

Национальный исследовательский университет
Высшая школа экономики
Московский институт электроники и математики

Департамент прикладной математики
кафедра компьютерной безопасности

Долгосрочное домашнее задание по математической
статистике

Дискретное распределение: *Биномиальное*

$$P(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, x \in \{0, 1, \dots, n\}, n \in \mathbb{N}, 0 < \theta < 1$$

Непрерывное распределение: *Эрланга*

$$f(x) = \frac{\theta^m}{(m-1)!} x^{m-1} e^{-\theta x}, x, \theta \in \mathbb{R}^+, m \in \mathbb{N}$$

Выполнил
Яськов А.С.

Проверил
Чухно А.Б.

Москва 2022

Оглавление

1	Характеристики вероятностных распределений	4
1.1	Описание основных характеристик распределения	4
1.1.1	Дискретное распределение	4
1.1.2	Непрерывное распределение	5
1.2	Поиск примеров событий	7
1.2.1	Дискретное распределение	7
1.2.1.1	Пример интерпретации распределения	7
1.2.1.2	Соотношения между распределениями	7
1.2.2	Непрерывное распределение	8
1.2.2.1	Пример интерпретации распределения	8
1.2.2.2	Соотношения между распределениями	8
1.3	Описание способа моделирования	8
1.3.1	Дискретное распределение	8
1.3.2	Непрерывное распределение	10
2	Основные понятия математической статистики	14
2.1	Генерация выборок выбранных случайных величин	14
2.1.1	Дискретное распределение	14
2.1.2	Непрерывное распределение	20
2.2	Построение эмпирической функции распределения	28
2.2.1	Дискретное распределение	29
2.2.2	Непрерывное распределение	33
2.2.3	Дискретное распределение	38
2.2.4	Непрерывное распределение	39
2.3	Построение гистограммы и полигона частот	39
2.3.1	Дискретное распределение	39
2.3.2	Непрерывное распределение	44
2.4	Вычисление выборочных моментов	48
2.4.1	Дискретное распределение	50
2.4.2	Непрерывное распределение	51
3	Основные понятия математической статистики	53
3.1	Получение оценок	53
3.1.1	Дискретное распределение	53
3.1.1.1	Метод моментов	53
3.1.1.2	Метод максимального правдоподобия	54

3.1.2	Непрерывное распределение	56
3.1.2.1	Метод моментов	56
3.1.2.2	Метод максимального правдоподобия	57
3.2	Поиск оптимальных оценок	58
3.2.1	Дискретное распределение	58
3.2.2	Непрерывное распределение	60
4	Проверка статистических гипотез	63
4.1	Проверка гипотезы о виде распределения	63
4.1.1	Критерий согласия Колмогорова (Смирнова)	63
4.1.1.1	Дискретное распределение	64
4.1.2	Непрерывное распределение	64
4.1.3	Критерий согласия хи-квадрат	64
4.1.3.1	Дискретное распределение	65
4.1.3.2	Непрерывное распределение	65
4.1.4	Критерий согласия Колмогорова (Смирнова) для слож- ной гипотезы (в условиях когда неизвестен параметр рас- пределения)	66
4.1.4.1	Дискретное распределение	66
4.1.4.2	Непрерывное распределение	67
4.1.5	Критерий согласия хи-квадрат для сложной гипотезы (в условиях когда неизвестен параметр распределения)	67
4.1.5.1	Дискретное распределение	68
4.1.5.2	Непрерывное распределение	68
4.2	Проверка гипотезы об однородности выборок	68
4.2.1	Дискретное распределение	68
4.2.1.1	Непрерывное распределение	69
5	Различение статистических гипотез	71
5.1	Теоретическое введение	71
5.2	Вычисление функции отношения правдоподобия	72
5.2.1	Дискретное распределение	72
5.2.2	Непрерывное распределение	73
5.3	Вычисление критической области	74
5.3.1	Дискретное распределение	74
5.3.2	Непрерывное распределение	76
5.4	Вычисление минимального необходимого количества...	77
5.4.1	Дискретное распределение	77
5.4.2	Непрерывное распределение	79

Домашнее задание 1.

Характеристики вероятностных распределений

1. Описание основных характеристик распределения

1.1.1. Дискретное распределение

Функция распределения

$$F_{\xi}(x) = P(\xi < x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} \theta^k (1 - \theta)^{n-k}, x \in \{0, 1, \dots, n\}, n \in \mathbb{N}, 0 < \theta < 1$$

Математическое ожидание

Математическое ожидание дискретной случайной величины вычисляется по формуле:

$$E\xi = \sum_{i=1}^{\infty} x_i \cdot P(\xi = x_i)$$

Пусть ξ имеет биномиальное распределение с параметрами n и θ . Это соответствует числу успехов в n независимых испытаниях Бернулли с вероятностью θ в каждом. Получаем, что ξ - сумма n независимых испытаний x_1, x_2, \dots, x_n , где $x_i = 1$ при успехе. Тогда, при $n = 1$ $E\xi = \theta$

$$E\xi = E \sum_{i=1}^n x_i = \sum_{i=1}^n E x_i = n\theta$$

Дисперсия

Дисперсия вычисляется по формуле:

$$D\xi = E(\xi - E\xi)^2 = E\xi^2 - (E\xi)^2$$

Тогда, при $n = 1$ $E\xi = \theta$ и $E\xi^2 = \theta$ (так как при возведении в квадрат получается та же самая случайная величина). Получаем $D\xi = \theta - \theta^2 = \theta(1 - \theta)$

$$D\xi = \sum_{i=1}^n D(x_i) = n\theta(1 - \theta)$$

Квантиль уровня γ

Квантилью распределения случайной величины ξ уровня γ называется любое

из чисел x_γ таких, что

$$P(\xi \leq x_\gamma) \geq \gamma \text{ \&\& } P(\xi \geq x_\gamma) \geq 1 - \gamma, \gamma \in (0, 1)$$

Найдем такой x , при котором выполняются данные условия:

$$P(\xi \leq x_\gamma) \geq \gamma \rightarrow \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} \theta^k (1 - \theta)^{n-k} \geq \gamma$$

$$P(\xi \geq x_\gamma) \geq 1 - \gamma \rightarrow \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} \theta^k (1 - \theta)^{n-k} \leq \gamma$$

Получаем x из решения данного уравнения $\sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} \theta^k (1 - \theta)^{n-k} = \gamma$

1.1.2. Непрерывное распределение

Функция распределения

$$\begin{aligned} F_\xi(x) &= P(\xi < x) = \int_0^x \frac{\theta^m}{(m-1)!} x^{m-1} e^{-\theta x} dx = \int_0^x \frac{\theta \theta^{m-1}}{(m-1)!} x^{m-1} e^{-\theta x} dx = \\ &= \frac{\theta}{(m-1)!} \int_0^x (\theta x)^{m-1} x^{-\theta x} dx = / * \text{ пусть } \theta x = t, dx = \frac{dx}{\theta} * / = \frac{1}{(m-1)!} \cdot \\ &\quad \cdot \int_0^{\theta x} t^{m-1} e^{-t} dt = \frac{\Gamma(m, \theta x)}{(m-1)!} \end{aligned}$$

Математическое ожидание

В случае непрерывной случайной величины математическое ожидание имеет следующее определение:

Если распределение случайной величины ξ имеет плотность $f_\xi(x)$ и $\int_{\mathbb{R}} f_\xi(x) <$

∞ , то $E\xi = \int_{\mathbb{R}} x f_\xi(x) dx$

Получаем математическое ожидание распределения Эрланга:

$$E\xi = \int_{\mathbb{R}} x \frac{\theta^m}{(m-1)!} x^{m-1} e^{-\theta x} dx = / * \text{ пусть } \theta x = t, dx = \frac{dt}{\theta} * / = \frac{1}{\theta \Gamma(m)} \cdot \int_0^{\infty} e^{-t} t^m dt = \frac{\Gamma(m+1)}{\theta \Gamma(m)} = \frac{m}{\theta}$$

Дисперсия

Дисперсия вычисляется по формуле:

$$D\xi = E(\xi - E\xi)^2 = E\xi^2 - (E\xi)^2$$

Найдем $E\xi^2$:

$$E\xi^2 = \int_{\mathbb{R}} x^2 \frac{\theta^m}{(m-1)!} x^{m-1} e^{-\theta x} dx = / * \text{ пусть } \theta x = t, dx = \frac{dt}{\theta} * / = \frac{1}{\theta^2 \Gamma(m)} \cdot \int_0^{\infty} \theta^2 \Gamma(m) e^{-t} t^{m+1} dt = \frac{\Gamma(m+2)}{\theta^2 \Gamma(m)} = \frac{m(m+1)}{\theta^2}$$

Таким образом, дисперсия будет равна $D\xi = \frac{m(m+1)}{\theta^2} - \frac{m^2}{\theta^2} = \frac{m}{\theta^2}$

Квантиль уровня γ

Квантилью распределения случайной величины ξ уровня γ называется любое из чисел x_γ таких, что

$$P(\xi \leq x_\gamma) \geq \gamma \text{ \&\& } P(\xi \geq x_\gamma) \geq 1 - \gamma, \gamma \in (0, 1)$$

Найдем такой x , при котором выполняются данные условия:

$$P(\xi \leq x_\gamma) \geq \gamma \rightarrow \frac{\Gamma(m, \theta x)}{(m-1)!} \geq \gamma$$

$$P(\xi \geq x_\gamma) \geq 1 - \gamma \rightarrow \frac{\Gamma(m, \theta x)}{(m-1)!} \leq \gamma$$

Получаем x из решения данного уравнения $\frac{\Gamma(m, \theta x)}{(m-1)!} = \gamma$

