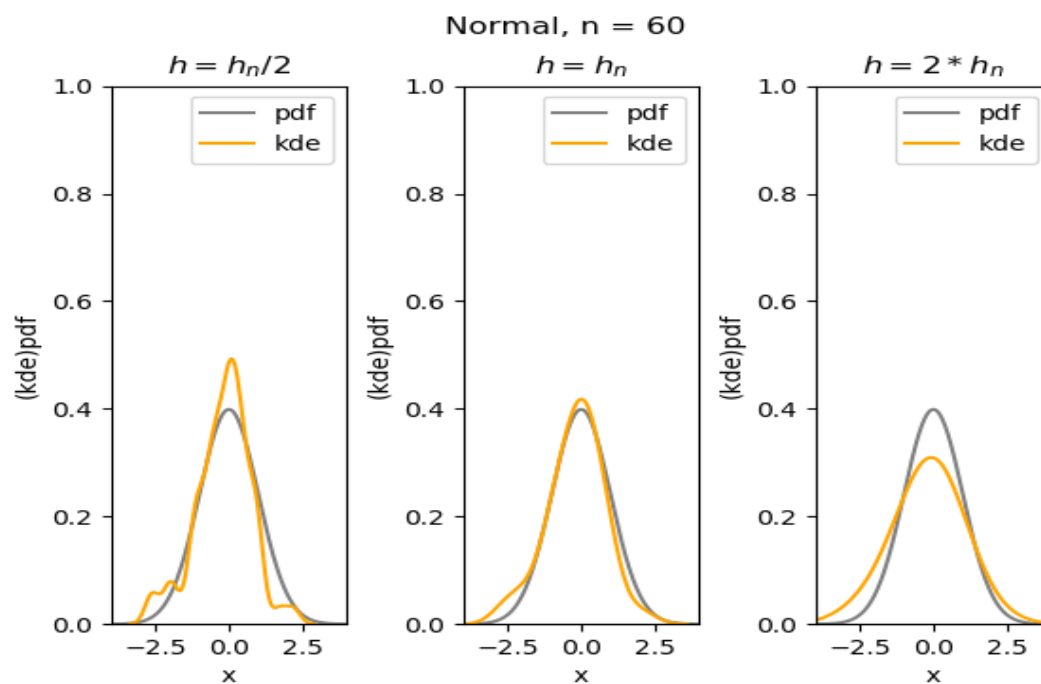


Рис. 17: Нормальное распределение, $n = 60$

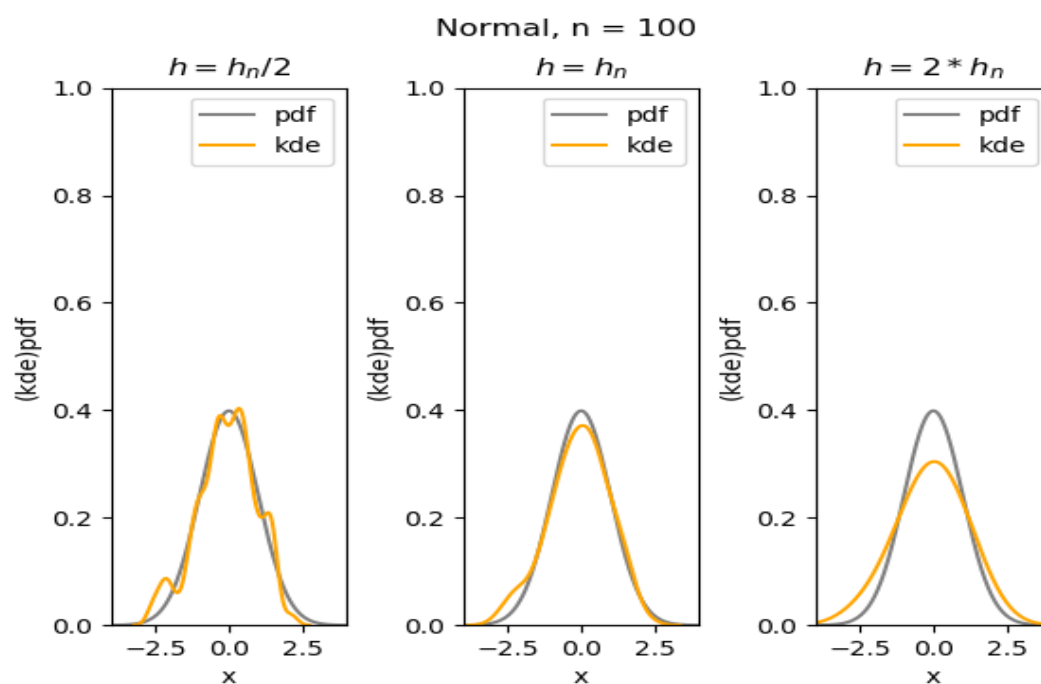


Рис. 18: Нормальное распределение, $n = 100$

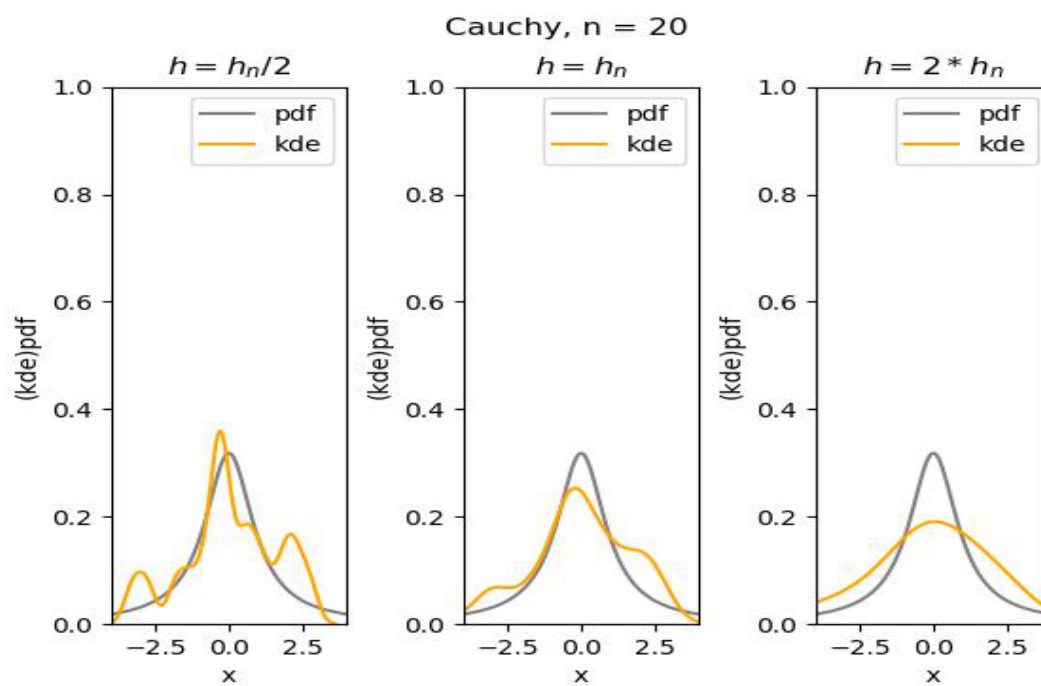


Рис. 19: Распределение Коши, $n = 20$

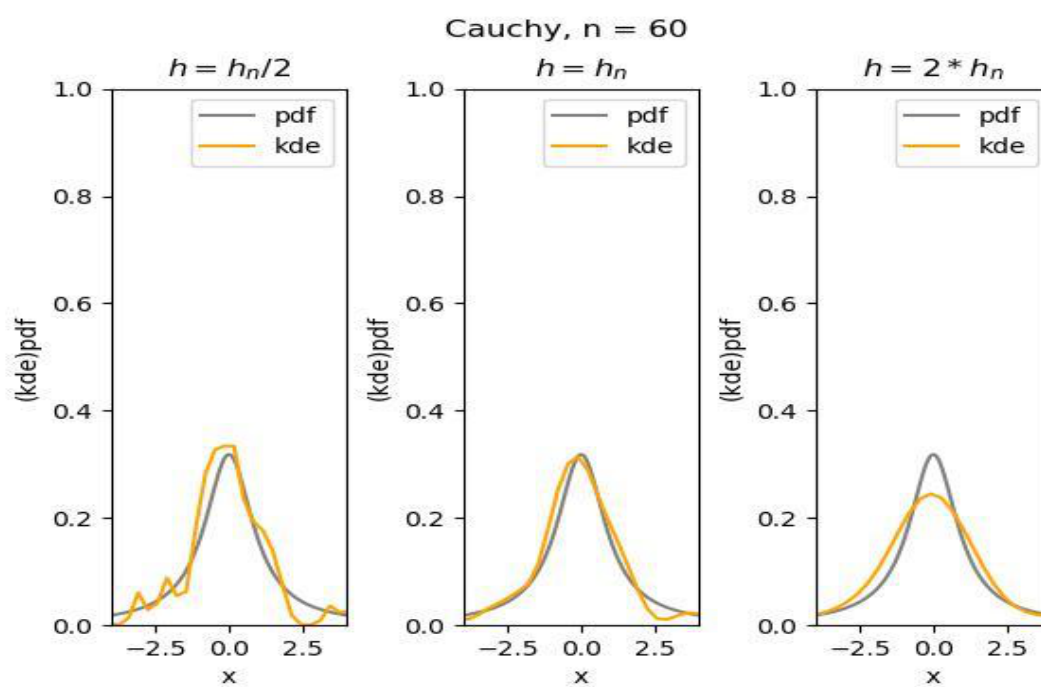


Рис. 20: Распределение Коши, $n = 60$

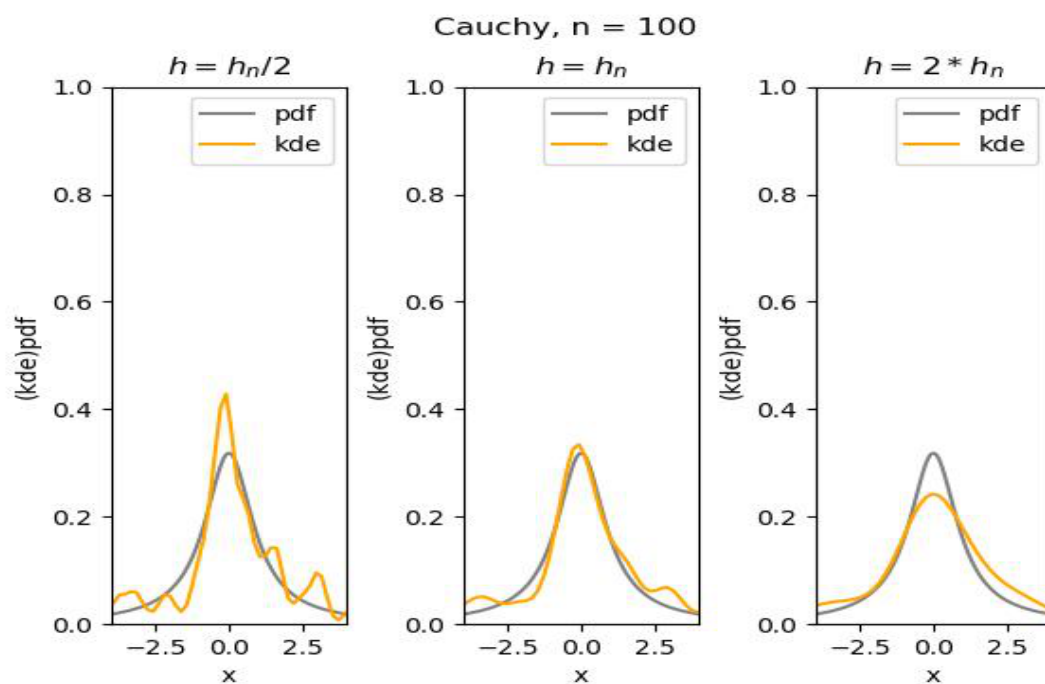


Рис. 21: Распределение Коши, $n = 100$

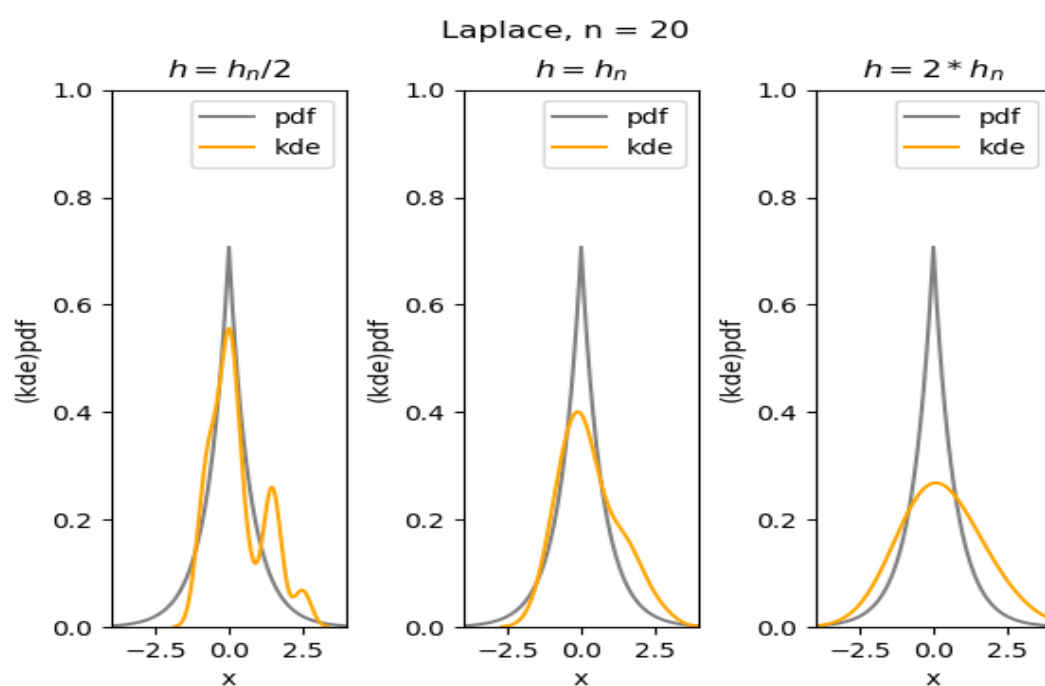


Рис. 22: Распределение Лапласа, $n = 20$

