

Домашнее задание 2

Дородный Дмитрий СКБ172

4 декабря 2019 г.

Содержание

1	Гамма распределение	1
1.1	Моделирование выборок	1
1.2	Эмпирическая функция	2
1.2.1	Практические результаты	2
1.2.2	Теоритический вывод	3
1.3	Вариационный ряд	3
1.4	Квантили	4
1.4.1	Практические результаты	4
1.4.2	Теоритический вывод	5
1.5	Гистограмма и полигон частот	5
1.5.1	Практические- результаты	5
1.5.2	Теоритический вывод	8
2	распределение Бореля-Таннера	8
2.1	Моделирование выборок	8
2.2	Эмпирическая функция	9
2.3	Вариационный ряд	10
2.4	Квантили	10
2.5	Гистограмма и полигон частот	11

1 Гамма распределение

1.1 Моделирование выборок

Выборки были смоделированы для параметра размера $\lambda = 5$ и параметра формы $\alpha = 3$

моделирование величин - сткоки 14-29

Ниже представлены по 5 смоделированных выборок размерами 5 и 10 величин:

$n = 5$

- 1) 1.382 0.759 1.826 1.778 1.926
- 2) 1.625 2.846 2.107 2.623 2.543
- 3) 2.749 1.326 2.437 2.761 1.228
- 4) 1.847 1.425 1.626 0.626 1.851
- 5) 2.816 3.370 1.766 1.472 1.421

$n = 10$

- 1) 2.741 3.378 1.785 1.375 1.491 1.441 1.615 2.553 1.042 1.234
- 2) 1.545 1.144 1.221 3.679 2.432 1.107 1.976 0.537 2.401 1.127
- 3) 2.867 0.769 2.173 0.803 1.952 2.081 2.082 0.986 1.476 0.610
- 4) 1.618 1.592 2.970 2.096 2.670 2.416 1.810 2.280 2.138 0.934
- 5) 1.736 1.064 2.338 0.999 1.494 1.008 1.235 1.781 1.804 2.155

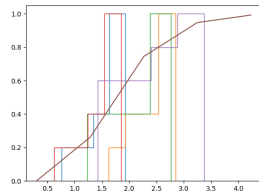
1.2 Эмпирическая функция

1.2.1 Практические результаты

Для генерации "реальной" функции плотности использовалась библиотечная ф-ция с табличной реализацией (ссылка ниже), т.к. PDF гамма-распределения не представим в элементарных ф-циях.

построение графиков ст.39-43, верхняя граница ст. 17-22

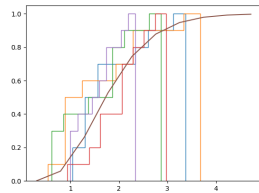
$n = 5$



Верхние границы разниц по всем эмпирическим функциям:

0.2 0.199 0.0 0.399 0.199 0.2 0.399 0.199 0.399 0.399

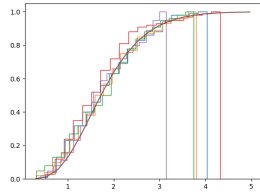
$n = 10$



Верхние границы разниц по всем эмпирическим функциям:

0.1 0.299 0.499 0.3 0.199 0.4 0.2 0.3 0.1 0.299

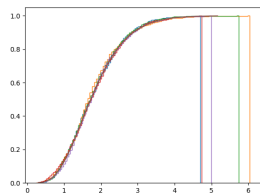
$n = 100$



Верхние границы разниц по всем эмпирическим функциям:

0.08 0.13 0.19 0.17 0.07 0.27 0.13 0.3 0.13 0.33

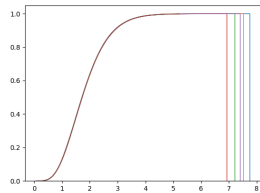
$n = 1000$



Верхние границы разниц по всем эмпирическим функциям:

0.222 0.178 0.043 0.092 0.055 0.234 0.143 0.19 0.096 0.102

$n = 100000$



Верхние границы разниц по всем эмпирическим функциям:

0.0023 0.005 0.012 0.0049 0.0028 0.0092 0.0027 0.0072 0.0091 0.0075

1.2.2 Теоритический вывод

Можно заметить, что с ростом размера выборки график эмпирический функции совпадает с графиком функции распределения, что подтверждает теорему о сходимости эмпирической функции распределения:

$\forall \epsilon > 0, n \rightarrow \infty$

$P(|\hat{F}_n(x) - F(x)| < \epsilon) \rightarrow 1$

$\hat{F}_n(x) \xrightarrow{P} F(x)$

Т.е. показано свойство состоятельности эмпирических функций

1.3 Вариационный ряд

Вариационный ряд ст. 33

$n = 5$

- 1) 0.759 1.382 1.778 1.826 1.926
- 2) 1.625 2.107 2.543 2.623 2.846
- 3) 1.228 1.326 2.437 2.749 2.761
- 4) 0.626 1.425 1.626 1.847 1.851
- 5) 1.421 1.472 1.766 2.816 3.370

$n = 10$

- 1) 1.042 1.234 1.375 1.441 1.491 1.615 1.785 2.553 2.741 3.378
- 2) 0.537 1.107 1.127 1.144 1.221 1.545 1.976 2.401 2.432 3.679
- 3) 0.610 0.769 0.803 0.986 1.476 1.952 2.081 2.082 2.173 2.867
- 4) 0.934 1.592 1.618 1.810 2.096 2.138 2.280 2.416 2.670 2.970
- 5) 0.999 1.008 1.064 1.235 1.494 1.736 1.781 1.804 2.155 2.338

1.4 Квантили

1.4.1 Практические результаты

Столбики - это 0.1, 0.5 и 0.7 квантили соответственно

квантили ст.35

$n = 5$

- 1) 11.071, 1.802, 1.876
- 2) 1.866, 2.583, 2.735
- 3) 1.277, 2.593, 2.755
- 4) 1.026, 1.736, 1.849
- 5) 1.447, 2.291, 3.093

$n = 10$

- 1) 1.304, 1.700, 2.647
- 2) 1.117, 1.761, 2.416
- 3) 0.786, 2.016, 2.128
- 4) 1.605, 2.209, 2.543
- 5) 1.036, 1.758, 1.980

$n = 100$

- 1) 0.919, 1.776, 2.351
- 2) 0.874, 1.787, 2.259
- 3) 0.822, 1.953, 2.305
- 4) 0.902, 1.697, 2.097
- 5) 0.992, 1.693, 2.273

$n = 1000$

- 1) 0.923, 1.741, 2.205
- 2) 0.919, 1.688, 2.113
- 3) 0.932, 1.741, 2.152
- 4) 0.890, 1.764, 2.177
- 5) 0.997, 1.760, 2.146

$n = 100000$

- 1) 0.927, 1.718, 2.149
- 2) 0.931, 1.725, 2.152
- 3) 0.927, 1.719, 2.147
- 4) 0.927, 1.725, 2.153
- 5) 0.928, 1.725, 2.144

Реальные значения квантилей:

0.1 квантиль: 0.929
 0.5 квантиль: 1.723
 0.7 квантиль: 2.149

1.4.2 Теоритический вывод

Как и ожидалось, с увеличением объема выборки разни́ца между выборочными и теоритическими квантилями уменьшается с ростом размера выборки.

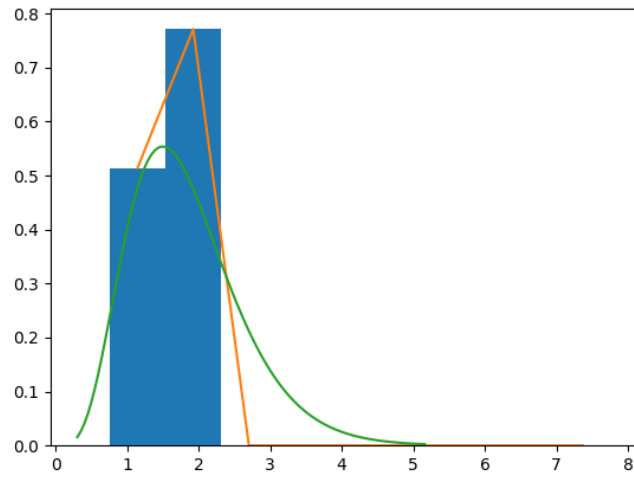
Можно заметить, что при увеличении размера моделируемой выборки квантили почти совпадают