2. Поиск примеров событий

1.2.1. Дискретное распределение

1.2.1.1. Пример интерпретации распределения

- ▷ Бросание кубика: Вероятность получения числа шесть (6) (0, 1, 2, 3...50) при броске кубика 50 раз; Здесь случайная величина X - это количество "успехов то есть количество раз, когда происходит шесть. Вероятность получить шестерку равна $1/6$. Биномиальное распределение может быть представлено в виде $\text{Bin}(50, 1/6)$.
- ▷ Бросание монеты: Вероятность получения количества голов (0, 1, 2, 3...50) при подбрасывании монеты 50 раз; Здесь случайная величина X - это количество "успехов то есть количество выпадений орла. Вероятность получить орел равна $1/2$. Биномиальное распределение можно представить в виде $\text{Bin}(50, 0,5)$.
- ▷ Дефектные изделия: Вероятность обнаружения количества дефектных изделий (0, 1, 2, 3...30) при проверке 30 раз; Здесь случайная величина X - это количество "успехов то есть количество раз, когда обнаруживается дефектный товар. Вероятность обнаружения дефектного элемента равна p . Биномиальное распределение может быть представлено в виде $\text{Bin}(30, p)$
- ▷ Человек, страдающий какой-либо болезнью: Вероятность обнаружения 0 или более человек, страдающих определенной болезнью, при обследовании 100 человек; Здесь случайная величина X - это число "успехов то есть число людей, страдающих какой-либо болезнью. Вероятность нахождения человека, страдающего какой-либо болезнью, говорит, p . Биномиальное распределение может быть представлено в виде $\text{Bin}(100, p)$

1.2.1.2. Соотношения между распределениями

- ▷ Если $n=1$, то получаем распределение Бернулли.
- ▷ Если n большое, то в силу центральной предельной теоремы (сумма достаточно большого количества слабо зависимых случайных величин, имеющих примерно одинаковые масштабы) $\text{Bin}(n, p) \xrightarrow[n \rightarrow \infty]{P} N(np, npq)$, где $N(np, npq)$ - нормальное распределение с математическим ожиданием np и дисперсией npq .
- ▷ Если n большое, а λ — фиксированное число, то $\text{Bin}(n, \lambda/n) \xrightarrow[n \rightarrow \infty]{P} P(\lambda)$, где $P(\lambda)$ — распределение Пуассона с параметром λ .

- ▷ Если случайные величины X и Y имеют биномиальные распределения $\text{Bin}(D, p)$ и $\text{Bin}(N-D, p)$ соответственно, то условное распределение случайной величины X при условии $X+Y=n$ - Гипергеометрическое $\text{HG}(N, D, n)$.

1.2.2. Непрерывное распределение

1.2.2.1. Пример интерпретации распределения

- ▷ Распределение Эрланга используется для моделирования времени между входящими вызовами в колл-центр вместе с ожидаемым количеством вызовов. Это позволяет колл-центрам знать, какой должна быть их кадровая мощность в разное время дня, чтобы они могли своевременно обрабатывать входящие звонки, не теряя денег из-за набора слишком большого количества сотрудников в течение данной смены.
- ▷ Распределение Эрланга широко используется для моделирования распределения времени клеточного цикла, которое имеет множество различных применений в медицинских учреждениях.
- ▷ Распределение Эрланга используется розничными торговцами для моделирования частоты повторных покупок потребителями. Это дает розничным торговцам и другим предприятиям представление о том, как часто данный потребитель будет покупать у них продукт или услугу. Это помогает предприятиям с контролем запасов, а также с персоналом.

1.2.2.2. Соотношения между распределениями

- ▷ При $m = 1$, распределение Эрланга равно экспоненциальному распределению.
- ▷ При $\theta = 2$, распределение Эрланга равно распределению χ -квадрат с 2 степенями свободы.

3. Описание способа моделирования

1.3.1. Дискретное распределение

Биномиальное распределение $\text{Bin}(n, p)$ с параметрами $n \in \mathbb{N}$ и $p \in (0, 1)$ можно задать с помощью таблицы распределения:

$$P : \begin{pmatrix} 0 & 1 & \dots & k & \dots & n \\ p_0 & p_1 & \dots & p_k & \dots & p_n \end{pmatrix}, \text{ где } p_i = \binom{n}{i} p^i (1-p)^{n-i}$$

Алгоритм для моделирования BIS (Binomial Inverse Sequential) является просто последовательным методом обратных функций, но в нетабличном варианте. А именно, накопленные вероятности s_k здесь вычисляются не заранее,

а в том же цикле, где проверяются неравенства $\alpha \leq s_k$. Точнее, вероятности p_k рекуррентно пересчитываются одна через другую, а накопленные вероятности s_k последовательно вычисляются через s_{k-1} и p_k .

Поскольку для биномиального распределения $p_0 = (1 - p)^n$ и

$$\frac{p_k}{p_{k-1}} = \frac{n - k + 1}{k} \frac{p}{1 - p}, \text{ при } k = 1, 2, \dots, n, \text{ то}$$

мы приходим к следующему алгоритму:

Входные данные: n, p .

Результат: ξ

1. (Инициализация) $c \leftarrow p/(1 - p)$; $s \leftarrow r \leftarrow (1 - p) \wedge n$; $k \leftarrow 0$; $\text{Get}(\alpha)$
2. (Пересчет вероятностей и поиск окна) While $\alpha > s$ do $k \leftarrow k + 1$; $r \leftarrow r \cdot c \cdot (n - k + 1)/k$; $s \leftarrow s + r$;
3. (Завершение) $\xi \leftarrow k$; STOP

Дадим краткие комментарии к алгоритму. Переменная s обозначает текущую накопленную вероятность s_k , а переменная r — текущую вероятность p_k . Поскольку первая накопленная вероятность равна p_0 , то инициализация переменных s и r одинакова. Переменная c заведена для того, чтобы не делить p на $1 - p$ в цикле.

Source code:

Данная программа моделирует 10000 случайных величин с Биноминальным распределением с параметрами $n = 35$ $p = 0.5$ и строит график плотности вероятности.

Листинг 1.1: Binomial distribution

```

1  import random
2  import matplotlib.pyplot as plt
3  import seaborn as sbn
4
5  def Bin(n, p):
6
7      c = p / (1 - p)
8      r = (1 - p) ** n
9      s = r
10     k = 0
11     x = random.uniform(0, 1)
12
13     while x > s:
14         k += 1

```

```

15         r *= c * (n - k + 1) / k
16         s += r
17
18     return k
19
20
21     n = 35
22     theta = 0.5
23     N = 10000
24     array = []
25
26     for i in range(N):
27         number = Bin(n, theta)
28         array.append(number)
29
30     sbn.kdeplot(array)
31     plt.grid()
32     plt.show()

```

Plot:

1.3.2. Непрерывное распределение

Плотность распределения Эрланга с параметрами m и θ задается формулой $f_{m,\theta}(x) = \frac{\theta^m}{(m-1)!} x^{m-1} e^{-\theta x}$. θ - параметр масштаба: если $\chi \in E(m, 1)$, то $\xi = \frac{\chi}{\theta} \in E(m, \theta)$. Достаточно уметь моделировать распределение Эрланга с $\theta = 1$, при $m \geq 1$ такое распределение может быть легко промоделировано с помощью формулы:

$$\xi = -\ln(\alpha_1, \alpha_2, \dots, \alpha_m), \alpha_i = rand(0, 1)$$

При $\theta \neq 1$ получаем $\xi = \frac{-\ln(\alpha_1, \alpha_2, \dots, \alpha_m)}{\theta}, \alpha_i = rand(0, 1)$

Source code:

Данная программа моделирует 10000 случайных величин с распределением

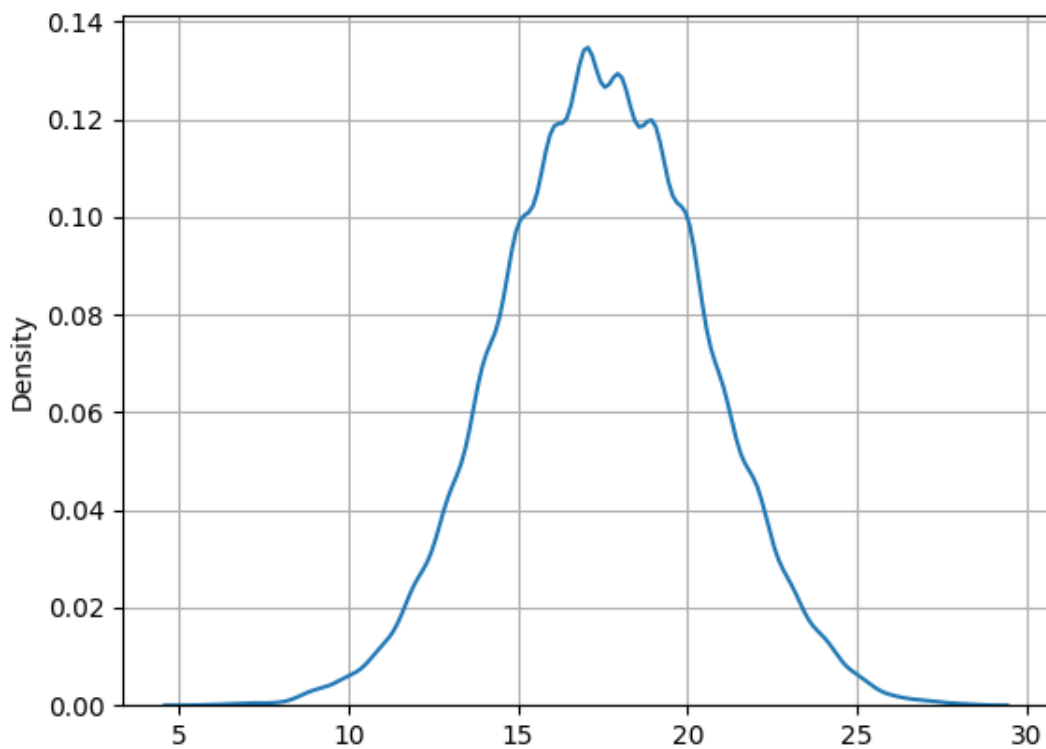


Рис. 1.1: Функция плотности вероятности

Эрланга с параметрами $m = 8$ и $\theta = \frac{1}{8}$ и строит график плотности вероятности.

