Санкт-Петербургский Политехнический Университет им. Петра Великого

Институт прикладной математики и механики

Кафедра прикладной математики

Отчёт по лабораторной работе №4 по дисциплине “Математическая статистика”

**Эмпирические функции и ядерные оценки**

Выполнил студент:

Мишутин Д. В.

Группа:

3630102/70301

Проверил:

К.ф.-м.н., доцент

Баженов Александр Николаевич

Санкт-Петербург

2020 г.

Оглавление

[1 Постановка задачи 4](#_Toc39161496)

[2 Теория 4](#_Toc39161497)

[3 Реализация 5](#_Toc39161498)

[4 Результаты 6](#_Toc39161499)

[4.1 Эмпирические функции распределения (ЭФР) 6](#_Toc39161500)

[4.2 Ядерные функции плотности (ЯФП) 8](#_Toc39161501)

[4.3 Таблицы. Средние модулей ошибок 15](#_Toc39161502)

[5 Выводы 16](#_Toc39161503)

[6 Литература 17](#_Toc39161504)

[7 Приложения 17](#_Toc39161505)

Список иллюстраций и таблиц

ЭФР. Стандартное нормальное распределение……………………………………………………………………………..……5

ЭФР. Стандартное распределение Коши…………………………………….…………………………………………………..……6

ЭФР. Распределение Лапласа…………………………………………………………………………………..……………………..……7

ЭФР. Равномерное распределение…………………………………………………………………………………………….…..……7

ЭФР. Распределение Пуассона………………………………………………………………………………………………………………8

ЯФП. Стандартное нормальное распределение при n=20…..……………………………………………………………..8

ЯФП. Стандартное нормальное распределение при n=60…..………………………………………………………………9

ЯФП. Стандартное нормальное распределение при n=100………………..………………………………………..…….9

ЯФП. Стандартное распределение Коши при n=20……………………………..…….………………………………..…….10

ЯФП. Стандартное распределение Коши при n=60……………………………………………………………………..…….10

ЯФП. Стандартное распределение Коши при n=100….………………………..…….………………………………..…….11

ЯФП. Распределение Лапласа при n=20….……………….…………………………………………………………………..…….11

ЯФП. Распределение Лапласа при n=60….…………….……………..……………………………………………………..…….12

ЯФП. Распределение Лапласа при n=100….…………….…………………………………………………………………..…….12

ЯФП. Равномерное распределение при n=20…………………….……..………………………………………………..…….13

ЯФП. Равномерное распределение при n=60….………………………..……………………..…………….…………..…….13

ЯФП. Равномерное распределение при n=100….………………………………………………………………………..…….14

ЯФП. Распределение Пуассона при n=20…………………….………..……..………………………..…………..………….….14

ЯФП. Распределение Пуассона при n=60….………………...……………………………………………………..………….….15

ЯФП. Распределение Пуассона при n=100….………………………..………………..………………………………………….15

Таблица. Стандартное нормальное распределение………………………………………………………………………….14

Таблица. Стандартное распределение Коши……………………………………………………………………………………..15

Таблица. Распределение Лапласа………………………………………………………………………………………………………16

Таблица. Равномерное распределение……………………………………………………………………………………………..16

Таблица. Распределение Пуассона………………………………………………………………………………………..…………..16

# 1 Постановка задачи

Сгенерировать выборки размером 20, 60 и 100 элементов. Построить на них эмпирические функции распределения и ядерные оценки плотности распределения на отрезке для непрерывных распределений и на отрезке для распределения Пуассона.

Распределения:

* Стандартное нормальное распределение:

|  |  |
| --- | --- |
|  | (1.1) |

* Стандартное распределение Коши:

|  |  |
| --- | --- |
|  | (1.2) |

* Распределение Лапласа:

|  |  |
| --- | --- |
|  | (1.3) |

* Распределение Пуассона:

|  |  |
| --- | --- |
|  | (1.4) |

* Равномерное распределение:

|  |  |
| --- | --- |
|  | (1.5) |

# **2 Теория**

*Эмпирической функцией распределения (ЭФР)* называется относительная частота события , полученная по данной выборке суммированием частот , для которых элементы статистического ряда меньше :

ЭФР является оценкой, то есть приближённым значением, генеральной функции распределения.