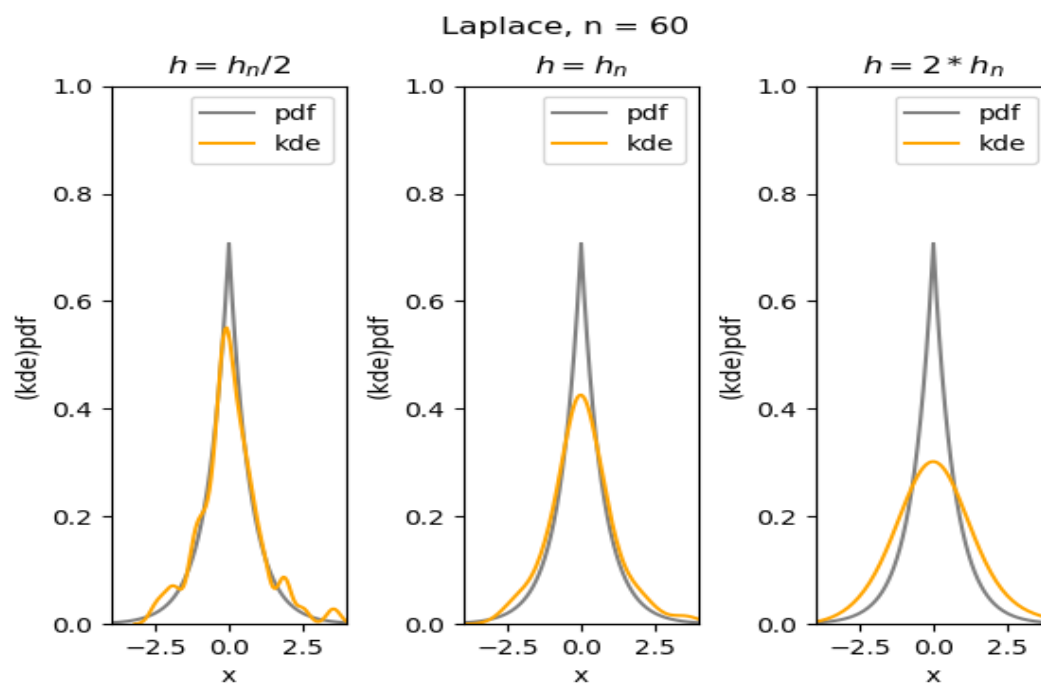


Рис. 23: Распределение Лапласа, $n = 60$

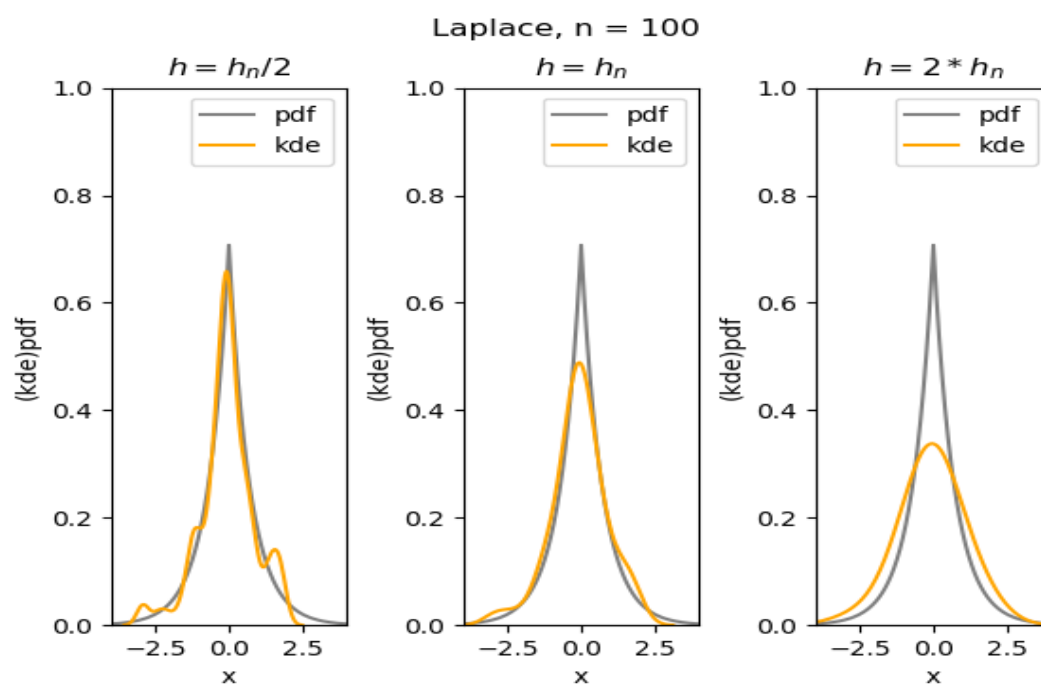


Рис. 24: Распределение Лапласа, $n = 100$

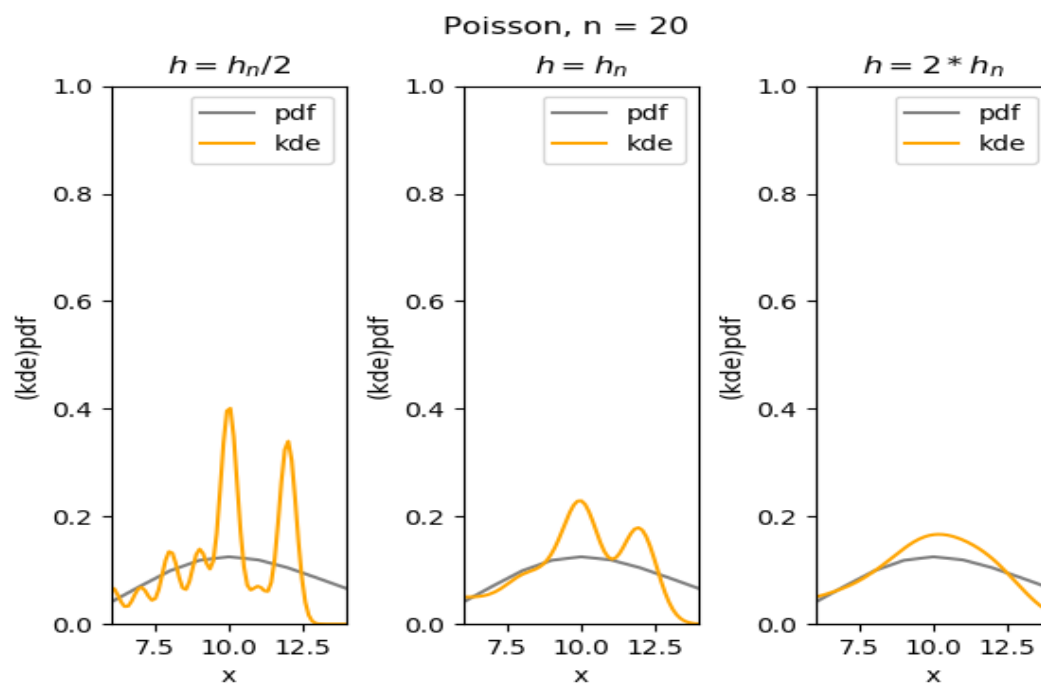


Рис. 25: Распределение Пуассона, $n = 20$

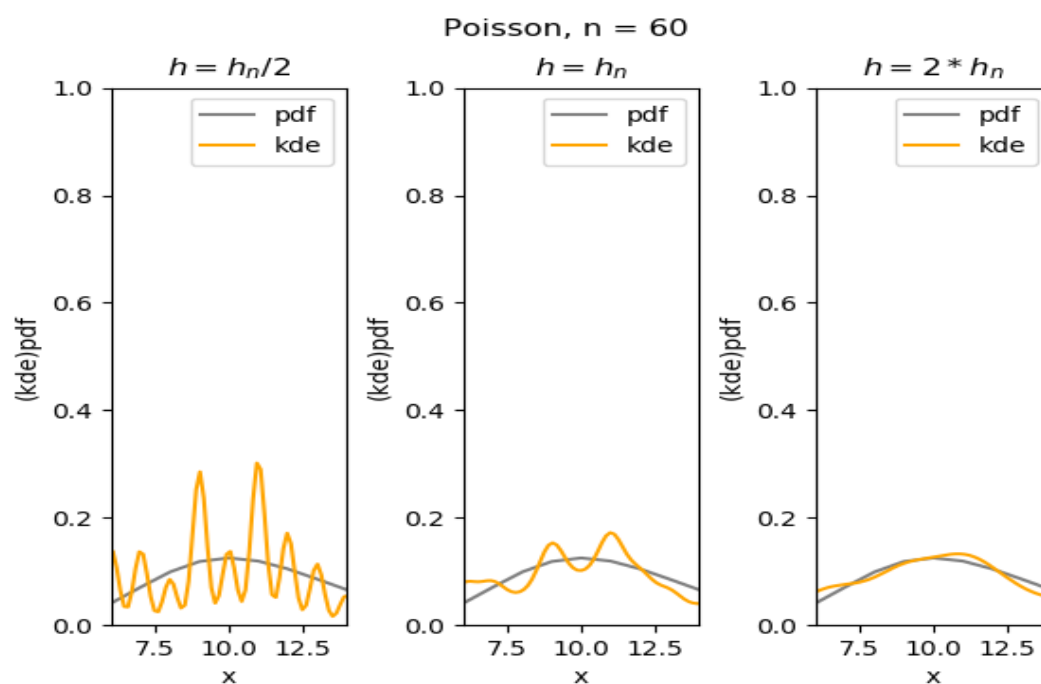


Рис. 26: Распределение Пуассона, $n = 60$

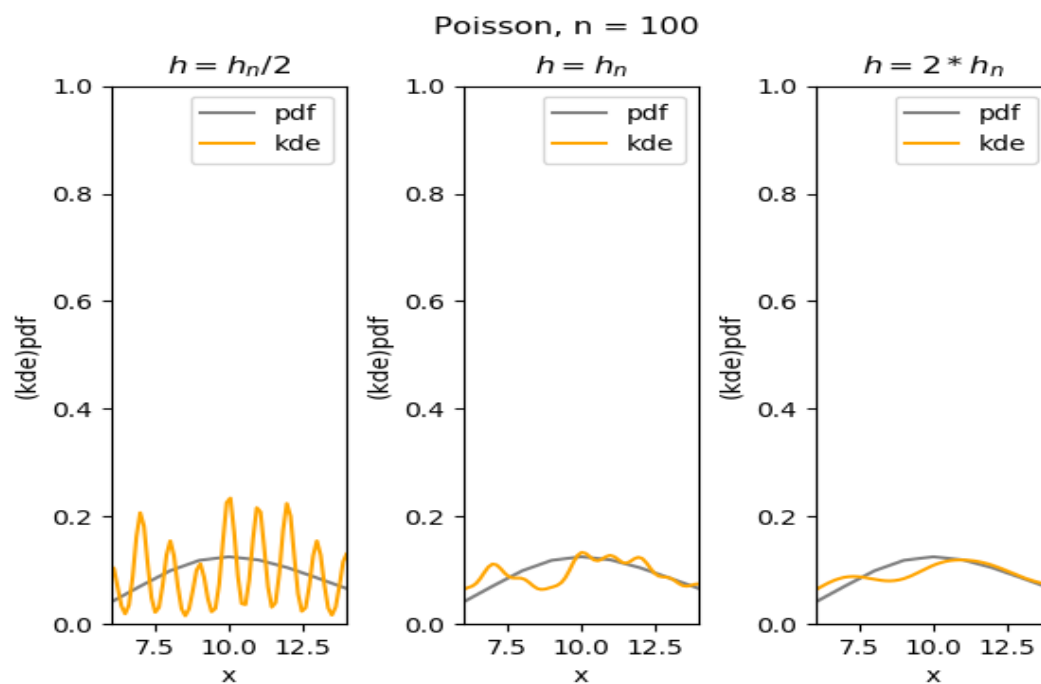


Рис. 27: Распределение Пуассона, $n = 100$

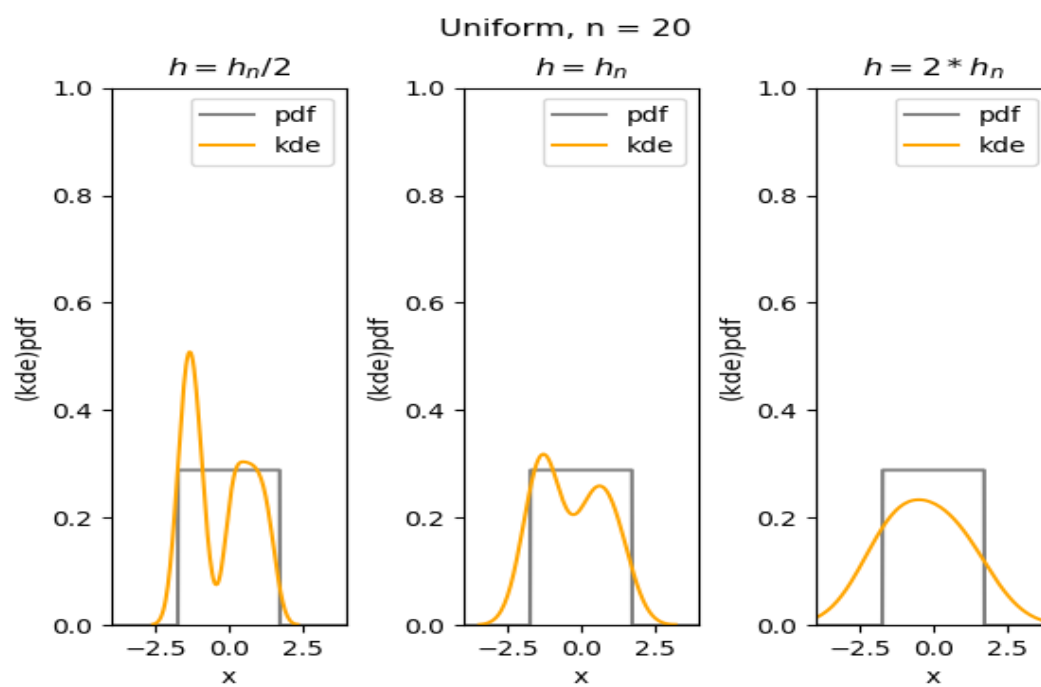