Листинг 1.2: Erlang distribution

```

1  import random
2  import math
3  import matplotlib.pyplot as plt
4  import seaborn as sbn
5
6  def erlang(a,b):
7      k = 1
8      for i in range(a):
9          k *= random.uniform(0,1)
10
11     return -math.log(k)/b
12
13
14  N = 10000
15  m = 8
16  theta = 0.125

```

```
17     array = []
18
19     for i in range(N):
20         n = erlang(m, theta)
21         array.append(n)
22
23     sbn.kdeplot(array)
24     plt.grid()
25     plt.show()
```

Plot:

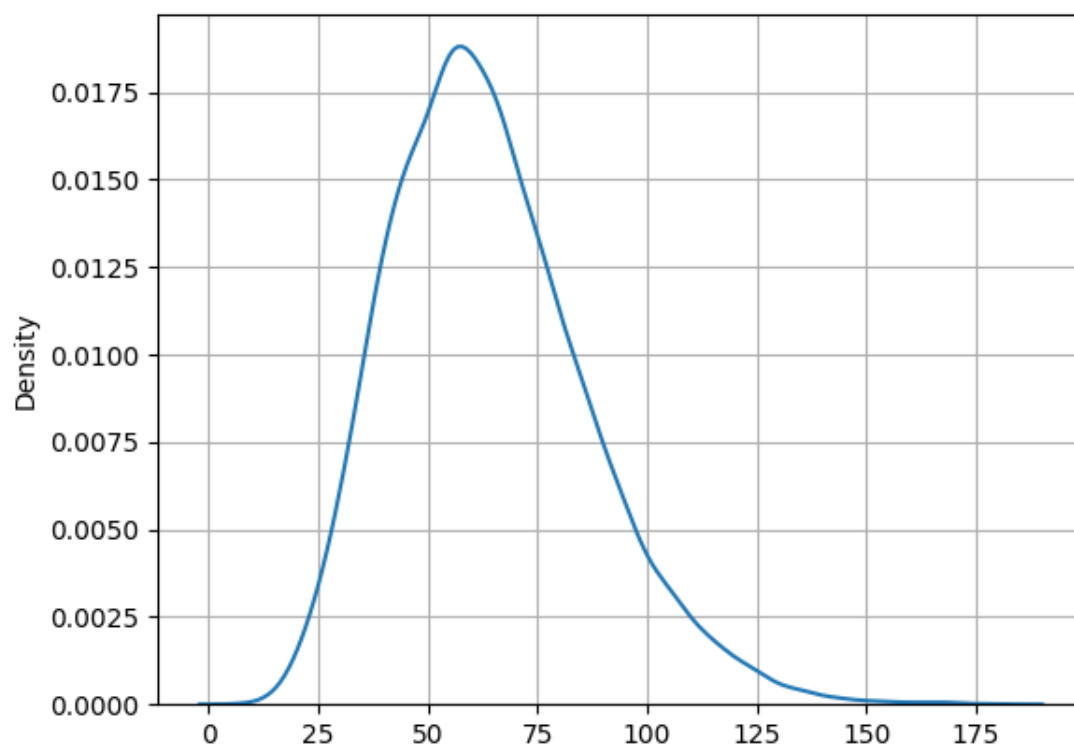


Рис. 1.2: Функция плотности вероятности

Домашнее задание 2.

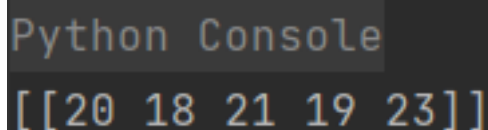
Основные понятия математической статистики

1. Генерация выборок выбранных случайных величин

2.1.1. Дискретное распределение

Генерация выборки биномиального распределение с параметрами $n = 35$ и $\theta = 0.5$

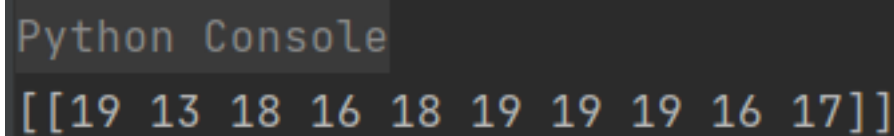
▷ $N = 5$:



```
Python Console  
[[20 18 21 19 23]]
```

Рис. 2.1: Выборка из Биномиального распределения $N = 5$

▷ $N = 10$:



```
Python Console  
[[19 13 18 16 18 19 19 19 16 17]]
```

Рис. 2.2: Выборка из Биномиального распределения $N = 10$

▷ $N = 100$:

```
Python Console
[[16 15 16 17 16 21 15 20 17 15]
 [17 15 18 18 18 17 17 14 19 13]
 [16 17 21 21 14 15 16 22 20 19]
 [15 19 18 18 18 24 18 21 20 20]
 [15 19 14 11 23 16 19 11 17 16]
 [16 16 15 16 16 20 15 17 16 23]
 [17 17 14 16 16 20 21 20 17 20]
 [18 16 15 15 18 18 19 22 15 16]
 [12 16 16 18 15 16 16 13 20 19]
 [16 14 15 17 16 17 19 21 17 16]]
```

Рис. 2.3: Выборка из Биномиального распределения $N = 100$

▷ $N = 200$:

```
Python Console
[[23 15 21 13 20 16 15 23 15 20 19 21 16 15 17 19 20 16 18 16]
 [22 18 19 14 21 21 17 16 22 23 15 22 25 20 17 23 13 22 18 15]
 [13 18 21 22 19 20 14 16 16 18 19 17 18 12 21 18 22 22 18 19]
 [19 18 14 24 21 19 19 20 17 20 15 16 15 16 19 19 18 14 19 17]
 [18 15 20 17 19 19 19 21 14 21 13 19 18 18 20 20 17 14 18 22]
 [18 14 18 20 13 21 12 18 23 19 19 12 13 13 16 8 21 13 15 12]
 [13 23 25 19 18 17 21 14 18 19 20 18 17 18 18 16 17 16 17 25]
 [13 15 15 19 12 17 16 19 17 18 18 19 18 14 12 19 20 16 15 18]
 [23 18 13 16 20 14 15 17 21 21 19 25 18 18 14 18 12 13 17 18]
 [17 17 14 21 17 16 18 14 19 16 22 20 16 19 19 15 17 19 21 20]]
```

Рис. 2.4: Выборка из Биномиального распределения $N = 200$

▷ $N = 400$:

```
Python Console
[[13 17 13 14 21 12 22 21 16 15 17 18 14 21 22 18 19 15 18 14]
 [17 15 24 20 15 17 16 16 21 21 20 19 18 13 18 20 17 16 15 17]
 [19 16 14 19 20 18 20 19 18 18 19 22 15 12 17 18 12 16 13 16]
 [18 11 16 17 13 19 19 16 16 14 17 23 18 17 14 21 22 16 17 14]
 [19 23 16 19 18 18 13 17 13 17 16 17 20 18 22 15 16 16 18 17]
 [20 14 21 14 13 14 16 17 15 17 18 13 18 20 15 16 14 14 16 15]
 [13 14 15 19 17 19 15 18 22 22 20 17 21 20 24 21 19 17 17 15]
 [17 20 18 19 13 23 16 17 21 18 19 20 20 19 13 16 20 19 15 17]
 [11 16 13 18 15 10 17 18 14 20 17 20 21 16 20 14 21 18 14 17]
 [18 12 21 24 16 14 17 17 18 13 19 20 22 22 16 20 15 11 18 20]
 [15 17 20 22 16 20 19 21 20 21 24 23 18 19 19 16 17 19 20 15]
 [14 22 15 19 21 17 14 17 17 16 12 17 18 16 19 15 18 21 15 16]
 [12 16 22 18 15 12 17 15 12 17 22 18 20 16 16 20 20 23 15 17]
 [15 20 15 19 14 16 21 10 18 16 19 21 22 14 22 13 19 10 16 24]
 [13 20 17 22 19 15 20 15 18 12 22 12 18 18 14 11 16 13 17 14]
 [22 19 18 15 19 20 16 23 19 15 16 18 15 20 10 15 16 16 12 17]
 [16 17 21 11 19 22 17 20 18 18 14 16 17 19 19 15 17 23 15 15]
 [20 21 22 16 17 21 15 14 20 21 17 15 17 14 15 12 13 24 20 15]
 [20 14 18 18 15 18 16 19 17 19 13 13 20 11 16 15 14 20 18 18]
 [17 16 13 18 20 20 16 21 19 18 15 18 20 16 20 17 16 19 19 22]]
```