1.5 Гистограмма и полигон частот

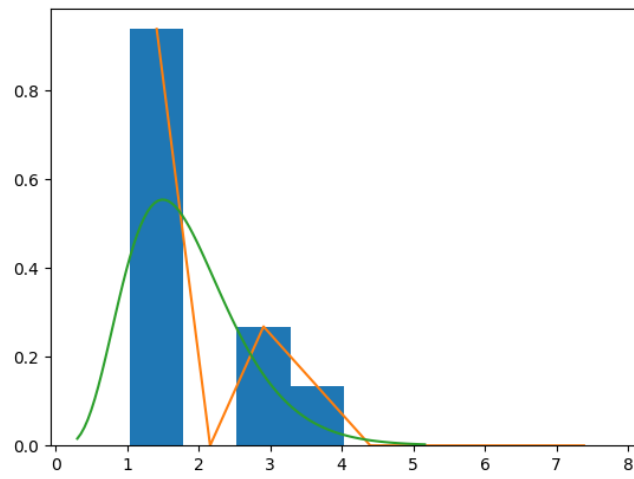
1.5.1 Практические- результаты

Зеленая линия - теоретическая плотность вероятности, оранжевая - полигон частот и синим цветом - гистограмма частот. Как видно, все три линии графика практически совпадают при большом объеме выборки.