Рис. 28: Равномерное распределение, $n = 20$

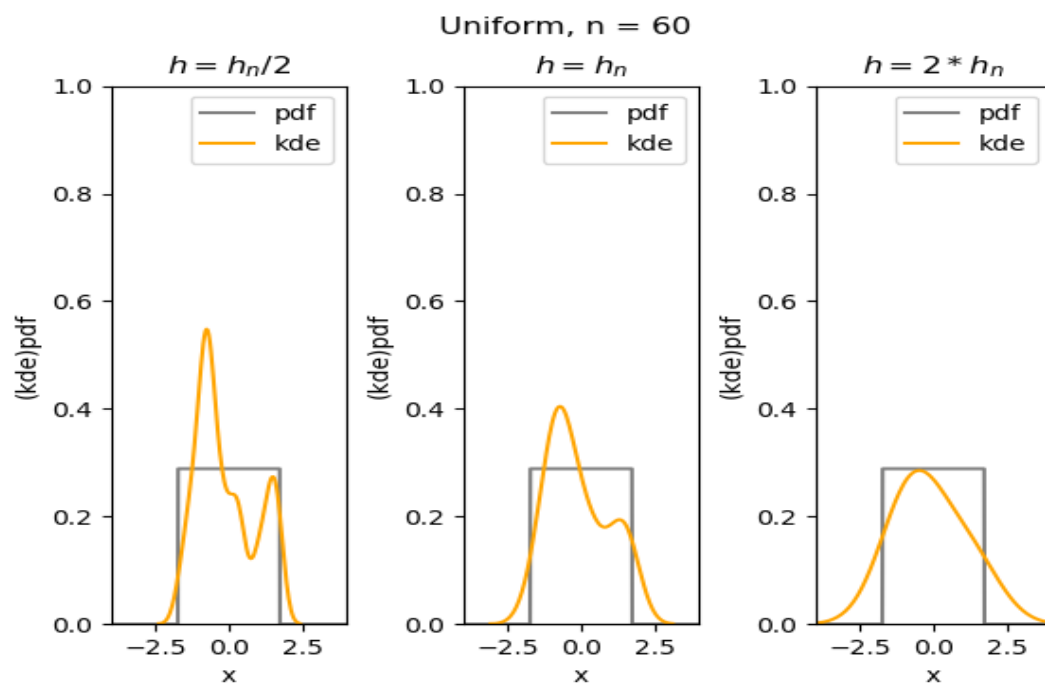


Рис. 29: Равномерное распределение, $n = 60$

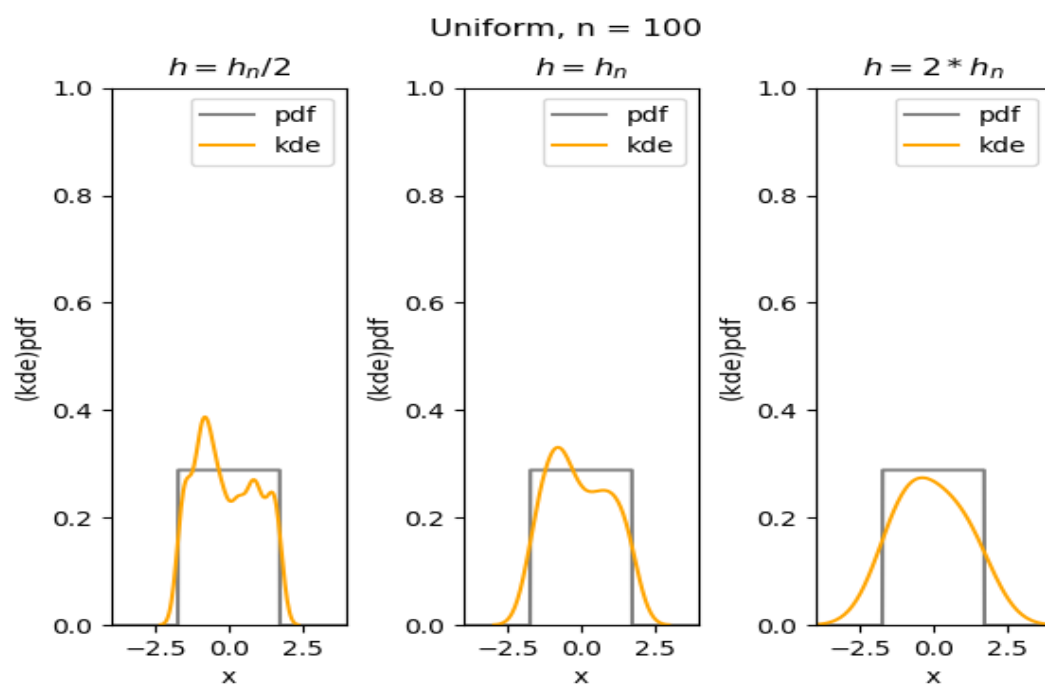


Рис. 30: Равномерное распределение, $n = 100$

5 Обсуждение

5.1 Гистограмма и график плотности распределения

По результатам проделанной работы можем сделать вывод о том, 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.

5.2 Характеристики положения и рассеяния