Рис. 2.5: Выборка из Биноминального распределения $N = 400$

▷ $N = 600$:

```
[15 21 16 20 23 13 16 16 13 15 17 12 19 16 18 19 18 18 11 19]
[13 20 19 17 10 18 24 18 18 19 19 17 20 15 18 21 19 15 15 17]
[20 20 18 18 13 18 18 18 14 18 17 17 18 19 17 20 13 15 17 8]
[17 17 7 13 14 18 17 20 17 23 15 20 13 21 17 19 21 21 19 13]
[16 14 25 14 17 15 22 19 19 22 15 16 18 20 17 13 23 19 14 16]
[13 18 17 20 14 20 14 17 21 16 20 12 21 20 18 19 21 15 17 15]
[16 15 20 19 18 17 18 16 14 16 20 14 20 12 19 18 16 17 22 20]
[19 13 17 17 22 16 15 24 16 23 14 15 13 18 15 15 19 16 14 22]
[21 17 13 14 21 16 19 20 23 19 18 13 17 17 12 19 15 22 19 18]
[16 18 19 19 16 19 12 16 17 19 20 16 20 12 20 19 15 22 15 16]
[13 21 22 22 19 21 16 12 13 18 13 16 16 19 19 17 16 19 16 15]
[20 19 14 14 22 22 25 19 14 17 17 17 17 15 19 17 21 12 17 16]
[20 14 20 14 15 18 16 14 17 18 21 18 19 15 18 20 12 17 12 21]
[19 18 15 17 16 18 12 16 17 19 20 15 18 16 16 18 18 20 21 15]
[14 19 17 21 13 16 21 20 16 17 19 14 14 11 20 22 20 13 15 16]
[18 16 18 20 16 18 22 17 19 13 21 15 18 14 19 20 20 13 18 14]
[12 15 15 15 20 20 21 20 17 17 14 21 15 17 15 20 20 15 18 19]
[20 20 17 18 21 16 14 18 19 17 23 12 16 21 20 15 22 24 15 14]
[17 11 21 8 19 18 14 18 16 19 15 21 19 16 17 18 15 16 19 15]
[19 19 14 22 16 14 13 15 14 20 21 14 23 17 21 15 12 20 14 14]
[17 17 20 15 20 10 12 16 18 24 19 15 17 16 14 18 12 19 21 18]
[18 18 15 20 18 16 20 19 14 12 23 19 13 23 22 16 19 18 17 17]
[15 15 19 18 16 18 17 23 17 13 21 13 14 19 19 17 15 16 21 11]
[17 15 17 17 19 18 14 17 20 22 22 14 19 15 18 13 11 15 18 20]
[22 14 16 20 18 19 15 16 14 20 17 19 18 18 17 9 16 20 18 19]
[17 20 19 14 14 14 15 16 16 19 17 20 21 15 24 24 9 13 15 23]
[21 17 17 17 18 18 16 14 19 17 21 18 21 15 27 17 17 24 19 18]
[20 25 10 19 16 22 12 15 19 17 20 15 19 15 19 18 20 20 16 18]
[19 23 16 16 17 17 19 17 20 14 25 15 15 21 17 19 19 22 17 20]
[16 17 18 15 20 22 17 19 22 16 18 12 20 19 17 19 14 23 18 20]]
```

Рис. 2.6: Выборка из Биномиального распределения $N = 600$

▷ $N = 800$:

```
[[19 16 21 18 20 20 16 23 16 15 19 17 15 20 16 16 14 17 21 16]
 [16 22 18 20 16 17 18 15 19 17 15 19 15 12 23 21 14 19 19 12]
 [16 13 20 23 15 21 20 15 21 11 17 14 16 22 20 22 18 22 17 15]
 [17 20 16 17 18 16 15 19 21 18 17 16 21 21 24 16 16 13 12 18]
 [15 18 20 16 19 18 18 23 13 21 16 17 20 20 21 16 17 17 21 18]
 [16 13 15 21 23 20 13 20 18 15 22 20 15 20 21 20 14 20 15 21]
 [22 19 14 15 18 19 20 18 11 15 17 13 16 20 16 15 19 22 18 16]
 [19 23 18 19 15 22 22 15 21 16 18 17 18 12 17 15 20 19 18 14]
 [16 20 17 18 22 14 23 20 16 16 17 12 19 20 12 14 20 16 16 15]
 [18 17 15 18 17 18 18 16 19 21 19 18 21 16 18 18 15 23 16 20]
 [17 20 13 15 22 16 16 16 19 14 19 12 18 15 15 22 17 18 20 25]
 [21 18 20 19 16 15 19 18 20 17 18 21 21 17 21 17 20 13 22 18]
 [16 20 16 16 22 16 17 14 16 14 15 13 19 15 16 21 17 19 16 18]
 [16 20 14 10 17 16 23 21 19 19 14 19 19 17 17 20 17 17 21 19]
 [19 17 15 17 17 20 17 19 14 18 21 19 17 16 24 20 15 20 21 18]
 [16 18 12 16 20 16 17 14 15 10 16 21 15 18 17 13 16 15 20 18]
 [19 22 16 16 20 17 20 16 22 18 16 20 18 19 17 13 21 15 11 13]
 [21 25 21 18 18 15 16 25 19 17 16 21 20 13 21 22 16 17 13 17]
 [19 17 18 13 20 13 16 19 16 16 13 22 16 22 15 17 21 14 20 17]
 [20 15 20 21 12 18 17 15 18 17 17 17 14 18 18 19 16 20 18 15]
 [16 19 15 14 19 17 19 16 15 22 16 15 13 16 19 20 16 22 15 14]
 [18 22 14 16 20 19 20 19 16 21 18 24 19 21 21 18 18 14 18 12]
 [18 16 17 17 13 17 18 18 18 17 18 16 16 17 17 15 20 14 16 15]
 [16 15 17 17 20 14 24 14 19 15 17 23 19 16 15 19 21 14 15 8]
 [18 21 20 18 15 22 19 20 15 16 20 10 23 16 18 16 16 18 17 22]
 [18 22 17 15 19 16 15 13 20 22 16 20 22 16 14 17 18 16 17 16]
 [20 25 16 14 20 24 12 19 15 15 15 21 14 17 20 17 18 12 19 18]
 [14 20 18 19 15 18 20 20 19 15 17 17 14 22 14 16 14 21 18 18]
 [14 17 12 20 20 15 19 15 17 19 18 16 15 20 18 15 15 21 20 12]
 [21 15 22 15 17 18 20 21 17 13 18 15 19 14 19 19 11 15 21 20]
 [14 23 16 19 17 19 14 16 21 13 21 21 17 19 20 18 9 25 14 13]
 [20 17 21 15 20 15 14 21 17 19 17 20 14 15 21 12 14 19 21 11]
 [16 18 18 17 15 22 12 20 21 15 17 19 19 17 18 11 13 22 18 20]
 [10 19 13 17 19 22 20 17 20 17 18 21 17 18 14 18 21 18 11 18]
 [19 21 12 20 16 15 13 19 15 19 17 19 20 16 17 20 24 17 17 22]
 [12 19 15 20 22 15 16 19 12 20 18 16 19 17 17 20 20 15 21 17]
 [19 22 19 18 11 17 18 20 18 14 15 16 18 19 14 13 21 20 19 19]
 [22 19 17 24 22 15 17 18 18 21 18 17 19 18 17 16 17 20 20 15]
 [17 21 17 15 14 18 19 18 15 16 19 18 19 20 19 22 10 19 19 20]
 [14 20 19 12 11 21 20 20 15 26 20 19 17 14 23 20 15 16 16 21]]
```

Рис. 2.7: Выборка из Биноминального распределения $N = 800$

▷ $N = 1000$:

```
[[25 14 16 13 21 17 14 17 17 22 22 21 22 21 22 17 19 12 18 16]
[18 10 19 19 17 19 15 19 19 15 18 19 23 21 16 22 16 22 19 20]
[15 20 17 18 12 15 17 16 14 18 15 18 21 18 20 17 19 12 17 16]
[20 19 17 13 17 19 25 20 12 14 20 17 19 22 19 17 18 16 19 15]
[23 14 15 15 15 22 17 17 19 18 17 16 17 16 17 15 23 17 13 15]
[13 23 17 17 26 19 17 15 20 18 18 12 19 19 19 18 20 20 15 16]
[16 20 17 18 19 18 13 19 19 19 14 18 21 15 20 14 20 16 17 15]
[15 20 20 20 17 13 17 19 17 11 19 15 23 19 19 17 21 16 21 13]
[19 18 19 16 17 18 14 21 18 14 19 25 23 17 18 14 17 18 18 15]
[19 18 21 16 13 14 16 19 18 14 13 18 21 17 20 15 19 16 17 11]
[21 16 21 21 13 21 18 19 16 15 14 16 14 18 17 18 11 12 20 12]
[22 21 17 16 14 16 17 19 18 19 16 23 19 14 22 11 16 17 19 20]
[21 19 22 9 21 17 20 17 17 17 19 23 25 14 19 21 17 14 17 16]
[16 20 20 15 20 15 20 15 18 23 18 24 11 17 21 18 14 15 18 24]
[15 14 18 22 16 17 19 20 15 17 13 18 23 19 21 18 15 18 20 17]
[22 21 17 19 14 16 19 16 20 18 14 14 18 22 12 19 17 17 20 19]
[17 19 19 18 17 18 17 16 21 19 15 16 15 11 21 21 15 20 15 17]
[18 20 22 17 18 18 18 17 18 18 17 18 14 16 18 20 17 15 16 20]
[23 16 19 17 15 19 15 17 19 13 13 19 20 21 17 19 21 10 22 16]
[20 14 18 17 18 16 11 17 21 20 15 21 18 20 18 23 14 19 22 11]
[17 21 18 14 16 18 13 13 20 16 19 16 13 17 22 20 14 19 12 16]
[19 20 19 19 23 15 23 15 21 22 17 20 15 17 21 10 15 12 15 17]
[21 18 20 17 15 16 18 19 19 15 22 18 19 14 20 18 19 14 19 17]
[21 14 16 19 16 16 16 20 19 18 22 13 15 19 16 20 20 14 20 23]
[18 17 20 21 17 14 22 22 12 17 18 17 16 17 18 23 18 20 17 21]
[14 22 18 19 24 18 21 19 17 21 16 15 17 14 18 18 19 21 24 19]
[16 14 22 13 20 18 22 17 18 18 17 16 15 18 17 20 18 14 11 18]
[15 18 22 18 18 18 14 17 13 21 17 19 18 18 17 19 14 16 18 17]
[19 21 15 23 19 18 21 18 13 12 22 20 13 14 18 15 22 15 21 22]
[17 15 17 21 19 10 18 15 13 21 18 24 17 13 22 22 19 16 18 17]
[22 15 19 15 15 14 20 22 22 17 23 16 20 15 18 12 23 16 19 20]
[17 12 19 18 15 16 22 19 16 26 19 15 16 17 22 16 22 19 23 18]
[19 15 20 16 16 18 19 17 12 17 16 16 19 21 20 16 16 18 15 19]
[13 17 22 14 19 16 24 17 11 20 15 17 18 17 15 20 18 14 9 14]
[19 19 21 17 21 16 19 20 11 19 18 18 18 18 19 14 15 16 19 19]
[19 20 14 18 20 12 21 19 19 13 16 13 18 17 15 9 15 23 13 16]
[21 18 18 12 22 15 23 19 13 18 19 23 16 11 11 14 18 11 14 16]
[14 15 15 16 10 13 22 21 19 15 19 21 21 20 18 18 14 15 15 19]
[15 16 18 18 18 13 18 10 16 22 15 18 20 21 18 18 17 20 20 17]
[16 17 17 21 17 19 20 17 14 16 12 13 11 21 15 18 21 19 20 20]
[12 19 19 15 16 23 12 13 19 16 18 22 19 23 21 15 16 18 21 15]
[19 20 16 20 16 11 15 9 15 20 18 15 18 16 18 16 12 21 19 19]
[15 16 11 17 10 20 17 19 17 22 16 17 19 17 17 17 22 12 14 17]
[16 17 16 22 15 15 22 18 19 21 17 18 18 15 16 19 23 23 15 24]
[17 17 15 21 19 13 18 19 14 23 13 16 12 16 14 10 19 16 25 14]
[19 23 17 14 16 14 22 14 13 15 17 13 14 15 19 17 19 18 18 17]
[18 20 20 13 17 18 20 16 14 16 16 16 13 17 19 18 17 19 15 22]
[25 15 17 20 18 15 19 22 20 22 14 20 18 20 19 17 14 13 18 16]
[16 14 18 17 16 18 18 20 22 17 14 16 16 21 15 16 15 18 18 17]
[16 18 16 21 19 18 17 15 15 21 16 20 16 13 19 15 14 13 18 17]]
```