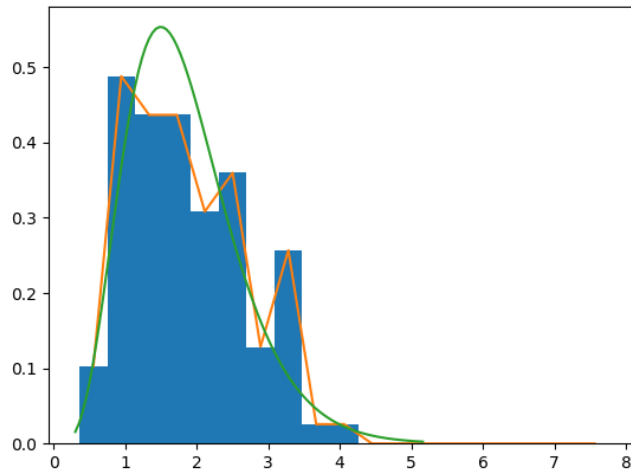
реализация ст. 72-78



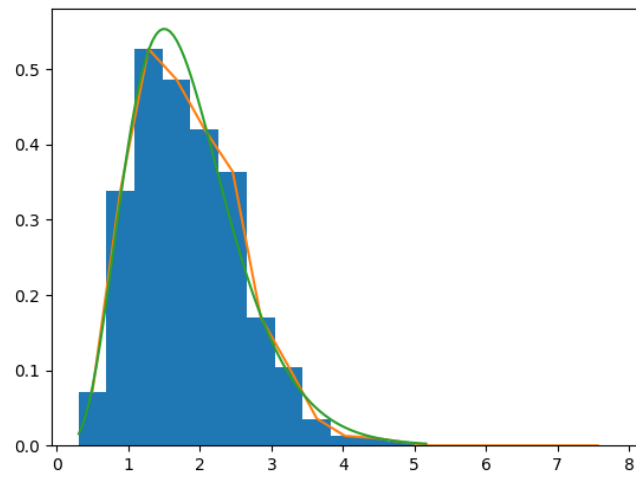
$n = 5$



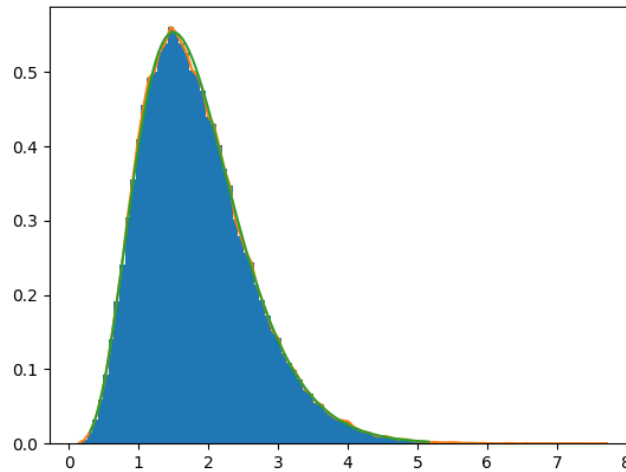
$n = 10$



$n = 100$



$n = 1000$



$n = 100000$

1.5.2 Теоритический вывод

Гистограмма является кусочно постоянной функцией, построенной по правилу

$$\hat{f}_n(x) = \frac{v_r}{n|\Delta_r|}, x \in \Delta_r$$

$v_r = \sum_{i=1}^n \text{Ind}(X_i \in \Delta_r)$, т.е. количество элементов выборки попавших в каждый интервал. Результаты построений подтверждают сходимость гистограммы к теоритической плотности с увеличением размера выборки:

Из закона больших числе относительная частота $\frac{v_r}{n}$ сближается с теоритической вероятностью как:

$$P\xi \in \Delta_r = \int_{\Delta_r} f(x)dx$$

что по теореме о среднем равно $f(a_r)|\Delta_r|$

где a_r - некоторая точка интервала r (например середина интервала). Таким образом, при больших n и достаточно мелком разбиении $\hat{f}_n(x) = f(a_r)$ т.е. гистограмма должна достаточно хорошо приближать график плотности, что можно увидеть на картинках.

2 распределение Бореля-Таннера

2.1 Моделирование выборок

Выборки были смоделированы для параметра скорости обслуживания $\alpha = 0.4$ и начального кол-ва клиентов $r = 4$.

ф-ция моделирования ст. 12-20

$n = 5$

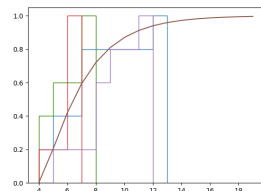
- 1) 4 5 13 7 7
- 2) 8 8 5 4 4
- 3) 4 8 5 7 4
- 4) 6 7 4 6 7
- 5) 5 12 9 8 8

$n = 10$

- 1) 7 5 4 9 11 4 8 7 4 8
- 2) 4 5 6 6 6 7 10 9 6 8
- 3) 6 5 5 14 5 8 9 12 5 4
- 4) 8 6 9 6 8 16 6 5 4 4
- 5) 6 9 4 6 6 4 7 7 4 5

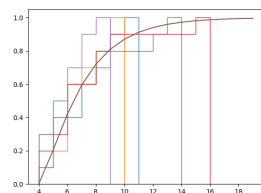
2.2 Эмпирическая функция

$n = 5$



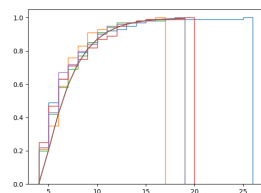
Верхние границы разниц по всем эмпирическим функциям:
0.4 0.2 0.3 0.8 0.2 0.4 0.4 0.4 0.6 0.8

$n = 10$



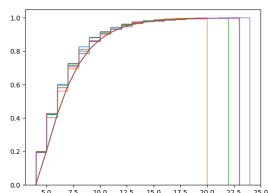
Верхние границы разниц по всем эмпирическим функциям:
0.2 0.2 0.26 0.3 0.3 0.13 0.2 0.2 0.3 0.23

$n = 100$

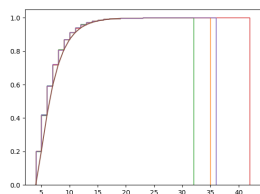


Верхние границы разниц по всем эмпирическим функциям:

0.145 0.075 0.05 0.065 0.07 0.12 0.08 0.05 0.09 0.06
 $n = 1000$



Верхние границы разниц по всем эмпирическим функциям:
0.041 0.013 0.039 0.024 0.034 0.026 0.021 0.026 0.015 0.026
 $n = 100000$



Верхние границы разниц по всем эмпирическим функциям:
0.0017 0.0045 0.0023 0.0017 0.0034 0.0014 0.0015 0.0033 0.0033 0.0013

2.3 Вариационный ряд

$n = 5$

- 1) 4 5 7 7 13
- 2) 4 4 5 8 8
- 3) 4 4 5 7 8
- 4) 4 6 6 7 7
- 5) 5 8 8 9 12

$n = 10$

- 1) 4 4 4 5 7 7 8 8 9 11
- 2) 4 5 6 6 6 6 7 8 9 10
- 3) 4 5 5 5 5 6 8 9 12 14
- 4) 4 4 5 6 6 6 8 8 9 16
- 5) 4 4 4 5 6 6 6 7 7 9

2.4 Квантили

В столбиках 0.1, 0.5 и 0.7 квантили соответственно

$n = 5$

- 1) 4.5, 7.0, 10.0
- 2) 4.0, 6.5, 8.0

- 3) 4.0, 6.0, 7.5
- 4) 5.0, 6.5, 7.0
- 5) 6.5, 8.5, 10.5

$n = 10$

- 1) 4.0, 7.5, 8.5
- 2) 5.5, 6.5, 8.5
- 3) 5.0, 7.0, 10.5
- 4) 4.5, 7.0, 8.5
- 5) 4.0, 6.0, 7.0

$n = 100$

- 1) 4.0, 6.0, 7.5
- 2) 4.0, 6.0, 7.0
- 3) 4.0, 6.0, 8.0
- 4) 4.0, 6.0, 7.5
- 5) 4.0, 6.0, 7.0

$n = 1000$

- 1) 4.0, 6.0, 7.0
- 2) 4.0, 6.0, 8.0
- 3) 4.0, 6.0, 7.0
- 4) 4.0, 6.0, 7.0
- 5) 4.0, 6.0, 7.0

$n = 100000$

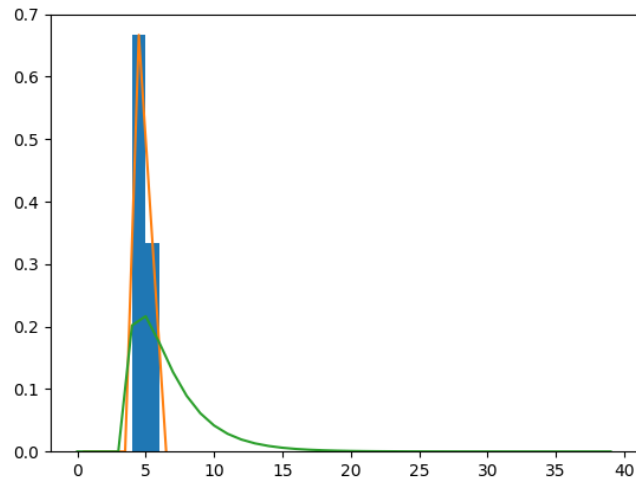
- 1) 4.0, 6.0, 7.0
- 2) 4.0, 6.0, 7.0
- 3) 4.0, 6.0, 7.0
- 4) 4.0, 6.0, 7.0
- 5) 4.0, 6.0, 7.0

Теоретически посчитанные квантили для тех же параметров:

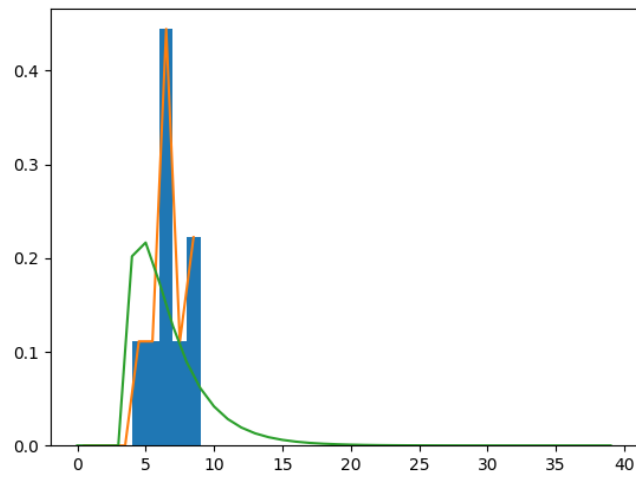
4 6 7

2.5 Гистограмма и полигон частот

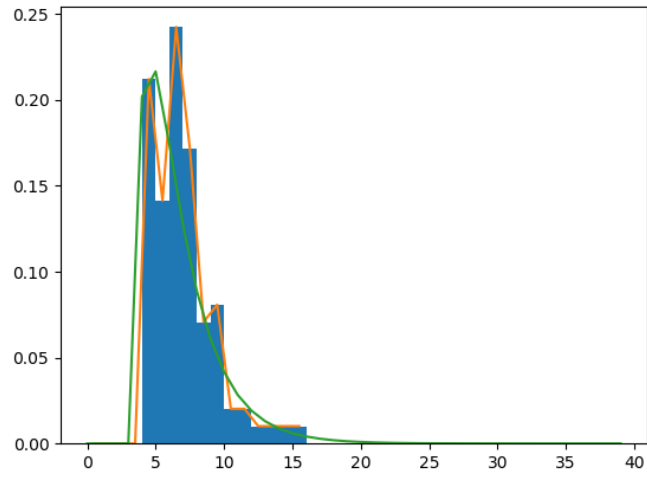
Зеленая линия - теоретическая плотность вероятности, оранжевая - полигон частот и синим цветом - гистограмма частот. Как видно, все три линии графика практически совпадают при большом объеме выборки.



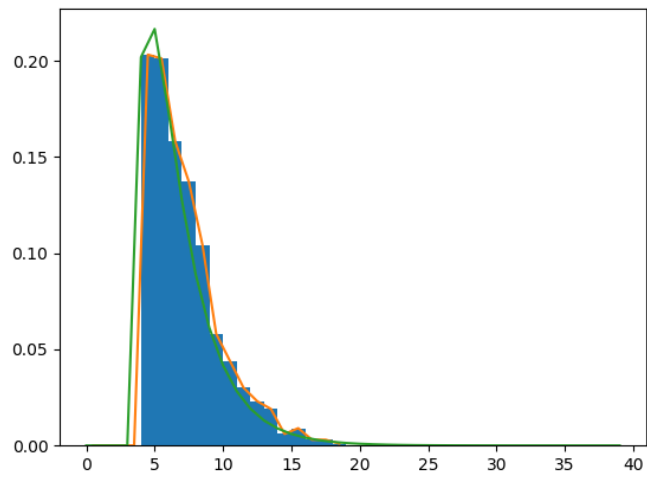
$n = 5$



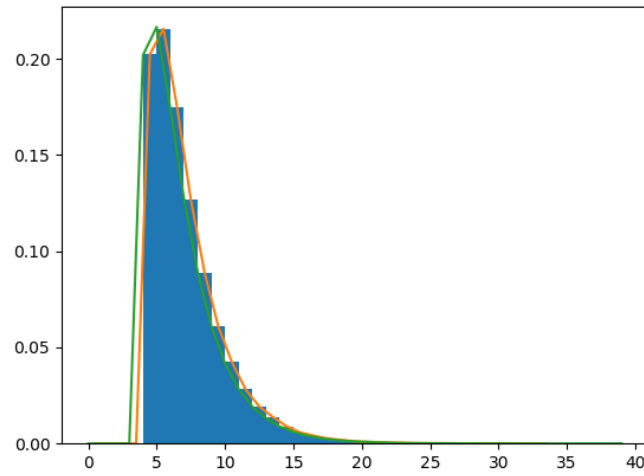
$n = 10$



$n = 100$



$n = 1000$



$n = 100000$

Для распределения бореля-таннера использовались аналогичные ф-ции, кроме генерации эмпирической. Для нее была написана следующая ф-ция:
[ст. 22-24](#)