*Оценкой плотности вероятности* называется функция , построенная на основе выборки, приближённо равная :

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

Здесь называется *ядерной функцией плотности (ЯФП)*, непрерывна и является плотностью вероятности, – любая последовательность положительных чисел, обладающая свойствами:

Гауссово ядро:

Правило Сильвермана:

где – выборочное стандартное отклонение.

# 3 Реализация

Был использован язык *Python 3.8.2*: модуль *numpy* для генерации выборок с различными распределениями и вычисления описательных статистик, модуль *scipy* для функций и плотностей распределений, модуль *matplotlib* для построения и отображения графиков и гистограмм, модуль *math* для математических расчётов, модуль *pandas* для оптимального хранения статистических данных и функция *display* из модуля *IPython.display* для их корректного отображения в таблицах.

# 4 Результаты

## 4.1 Эмпирические функции распределения (ЭФР)

Рис. 1 ЭФР для Стандартного нормального распределения

Изображение выглядит как карта

Автоматически созданное описание

Рис. 2 ЭФР для Стандартного распределения Коши

Изображение выглядит как карта, текст

Автоматически созданное описание

Рис. 3 ЭФР для распределения Лапласа

Изображение выглядит как карта, компьютер

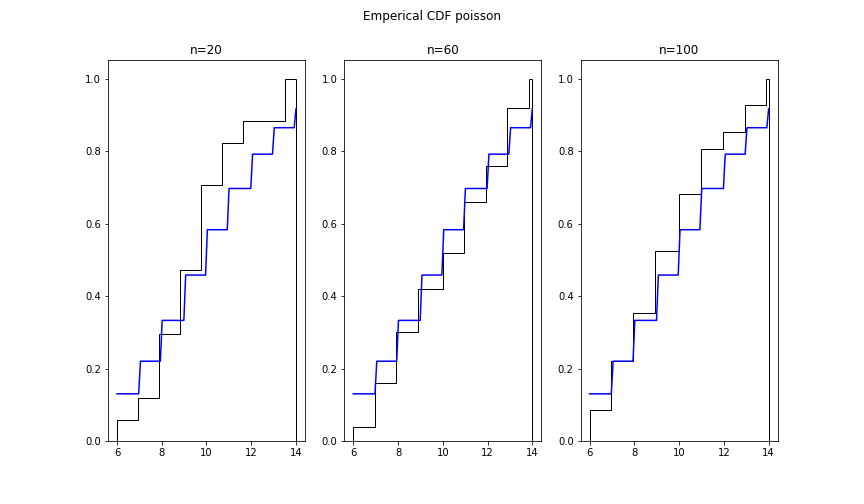
Автоматически созданное описание

Рис. 4 ЭФР для Равномерного распределения

Изображение выглядит как стол, компьютер, стоит

Автоматически созданное описание

Рис. 5 ЭФР для распределения Пуассона



## 4.2 Ядерные функции плотности (ЯФП)

Рис. 6 ЯФП для Стандартного нормального распределения при n=20

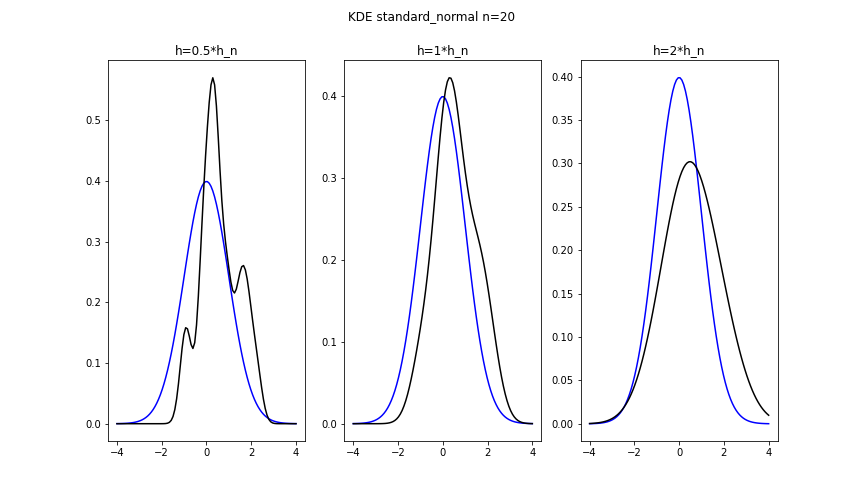


Рис. 7 ЯФП для Стандартного нормального распределения при n=60

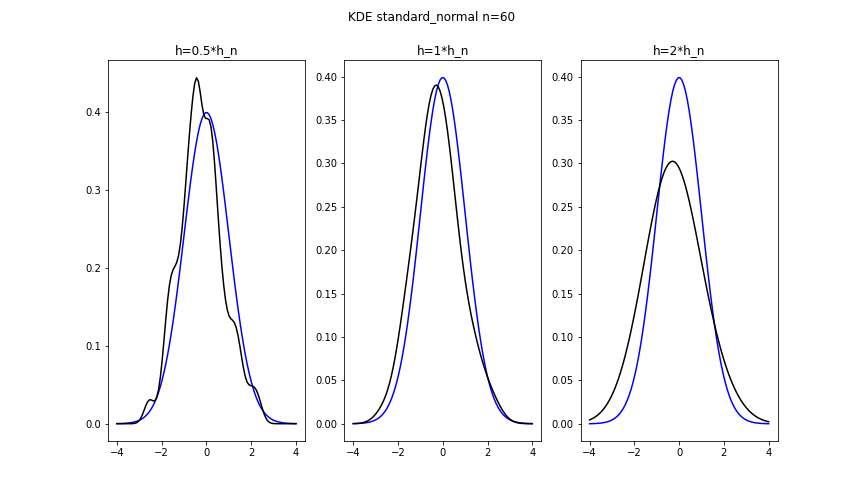


Рис. 8 ЯФП для Стандартного нормального распределения при n=100

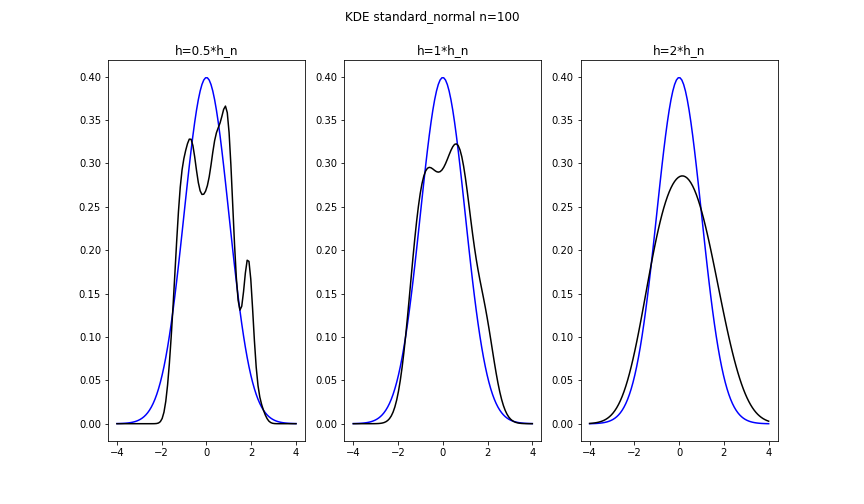


Рис. 9 ЯФП для Стандартного распределения Коши при n=20Изображение выглядит как карта