Рис. 2.8: Выборка из Биномиального распределения $N = 1000$

Результаты выборок хранятся в Bin.txt

2.1.2. Непрерывное распределение

Генерация выборки распределения Эрланга с параметром $\theta = \frac{1}{8}$

▷ $N = 5$:

```
[[48.39261689 41.17422288 43.36944286 71.7624318 59.32402  ]]
```

Рис. 2.9: Выборка из распределения Эрланга $N = 5$

▷ $N = 10$:

```
[[ 63.80993491  54.16008744 134.80767155  56.04152116  80.40087534
  51.05939917  80.53562402  62.44990154  46.42689598  36.69552625]]
```

Рис. 2.10: Выборка из распределения Эрланга $N = 10$

▷ $N = 100$:

```
[[124.23314271 106.47509005 46.14745653 75.78720183 95.58728627
  64.68040426 85.83208567 62.28239237 56.08199766 118.17871757]
 [ 96.47319334 80.72717939 49.76364468 76.60251507 67.10642779
  58.9404096 69.75760367 49.36183184 59.97810404 70.45526483]
 [ 61.60191596 58.17241811 57.87455806 47.40087086 54.64136078
  56.46416734 75.89621456 47.38864508 102.6029585 50.29170088]
 [ 50.52990954 73.56636523 49.0868352 53.12586551 63.83007289
  67.89954713 31.99110169 26.74140182 103.99213855 40.95348919]
 [ 37.07156388 66.59650772 81.51426884 109.62390448 65.93429249
  49.88774331 59.49859627 37.3372381 38.4843242 95.50372123]
 [ 67.84702712 71.13275511 62.04107867 73.87210425 43.97240255
  101.4922601 49.31419605 58.61433417 45.67464134 33.15063135]
 [ 52.78012984 86.57540562 61.10612051 71.67643889 70.68144014
  63.91240675 77.42901031 86.73910894 58.05474957 62.59241982]
 [ 73.63987462 53.34466804 55.22587636 50.08151202 37.6911579
  42.5679942 57.90129544 76.0365481 89.30161017 45.06074721]
 [ 73.58130624 60.78471952 68.92251578 40.13993452 48.23325034
  80.31903652 64.3002752 34.16859684 41.24280229 52.06320563]
 [ 76.66005329 90.50531434 101.43913059 37.18453657 38.5276091
  57.40316327 57.32157155 98.35852863 44.32874197 115.85375992]]
```

Рис. 2.11: Выборка из распределения Эрланга $N = 100$

▷ $N = 200$:

```
[ [ 86.34920105  58.52581413  61.04825852  59.77713382  69.22544898
   52.37652267  34.39338473  42.14762212  50.65284774  61.71265749
   60.31163289  46.00218222  78.51716635  61.09301056  84.27095282
   46.55045611  91.12448136  92.21359987  35.33156583  69.17983985]
 [ 28.40420579  59.486614   58.13867504  45.47663264  67.38446645
   71.80802978 113.13501867  91.48169464  70.41970622 107.38128074
   54.34993617  44.60786249  52.4658485   92.29458023  58.40269515
   47.55766342  57.56684239  64.70768075  50.95919427  54.37887566]
[127.34754645  53.06735285  65.12652679  68.64883228  29.37373817
   51.92629477  75.40159824  52.80771519  29.78812658  92.6464579
   40.91488442  65.36181813  51.06975974  47.72345902  60.47636265
   31.3750001   43.15756802  76.54498687  41.14384514  93.27420519]
 [ 92.94819452  25.80481292  54.71043196  68.38477244  92.40233891
   92.66386595 103.62460527 110.57501042  71.23476927  50.43114766
   56.00118143  48.34822931  30.5572743   66.00941452  86.96491442
   53.78040201  54.72833141  88.70936652 101.89562467  35.68730533]
 [ 37.17134233  51.24462813  56.43327349  69.00824039  30.38908798
   69.36557567  66.67192645  90.22679804  44.36434982 135.2125155
   66.81832877  38.02777253  94.85995201  86.1809936   97.27905865
   49.17211675  78.11719265 100.48455818  62.59645627 106.15913758]
 [ 51.06387426  71.86117798  47.19716052 103.47826105  67.1326484
   69.98483833  58.02737342  85.00622177  40.41552699  52.96866269
   51.17940607  53.10242641  71.26907213  46.49586328  72.80514575
   83.44695384  71.61745924  71.37687777  55.213717   79.70008657]
 [ 46.68632063  64.71654231  33.82225925  83.4544095   41.10070057
   48.35884699  48.1232872   64.47463175  57.3208671   67.00539553
   54.81651437  69.84194692  83.46145518  79.37344853  52.04053186
   92.72213178  48.25655267 117.86562094  48.75498563  64.89835664]
 [ 54.5749121   81.04960295  45.85139938  87.43041931  60.14325743
   64.99477905 102.88630574  57.16446274  95.47993244  63.36298552
   72.72590543  58.61354457  47.26428749  60.6789475   45.47157035
   38.21136029  76.97837696  48.43321079  70.14767782  51.27456287]
 [ 52.56838601  83.4563861   108.31524203  76.59234576  43.10622766
   71.16037164  37.5447593   72.95688596  45.08357152  19.97161044
   33.39062131  42.88320858  24.02320274  57.15701744  60.03628436
   49.03372884 133.24152979 126.38096721  50.41340958  53.38681949]
 [ 51.58578531  93.27532002  35.49809108  47.5494901   102.50892719
   113.49336546 124.58512132  62.36044692  67.31152813  43.35673858
   102.04743311 118.25262802  87.0972909   117.1345623   84.42544944
   46.2709443   35.70880748  84.05198282  78.33836575  30.27214189]]
```

Рис. 2.12: Выборка из распределения Эрланга $N = 200$

▷ N = 400:

```

[[ 33.09504301 31.24769992 55.09025973 67.59840835 47.96521811
  57.36078542 27.6171403 33.06771174 83.29812202 70.0715908
  98.1856946 95.84746676 40.47120813 82.59425289 54.29403494
  54.26289167 78.86632266 71.07947072 66.82138338 121.51429058]
[ 55.94237941 53.15573028 52.52143948 111.42236313 48.69162743
  64.44502838 37.79724852 75.23216728 65.30605987 59.89084214
  34.6297426 29.4835704 55.42791321 61.33544646 86.50419723
  92.62626459 36.96637154 37.21362911 44.88753035 65.13137992]
[ 93.28966579 47.45093916 19.36895563 28.78824817 96.20368274
  50.20294764 102.3988646 31.80148539 56.74749008 64.13503177
  69.2234156 60.09273522 103.34260834 29.54941405 46.28300662
  34.0545817 61.81078308 124.9486724 63.62977721 62.30083733]
[ 87.30943321 72.14765159 78.8113527 62.85847404 41.2932576
  79.63177943 60.63433829 134.14208556 111.95550434 55.30360056
  102.41370779 57.03949987 75.6464405 102.0906933 89.29966059
  68.91761281 87.44956216 31.07325171 109.28048674 72.99830242]
[ 86.32524697 43.14366479 66.58111808 37.78275268 41.06678332
  40.38002293 66.29954506 44.84519692 99.88010109 54.79335372
  122.80964294 84.43191352 40.72821857 81.51052595 41.29561877
  94.5775836 49.03511381 62.49656236 99.83151842 78.61000925]
[ 37.09801525 48.1943211 81.26796475 61.8353469 85.24268311
  56.70157471 72.19388618 61.56815836 41.53269839 52.37651018
  60.76623577 64.61318818 63.12421219 90.8108721 45.41182994
  70.10619792 52.2660356 51.50131889 29.12019805 117.37229018]
[ 29.61008388 43.83680386 40.77413384 71.23124476 150.04872126
  80.46071478 65.83977443 31.80511493 76.12768673 77.21379518
  57.89104461 54.83810075 61.80632995 83.79232144 68.08722951
  46.15344806 44.16269124 108.28140787 53.85378626 48.05490533]
[107.53220266 62.22909592 80.82525636 21.25208824 90.48551168
  200.14971131 48.01332042 52.56855476 65.83307288 69.98279349
  63.56737399 61.72260906 42.31703675 96.76488256 53.36728558
  93.37490706 88.18919521 55.15543596 74.7282797 46.30559796]
[ 48.65343342 25.40827621 90.17299764 106.19074846 65.45834987
  50.39841821 52.10864366 85.77769435 62.31147973 47.11270666
  50.65236243 48.13488104 67.46593915 70.66145429 30.94110114
  109.10389815 72.78360957 58.24885981 75.3196629 57.16525289]
[ 89.14290425 61.54151924 78.80398677 40.98677585 45.14019063
  73.20839492 42.78456494 70.25146286 53.85527905 49.99945634
  103.94340434 36.08024066 66.71880348 108.00498323 97.44745872
  77.03865743 60.51735571 51.07734469 108.60239971 77.72772523]