Заметим, что для распределения Коши $D(z)$ - быстро растет с ростом размера выборки, по которой берется z_R или \bar{x} , что отличает его от других типов распределений, для которых дисперсия характеристик рассеяния убывает с увеличением размера выборки - это происходит из-за характерных для распределения Коши выбросов, которые мы уже могли заметить на гистограммах в Задании 1.

5.3 Доля и теоретическая вероятность выбросов

Сравним долю выбросов определенную экспериментально с результатами, полученными теоретически. Видим точное соответствие с теорией для равномерного распределения - вероятность нулевая и выбросов мы не получили.

Результаты для выборок, сгенерированных в соответствии с законами распределения Лапласа и Коши, оказались близкими к теории, а доля выбросов для распределений Пуассона и Нормального ниже соответствующих теоретических оценок.

Заметим, что все распределения дают для большей выборки (100 элементов) результат ближе к теории, чем для меньшей выборки (20 элементов).

5.4 Эмпирическая функция и ядерные оценки плотности распределения

Можем наблюдать на рисунках (11) - (15), что ступенчатая эмпирическая функция распределения тем лучше приближает функцию распределения реальной выборки, чем мощнее эта выборка. Заметим так же, что для распределения Пуассона и равномерного распределения отклонение функций друг от друга наибольшее.

Рисунки (16) - (30) иллюстрируют сближение ядерной оценки и функции плотности вероятности для всех h с ростом размера выборки. Для распределения Пуассона наиболее ярко видно, как сглаживает отклонения увеличение параметра сглаживания h .

В зависимости от особенностей распределений для их описания лучше подходят разные параметры h в ядерной оценке: для равномерного распределения и распределения Пуассона лучше подойдет параметр $h = 2h_n$, для распределения Лапласа - $h = h_n/2$, а для нормального и Коши - $h = h_n$. Такие значения дают вид ядерной оценки наиболее близкий к плотности, характерной данным распределениям.

6 Приложения

Код программы - GitHub URL: <https://github.com/kaustika/Statistics2020>

Список литературы

- [1] Histogram. URL: <https://en.wikipedia.org/wiki/Histogram>
- [2] Вероятностные разделы математики. Учебник для бакалавров технических направлений.//Под ред. Максимова Ю.Д. — Спб.: «Иван Федоров», 2001. — 592 с., илл.
- [3] Box plot. URL: https://en.wikipedia.org/wiki/Box_plot
- [4] Анатольев, Станислав (2009) "Непараметрическая регрессия Квантиль, №7, стр. 37-52.