Автоматически созданное описание

Рис. 10 ЯФП для Стандартного распределения Коши при n=60

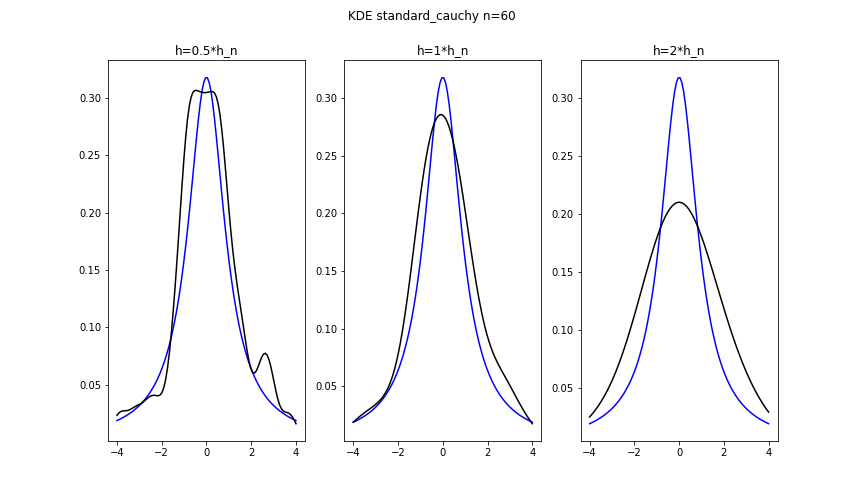


Рис. 11 ЯФП для Стандартного распределения Коши при n=100

Изображение выглядит как карта

Автоматически созданное описание

Рис. 12 ЯФП для распределения Лапласа при n=20

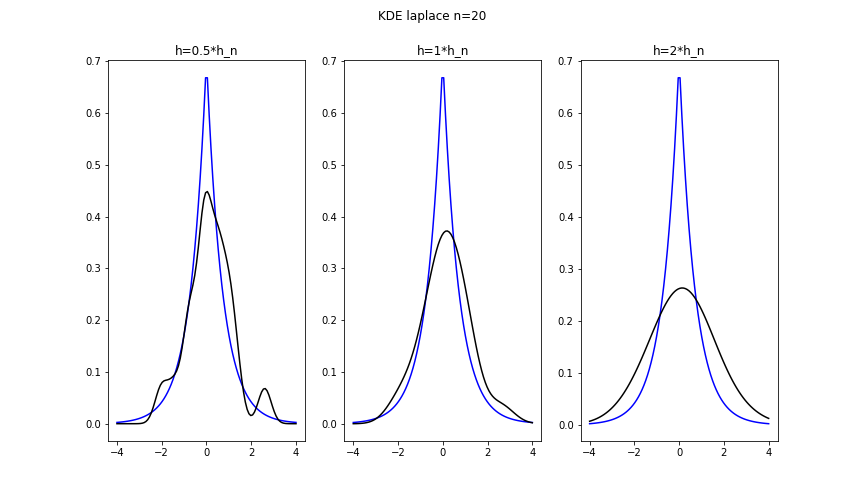


Рис. 13 ЯФП для распределения Лапласа при n=60

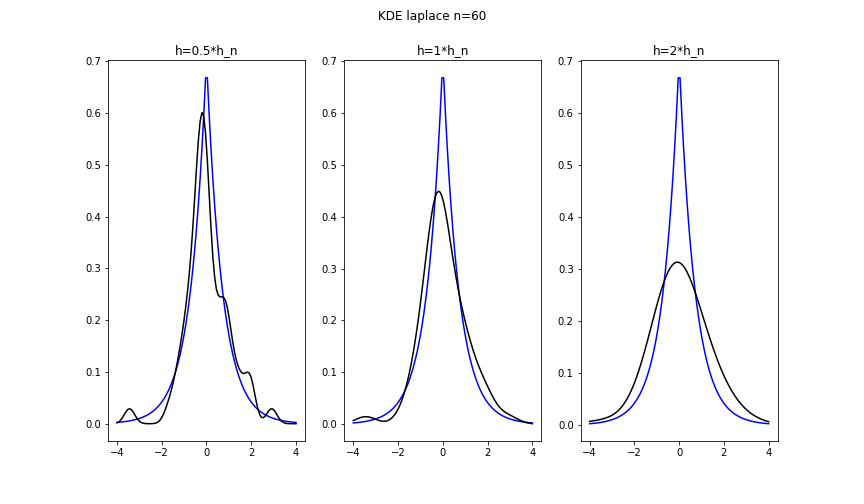


Рис. 14 ЯФП для распределения Лапласа при n=100

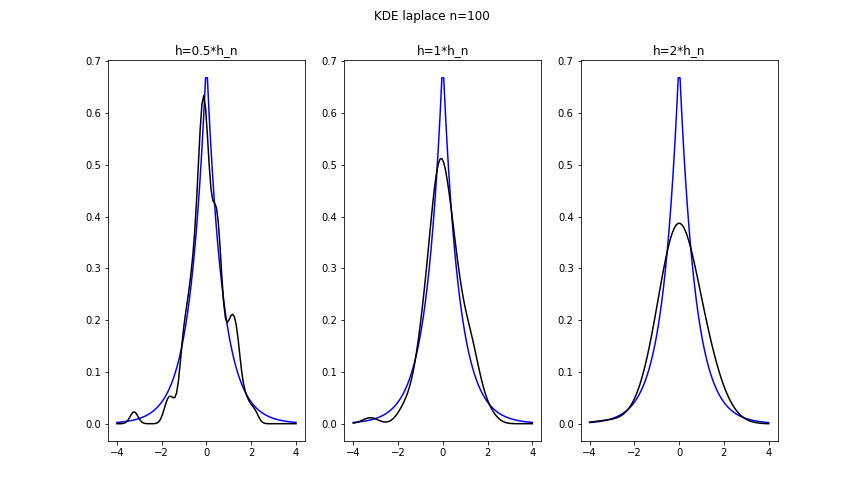


Рис. 15 ЯФП для Равномерного распределения при n=20Изображение выглядит как карта