```

```

[ 72.59433937 53.09889889 53.84866872 36.92083736 95.63642414
 70.82696356 94.23720236 17.6920909 66.81746062 80.38678397
 57.65653678 101.51414054 45.38460204 49.1357888 38.95472687
 73.24598975 40.88762323 62.88971263 38.93793252 103.37626681]
[ 57.16159211 99.66652396 37.85281371 28.63067332 75.93875697
 39.92550084 58.72378731 52.8956565 50.50188551 26.75571552
 84.98546611 113.33812536 44.78405675 76.6763764 115.15368054
 44.63627027 35.25449051 41.28744408 45.74887841 86.90647758]
[ 43.6466159 78.27213596 67.8046327 39.34795479 120.37446786
 63.51704806 52.52412465 74.58534745 53.77479359 100.91852163
 75.95108885 102.79570235 86.1984971 113.11555857 46.89675281
 51.98966813 95.2400463 70.04029718 55.89042288 41.53772047]
[ 47.8558514 40.39519214 50.45900414 79.43292731 48.04873008
 83.05129625 81.09648342 76.56623415 73.14097852 45.59438462
 48.17208895 60.59843355 49.63421353 83.27734945 86.95545455
 72.41745006 67.06563686 82.25316733 87.48882683 54.55984343]
[ 57.25270138 28.43030939 65.22029678 55.60552905 64.58681161
 74.52926617 34.53126831 43.25571034 52.96294719 77.01825715
 105.4993497 53.75883009 66.66646814 62.66973272 97.75653278
 38.86312217 55.94872184 94.2960619 35.17912295 58.9489495 ]
[ 43.27309821 49.44939498 51.6246552 56.87777025 32.72797539
 69.26883626 60.99937481 37.23121751 56.50623879 23.80380677
 47.18057311 58.39648235 25.92581523 71.9556719 71.12335521
 70.01067355 54.66313056 71.72026506 42.75288801 72.31149587]
[ 70.25656051 56.0940597 52.38208516 55.88701473 59.94230369
 56.03735913 54.34779 59.76841587 45.72278324 34.46969443
 100.64182602 45.61810386 42.21786351 58.91941001 72.33453281
 90.03090488 44.29244321 62.35036888 81.20304079 83.82948925]
[ 76.58171613 46.46874546 87.5121103 44.90049807 34.46030691
 44.88327192 34.16505147 62.21763463 62.51364322 58.91705203
 56.71632602 66.41070692 71.42655978 70.88301205 55.8510526
 56.45328379 17.0223768 55.56453303 43.7658012 42.77828977]
[ 59.41562027 67.98054614 48.28718221 33.79486752 76.69064412
 93.65483995 57.45743465 59.47254605 61.18449787 28.9366448
 57.33030557 98.73736971 52.57126488 95.28733531 48.3365648
 24.2154238 57.80779897 44.29055697 50.89570367 51.51636825]
[ 47.37111491 57.53087452 91.54626912 111.9160897 100.75572525
 72.02926822 61.39786111 47.29934289 58.08023241 64.49794402
 66.17515198 50.00903511 83.92316311 72.21848603 73.27730861
 38.19349476 73.71665433 60.23815525 38.61977494 53.76944461]]

```

Рис. 2.13: Выборка из распределения Эрланга $N = 400$

▷ N = 600:

[62.55278514	47.57928393	77.30494151	62.27174804	40.45728293	51.9837304	50.6433313	79.07183399	39.42620018	60.12979524	53.87693027	55.98363201	43.6650706	73.88715788	49.92408095
46.99321162	40.62460697	81.51753757	47.22757051	70.27652669	50.13199394	92.17404198	55.86202092	31.91359718	109.1916706	60.57709385	98.33184082	48.86157968	44.82929253	65.00705628]
[63.15137108	58.19592478	61.6139178	32.13346785	54.19082102	111.16369792	61.87944817	65.62842999	36.08012638	99.35813785	72.70857544	75.82502564	40.99630828	46.61211809	49.64026498
25.81770614	88.27769129	78.5685025	65.09150273	58.3603188	72.95379897	61.46803848	82.91959018	47.62103907	38.39696758	55.99523885	74.06912273	79.34309391	47.92315209	61.43653031]
[49.11418878	49.53866583	76.86614008	83.47346486	68.34840889	151.52180121	71.07367613	76.52531079	89.94559382	60.94843408	37.15089216	69.46571735	35.9855885	54.08467652	59.90732302
37.97207392	93.68338682	56.47264003	47.05112311	76.96086815	68.67972678	40.26438244	108.17533793	54.3272475	72.37715633	58.15682963	31.90762917	83.06030776	22.56504489	119.25726088]
[59.85466999	77.08306612	99.41108962	67.40316087	59.24334669	88.64202915	68.32011557	127.95646256	53.80746224	74.41841111	36.40576975	52.61432277	66.23256027	76.4965919	66.98437
142.34410021	39.25232396	58.89431859	34.60458927	76.7442326	20.80418915	35.72239699	57.12539726	36.26166342	85.46175825	111.99331427	31.99315978	68.35894366	36.02293508	46.36298754]
[92.15487063	49.3865712	68.23112439	66.55080649	107.73226222	79.18273566	84.27675365	42.14241873	86.76253335	64.59053221	98.40723977	75.00015392	44.02039597	48.63729687	23.41819214
92.31383767	42.27382269	45.25254399	53.47231447	86.9292813	85.50333807	73.09562552	63.90627245	57.6094814	75.89348897	57.40829083	76.01973466	55.81853016	61.94071153	49.39790982]
[72.97405162	38.79582277	46.8221443	54.46647162	47.29117838	44.07871267	36.72029258	69.21214924	48.65943244	30.6560351	52.92614886	86.38182384	53.36015571	42.58694225	51.31942823
61.61252588	42.67756461	70.96263766	23.70308598	49.48115758	51.97788551	62.90263645	101.24900414	36.32025765	74.23776593	70.32834337	67.18774735	52.10894617	81.44196839	44.362103]
[104.03994642	63.18310113	92.65729409	56.32328134	42.53257978	104.43157699	54.44986807	68.12315335	31.8411457	87.89128168	36.04505924	40.39755237	98.08321682	55.55964191	65.61040857
119.55358729	40.8713143	41.28100173	63.70376468	35.56754652	68.48748245	70.18497504	62.72635233	56.38265724	52.9108645	39.05344969	109.30862221	74.20498562	83.48929181	64.8322808]
[31.19505311	54.25598591	69.63286993	124.44118759	51.44706706	67.912849	87.317769	51.31801515	96.51095753	27.22820408	76.63016157	76.48940056	56.42389474	48.50954195	27.59205121
67.57590637	53.36885976	58.58897692	42.70493894	77.72007744	66.74441835	49.91304238	126.83791828	77.08216921	70.40044994	50.42175688	40.05717441	53.69871596	57.34149508	51.18276296]
[51.26393909	57.76874401	61.51317914	51.51480054	78.23780491	72.87457975	47.07695932	72.52656905	97.59081308	53.60961	67.07114924	38.1840837	79.81734284	47.59938907	62.90345678
51.7231677	45.73096562	94.48115631	54.23004745	40.44807406	66.55477984	58.92975279	49.79687645	72.72855078	35.08809401	62.74714125	66.347257	52.32876262	49.10427752	48.55725369]
[47.71273584	32.88266554	76.71606088	66.98385957	36.3264866	65.42451582	52.16075273	120.7585308	49.74480754	91.67291848	83.46552124	75.30540539	83.58689188	103.91386543	59.7794521
50.96045976	49.46348026	36.40975942	42.74357981	28.68056029	21.03118752	66.23761016	77.99422624	104.9348511	55.67240519	94.98736379	29.61643678	29.59495156	88.61231983	52.97397301]
[88.76808096	53.62312966	72.1089951	49.89983172	99.15454083	55.43534564	111.02791012	35.69709767	43.89139603	42.20708705	94.733235	57.54272877	70.98300929	55.14191573	77.03117307
68.58916778	76.75680671	53.27712351	39.57797031	74.18626449	82.4145854	94.25630781	73.42054972	62.80344705	53.36709226	45.28150298	39.90689469	52.4874468	38.37452413	80.33222066]
[45.09987311	43.4141516	21.32608945	46.70022774	68.2131392	72.81280982	82.09476736	104.03910899	60.24566884	40.09496527	55.7927473	70.47055344	22.05198633	55.07270177	91.5896745
35.76055954	41.08797947	85.32822438	69.99566485	99.41969028	57.14259644	104.0428134	52.48957104	50.20454212	56.15623723	76.37431782	53.03528559	76.22593594	69.0370319	51.18384118]
[46.62686365	58.27683678	26.80154217	66.46311831	85.2230521	100.41228987	71.57957538	65.11679322	56.32306151	81.87315718	48.59108452	70.95095288	50.55908254	55.75331418	61.24503084
114.13227511	105.31005169	60.80322774	52.47956601	50.68690135	93.48267881	73.72378024	47.06471656	44.49346151	45.90469052	76.54720469	75.33876334	69.15088936	41.89406627	68.1826813]
[65.94813258	66.94689939	79.15028141	70.46779959	74.45139945	54.62828726	80.34677481	66.80964794	55.64146391	48.2936454	60.57361768	50.67497956	39.16980648	55.9475368	73.14571094
88.29217716	39.29577286	70.10736389	56.2619552	57.10950311	59.90490401	36.63317441	59.67413891	25.5034407	40.26617337	36.21103149	58.92820845	55.28169944	19.46076581	60.48911096]
[99.50724248	54.67357593	140.14420405	50.90097107	60.32713403	78.60413725	52.54061458	52.41944491	95.65906411	72.40867757	51.05689439	77.87530942	51.40526382	64.27372492	91.63527217
24.82725931	65.25616173	92.85210639	76.30139455	72.89518368	87.67532348	72.51808857	32.4209726	41.13839048	88.85023775	30.18236061	50.55921651	58.05967486	80.66153505	74.72653383]
[51.60871914	51.70250124	58.27983145	81.55492559	80.82799724	71.27149443	78.42067956	80.73954053	102.28159976	54.5174476	80.89101024	41.18438145	110.72280919	105.89218182	59.74695565
69.66309339	48.72218732	41.96708965	64.71142623	48.6211045	48.39472851	75.110838	38.37416491	94.49975395	45.27753378	75.29048294	85.30841297	51.15462517	30.96425433	108.47961724]
[62.86410541	40.55558514	71.88966776	81.91716535	53.09380002	48.71377943	79.35962079	74.23211424	68.55670453	64.85465065	35.95677571	97.44014179	67.35802779	58.56159275	53.75024843
54.2780361	100.9682425	46.02161094	118.44319494	48.72006635	40.64988388	58.37906813	57.8876799	66.29115681	50.12966767	50.11705642	42.43070119	39.98007984	72.72287424	49.24285813]
[59.44707403	74.06222223	55.36644111	66.16733848	42.88760832	58.44880253	42.1510713	76.85521435	73.41859366	39.71161853	92.35270417	125.33198431	38.43194141	111.7983461	26.73430217
67.77001243	38.02552226	42.24955831	35.67453676	75.52323529	37.45445306	81.47113519	86.71924097	58.59225856	54.0671676	94.80102799	62.20922119	39.54302896	112.79071692	50.30903944]
[28.06461578	58.41545527	70.60316558	34.75957519	60.76464978	48.60593953	82.78996897	57.62752921	39.73514146	84.58398597	79.86881678	70.72449329	79.19420757	75.63109744	78.16570446
57.19520125	79.56787815	62.2480538	56.86808805	62.3330456	72.20909336	57.59323652	87.8198631	51.83610953	67.39062736	31.52846654	64.79879309	83.92076395	70.2700192	32.98564937]
[77.09565914	53.07480005	49.54721653	84.35523417	58.37926478	65.93723622	62.97221076	82.91585661	68.11764165	68.44817882	70.85530546	45.23740209	65.09443354	55.54661707	57.03307988
74.6346761	26.24839912	67.18095424	96.36179309	39.93442241	24.15064836	89.07225219	55.40612444	73.44738323	97.75129105	38.55893547	88.77621685	52.66290587	98.55494215	26.08356208]]