Автоматически созданное описание

Рис. 16 ЯФП для Равномерного распределения при n=60

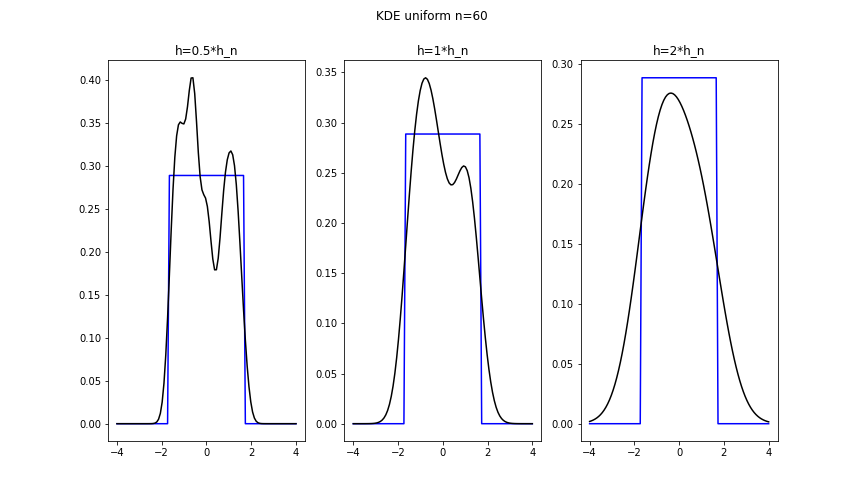


Рис. 17 ЯФП для Равномерного распределения при n=100

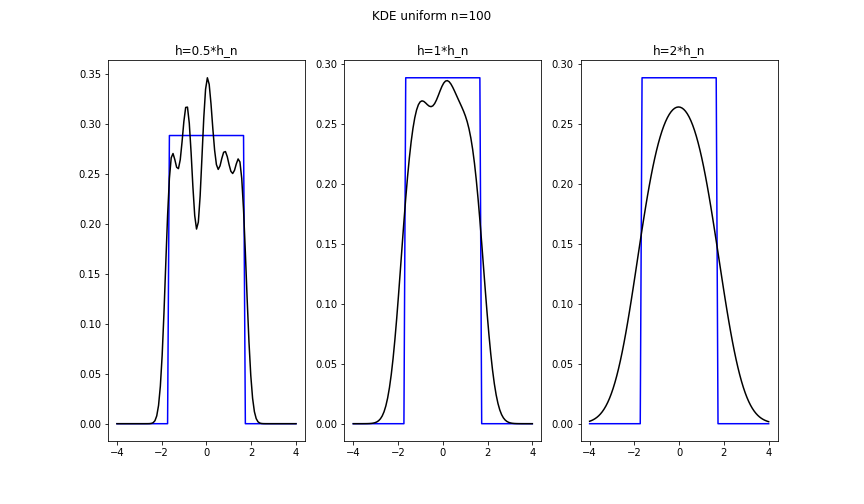


Рис. 18 ЯФП для распределения Пуассона при n=20

Изображение выглядит как карта

Автоматически созданное описание

Рис. 19 ЯФП для распределения Пуассона при n=60

Изображение выглядит как карта, текст

Автоматически созданное описание

Рис. 20 ЯФП для распределения Пуассона при n=100

Изображение выглядит как карта, текст

Автоматически созданное описание

## 4.3 Таблицы. Средние модулей ошибок

Строки – мощности выборок, столбцы – коэффициенты при , маркером помечены минимальные значения.

Таблица 1 Стандартное нормальное распределение

|  |  |  |  |
| --- | --- | --- | --- |
| **E(z)** | **0.5** | **1.0** | **2.0** |
| **n=20** | 0.061887 | 0.051959 | 0.052813 |
| **n=60** | 0.028175 | 0.026624 | 0.039276 |
| **n=100** | 0.035016 | 0.028130 | 0.037715 |

Таблица 2 Стандартное распределение Коши

|  |  |  |  |
| --- | --- | --- | --- |
| **E(z)** | **0.5** | **1.0** | **2.0** |
| **n=20** | 0.036591 | 0.037148 | 0.052669 |
| **n=60** | 0.022031 | 0.021232 | 0.039087 |
| **n=100** | 0.046387 | 0.042639 | 0.048923 |

Таблица 3 Распределение Лапласа

|  |  |  |  |
| --- | --- | --- | --- |
| **E(z)** | **0.5** | **1.0** | **2.0** |