Рис. 2.14: Выборка из распределения Эрланга N = 600

2.1. ГЕНЕРАЦИЯ ВЫБОРОК ВЫБРАННЫХ СЛУЧАЙНЫХ ВЕЛИЧИН 25

▷ N = 800:

[58.6116164	74.99089629	52.94698857	58.47402749	107.89815674	47.79360808	76.6710941	75.70001046	91.33345543	78.2530757	69.05877213
	69.16962795	77.02347834	79.18686236	52.27056222	66.26741157	56.05236159	56.78511013	80.16322268	61.05065902	34.46063942	45.1774703
	67.14808779	50.14027131	32.34716292	27.82400344	83.22944313	38.83527722	88.34424753	92.42654587	42.08675901	34.44008033	71.16573411
	59.58024673	72.09642661	99.77976117	92.40890291	79.45251444	87.21280278	66.60116611]				
[75.35548371	101.26567346	52.59775178	62.50970158	57.58209985	66.39650319	69.29106329	76.92160403	78.14913403	59.26029614	79.2030496
	72.28077807	44.74899276	73.74627328	51.81270937	55.4229425	67.15442265	82.84684221	80.72119462	77.58149052	36.86391788	51.93091058
	53.65769858	34.83096848	36.60526011	74.27984765	86.28729371	50.18506423	43.96275541	101.56016017	73.52223945	67.49480099	78.61520238
	61.86024801	95.98236328	67.34292585	48.80199603	85.69696031	77.96180078	62.06316066]				
[81.49939915	93.74516121	39.59337852	84.37087365	35.88657564	68.27786643	84.44247493	45.01659848	86.21301268	66.88020569	47.07643066
	30.15340218	56.38741634	77.6279999	46.35009874	70.15974093	96.95887864	72.47189455	81.73878112	53.68338046	54.70092499	30.07663126
	100.0225901	89.15615722	82.84432108	22.36946504	54.08423093	37.40220254	98.9502788	31.51810329	85.7528123	97.2206791	29.89430314
	72.91199327	114.53224927	28.37080344	74.00009275	44.59658778	48.89799062	72.14156814]				
[65.05061487	57.41940925	97.87009005	105.71646599	146.71752438	66.88026148	54.43062733	88.1775214	34.74492022	65.50075838	67.41134587
	89.06492131	70.09081353	68.28724809	34.47544531	90.39080892	42.30905539	31.94612879	50.35926366	59.8533744	37.49938869	72.68635807
	79.77603773	43.56352692	62.56643852	66.62906823	106.08539214	54.89229511	110.51846276	65.8619998	29.29162959	65.3603951	149.70463095
	151.14476487	22.77869381	81.26470976	40.14046195	82.97299678	55.44380087	51.91274141]				
[44.69809507	48.10052392	60.42708895	80.30012123	83.59280771	39.66865536	68.17609603	99.07786355	59.93749573	85.9944493	67.00113636
	42.04767652	56.05060157	58.60285295	33.46466891	87.78718794	47.20725406	95.32994468	75.29401178	53.44209049	42.99841816	72.28934971
	88.5321244	55.91333071	65.80139652	71.77241902	55.41070118	65.4664864	72.082816	117.44854344	40.94348292	57.20926746	68.41098873
	39.05699123	31.66967911	78.05227214	72.24147992	31.26807951	47.99184408	65.70547677]				
[46.19723898	100.69441324	41.33296401	52.05228641	53.49238626	87.1681503	57.25587125	99.22118518	45.49320791	82.4563306	104.86396812
	54.91924357	64.15233356	68.72373822	52.01321139	73.62407682	84.83090092	80.49514947	92.16327374	61.20807837	96.66792897	41.36408534
	50.58161902	57.53166557	83.51809687	17.37665874	72.69311725	37.24379613	34.73221319	39.52038122	55.89646192	47.14183034	56.50858569
	32.40698851	48.56479932	44.27220911	48.84681313	80.75052565	98.70311016	84.61663436]				
[62.74357224	82.65136629	150.77371165	77.62377873	85.45600798	75.58444783	35.77780729	75.34592761	76.77965969	47.95648438	122.44027927
	91.12654667	75.18490929	28.13679477	75.02630164	36.6718931	60.42164102	81.39039381	65.21161181	37.26809442	47.59499459	62.69960966
	78.20161519	63.14782222	81.36401629	51.64744064	74.00665556	29.70284674	62.26212721	39.49984543	68.74555634	82.15044977	40.60726918
	44.13606063	80.6430065	63.58463568	104.70558161	66.56782722	50.8907718	95.27107449]				
[32.33571583	47.75980577	70.24805784	42.08887998	45.43838704	58.3382121	57.21226766	79.56606154	35.08044978	106.3619671	75.81061832
	31.2187128	78.5976671	40.08747576	71.67623428	118.90608194	55.48509087	88.51115986	50.88076478	47.39713184	101.65069545	33.09847875
	79.23717588	73.90045187	83.01381575	100.09380811	78.73772319	26.55246792	83.69545795	85.45922617	48.16565228	50.16695491	111.29394933
	66.18832284	52.60435472	53.07251994	59.55463933	98.42411916	41.08595965	63.35226117]				
[60.31233771	43.02575933	53.32685209	47.79100263	112.97193019	50.39460967	58.13957537	71.56144764	62.54945709	92.95905403	98.81991997
	61.3973198	38.32026337	50.01820115	35.2933732	39.18699601	69.14175673	71.68526832	60.32419297	69.80448909	74.32661133	50.20619379
	32.70938679	65.71818679	37.02191131	35.98662726	83.49179288	76.68863533	96.67761477	63.12776112	41.64383592	62.77023133	51.14499574
	62.59083253	30.68242682	115.59449222	66.88654005	55.85057924	37.48906635	65.31739873]				
[76.5526197	44.85831591	59.26721299	85.07347423	61.30601613	65.92902676	57.55808875	40.02410662	42.67984797	56.33845432	49.92347341
	110.75781513	43.54316408	49.39706744	65.56693023	111.10849191	46.84327814	51.11404946	91.69033668	32.88010018	71.80501775	95.30901711
	48.42300109	49.88709191	75.82088123	55.82426234	59.46304157	66.81829621	128.12363168	68.00942079	57.54844961	83.79199999	57.48475017
	57.92903524	73.18985	48.65054156	47.18815382	47.02089137	38.26197498	50.90407668]				