| **n=20** | 0.035431 | 0.040874 | 0.067724 |
| **n=60** | 0.032295 | 0.030991 | 0.052448 |
| **n=100** | 0.024103 | 0.023139 | 0.034833 |

Таблица 4 Равномерное распределение

|  |  |  |  |
| --- | --- | --- | --- |
| **E(z)** | **0.5** | **1.0** | **2.0** |
| **n=20** | 0.041468 | 0.039627 | 0.052521 |
| **n=60** | 0.030078 | 0.033001 | 0.051592 |
| **n=100** | 0.025169 | 0.030812 | 0.054162 |

Таблица 5 Распределение Пуассона

|  |  |  |  |
| --- | --- | --- | --- |
| **E(z)** | **0.5** | **1.0** | **2.0** |
| **n=20** | 0.036118 | 0.020863 | 0.002852 |
| **n=60** | 0.022694 | 0.013810 | 0.011530 |
| **n=100** | 0.025138 | 0.017411 | 0.006402 |

# 5 Выводы

Чем больше выборка, тем лучше ЭФР приближает эталонную функцию распределения.

Исходя из данных в таблицах, для распределений

1. Стандартного нормального: лучшие оценки при
2. Стандартного Коши: для малых выборок лучшая оценка будет при , при больших –
3. Лапласа: для малых будет , для больших –
4. Равномерного: для малых – , для больших –
5. Пуассона: лучшие оценки при

# 6 Литература

[Основы работы с *numpy* (отдельная глава курса)](https://stepik.org/course/401)

[Документация по *scipy*](https://docs.scipy.org/doc/scipy/reference/index.html)

[Pandas обзор](https://www.dataquest.io/blog/pandas-cheat-sheet/)

# 7 Приложения

[Код лабораторной](https://github.com/MeShootIn/matstat/blob/master/lab_4/lab_4.ipynb)