[64.37140327	38.70323289	88.86278252	39.8516678	54.24238479	57.47633533	50.30191425	44.40812938	64.87388243	55.6574805	58.21875896
51.9360598	94.71303377	61.91235809	26.07313533	72.72177628	45.75954134	77.34375526	56.15370495	44.215268	73.98324162	45.73456384
56.92810156	62.6053756	72.11214103	54.44959361	104.14337903	42.17762712	43.6208741	85.47183985	52.10626519	69.03160285	79.36758461
18.0574119	81.40768136	76.38045315	79.36992288	97.79868851	98.14776411	69.91214558]				
[36.25090485	54.08077473	70.63639511	100.27322262	58.40642883	45.10550532	105.66239078	48.72147548	57.06732994	45.88492216	70.37568756
89.08731916	33.44187965	88.32539737	132.01992727	36.76339189	42.44425858	75.09673165	48.44034808	169.80679766	41.14503135	20.35890169
99.8584462	87.19678136	86.92959148	46.97249676	64.68744212	76.019387	58.90567645	50.27559647	45.24250296	74.44184575	68.5416336
66.58691302	52.45884913	62.03449063	74.05834853	70.25060177	42.84526822	50.56636208]				
[101.474597	76.50865718	93.85299605	59.26465212	53.00405819	56.8899778	32.70866287	99.26808688	59.8176566	69.27534348	89.79107985
56.67284545	54.72556678	92.92407555	44.20076289	136.99156818	64.10375681	40.23348249	49.74036083	46.15975187	91.02963465	94.02327965
33.20818124	84.23274399	59.90099927	75.29607468	35.83611691	120.76626792	46.51936922	23.48791573	44.24206903	44.6898594	103.16183312
41.85038063	51.12786623	35.59564831	86.4307339	28.29538161	37.93100056	64.79015531]				
[58.06453683	76.87457336	125.02921153	64.88026582	67.93441205	78.98385024	93.09838196	62.04377521	77.85364636	61.06067243	63.99139941
62.39754715	30.33557426	85.60438415	37.19753913	53.25896771	121.96840604	74.94478555	79.72618892	90.0115038	50.53884351	59.57628513
115.57390287	39.10784994	78.62107107	72.69399914	61.10719063	101.46232794	44.85652548	53.0092023	55.84753753	39.46502047	52.61427444
48.55478304	45.17855358	91.25124424	73.86559878	70.145818	49.32453345	54.27886751]				
[41.02691267	66.56523213	77.6377161	58.69122392	44.17247734	88.75369552	52.26916624	118.15204438	81.68743942	66.01967891	34.70240492
34.30755074	34.9307326	79.05012334	42.33892806	66.68455111	41.14341545	88.93162838	60.39756144	104.2061942	47.51764732	65.68695287
47.56524662	31.6708605	63.77309032	37.08891104	83.01329367	52.86505622	84.1476001	79.51001033	57.24967974	73.35862263	118.54663925
74.09870066	85.21689031	80.85119982	55.13696201	80.80948411	68.21313585	79.79301445]				
[104.15266798	38.24388657	45.27248103	53.52365137	78.12420579	42.77566007	66.03447431	47.09912667	89.56975784	52.30450343	107.45941439
60.87691402	54.40394499	57.6858867	55.47751817	34.70669077	51.1471069	57.08364834	94.83205896	25.42173345	41.16842587	48.34745753
84.70149171	62.08829401	58.67209313	54.08477477	26.91480427	24.28262746	62.51741584	53.31866803	34.91531568	69.9867108	56.63651616
49.44619572	72.18475756	56.01165625	51.41670186	84.84615374	52.13120562	49.60551596]				
[66.67027527	77.15223703	118.44841814	29.11420829	72.4283381	62.13528655	51.03619833	28.62544916	71.58657956	55.83245085	57.19491094
37.45361548	77.87460646	72.76061557	62.5000095	46.20131576	148.62105183	76.78304874	33.73895156	26.60987429	63.4325312	61.32888137
42.25515018	28.14805444	41.48748738	35.71138703	57.60564593	49.95589376	25.11859813	66.66410788	67.96489585	32.44619291	62.88193865
66.26345509	98.55980656	95.34271163	115.71195649	63.86089594	55.31035962	69.44609488]				
[74.60984437	85.47028735	58.91233413	63.94259127	37.05482238	63.62295955	39.61305538	90.66383725	110.05812822	128.51274871	66.34103163
122.37606203	46.912829	55.81798838	121.02030388	42.99522873	63.183219	53.21567394	43.84605347	45.61977286	80.87036299	71.81535253
57.78079224	58.28508577	60.26373424	62.53455587	135.87802809	29.95895447	94.92952755	69.33367057	24.85406634	58.33076594	54.34919448
86.36712676	65.11187368	42.39336019	59.41944548	53.26407125	58.54021567	114.64049864]				
[64.596171	40.90332347	93.08833518	92.03001526	111.89086198	62.39904588	49.45323443	126.40940941	23.0658205	81.84876302	58.23432107
73.53489446	56.55030209	67.64353556	54.26242925	89.09489928	48.56399073	61.10351026	77.56788298	56.89604719	38.01945137	44.15517805
78.48351114	68.56578977	66.05660925	76.75149779	50.626047	86.79937966	46.34431759	28.86167621	86.58017997	41.63821876	77.17724138
85.66364083	51.02386833	59.35451128	56.71147543	78.72457165	86.07649728	82.4048339]				
[105.0788451	86.78634245	102.24465061	69.42534016	65.7515989	51.87325372	52.528549	36.90945029	60.25408992	35.50916426	72.58713005
54.19433594	86.713656	68.86442936	39.3490895	47.01325285	53.71883026	77.49039305	74.98694399	64.24443364	73.4611264	44.39214409
76.40460092	91.86706523	85.46740371	65.66799048	55.40928893	104.25468707	79.30262658	65.13405327	53.4956799	66.46981001	71.11297236
65.45492799	64.60851775	67.60611438	77.16744671	73.2768256	53.57708462	62.92170499]]				

Рис. 2.15: Выборка из распределения Эрланга $N = 800$

2.1. ГЕНЕРАЦИЯ ВЫБОРОК ВЫБРАННЫХ СЛУЧАЙНЫХ ВЕЛИЧИН 27

▷ N = 1000:

[58.00094425	62.03649087	63.79952489	43.93628329	92.27937973	62.03657341	54.502684	34.13081719	75.12619525	46.62919534	64.45246249
	56.53950039	63.0415339	64.42597853	89.6314779	77.12201006	42.56349393	52.16660246	40.47641148	76.1951719	69.56267344	34.22820696
	52.68640732	93.15420999	77.90603509	75.61658885	50.30135609	44.92149018	65.45514643	61.10768581	107.68769338	26.13176751	60.07234197
	47.35984947	97.15227328	83.24466783	77.05151263	66.27172214	69.21075283	68.25281114	83.20234298	35.45704553	85.923835	27.45625706
	40.38364823	46.48642066	70.66520951	70.65524928	69.94406434	66.90865536]					
[33.52415956	16.78488515	61.57829125	52.07977585	48.05607657	146.54224231	88.87222685	106.17627539	45.76022776	69.88604766	65.47700218
	47.06626033	72.13282266	30.35622651	87.59169752	51.00010326	68.8619228	45.03519183	103.32489332	39.50388536	62.77049567	59.77126049
	62.71646631	57.88998499	115.42593054	83.35318279	79.53151514	65.27307465	37.65811367	77.77887437	40.93526874	78.96638976	135.95175866
	80.93801608	17.40064939	79.29870672	36.90578026	57.56667501	55.56910008	86.38357914	68.58466731	66.65820679	69.15835073	68.51444077
	43.12569555	95.25741059	89.93357756	65.74350014	53.91242908	41.42693324]					
[30.49186423	41.24586536	69.1611323	48.07204351	64.56640078	19.70308708	44.47233708	76.05644081	68.73137291	66.32758428	61.80752471
	82.14238947	99.23942003	64.46217777	41.5652855	132.48754898	42.5481734	72.27070806	61.70101922	58.6702954	49.95629386	23.18781243
	57.56543526	64.08951389	28.40504855	63.19199078	105.62349576	34.05416328	46.45863919	31.37116107	76.22156859	85.2564909	37.26885325
	90.27645958	37.04681266	54.11249732	26.38302314	61.48132953	68.05213307	64.15378315	73.7923975	74.22371218	61.42441065	65.2739956
	62.52016984	132.5737577	55.2289429	67.09798711	97.47275975	51.99772774]					
[56.84254413	26.43279199	72.93240299	66.39484072	32.2037659	34.25298634	64.94129972	48.49111856	76.78922097	68.55538992	39.65206673
	41.26225224	105.9552955	75.06076559	101.74210018	78.03750846	55.59519603	26.66326061	69.54434092	33.99101679	60.94526252	65.01357195
	77.64837609	55.76960501	89.73356435	62.04385396	97.28739242	60.42580031	52.94702468	63.70004195	56.60692726	54.49436166	52.69524872
	34.76080554	84.86469425	40.80662115	85.75039771	28.71386936	87.90315035	70.22704098	104.26854204	41.55126325	45.07097854	86.75372031
	61.52116118	47.75579987	47.15686882	61.37332675	69.4131654	64.5568226]					
[44.43950027	78.29340305	73.39445876	47.95281958	131.9234124	49.42050946	32.34741549	65.0854282	83.12091171	58.1206047	29.13839075
	43.73865225	49.4130737	59.87998301	52.19062817	63.41028904	113.9200817	39.10832282	83.77518756	60.86111906	72.52600742	49.93309585
	59.42958156	56.38895604	53.63878022	49.59051716	49.61747439	36.65198609	123.89999996	84.553803	53.66269017	71.99178104	50.88845562
	52.10108141	87.62256629	47.63910542	62.82904083	41.82461404	45.44583659	120.8363286	89.8437987	37.23475086	50.34961733	63.8968291
	42.80581767	60.93669111	49.83906473	28.69904534	74.48157222	51.00984661]					
[54.77221064	75.46105332	60.57503714	55.46346735	63.42190882	81.71762122	60.9897855	77.40931201	95.00865188	33.45110097	56.26711228
	60.77738175	29.72402244	34.32769832	86.59308244	34.20691039	65.13668704	61.27092421	90.05172386	103.32407693	32.712674	77.7274109
	47.21632939	74.92214208	79.69481372	69.08583998	57.4925072	68.89902302	68.30065812	59.62059568	70.93001904	61.63635443	59.31436049
	93.76419108	77.59049315	64.70119296	64.90258088	45.26314521	107.95559139	81.48722534	62.8611933	73.40956217	53.93289464	31.87419914
	56.56534063	38.927866	46.74737186	66.78706866	52.67133888	45.14372815]					
[47.28349261	32.66321872	38.20321783	155.38046705	42.01191492	89.92470236	60.67248488	49.30449036	44.22247311	36.50157127	58.18580023
	66.2116345	37.28281063	67.65875345	48.27441822	75.6072574	75.76873453	41.63262991	35.58132845	56.51074098	84.8822791	36.35034547
	74.93943469	67.57357545	49.01289678	86.00817673	68.13008576	40.02800708	38.05605828	44.21259634	53.68388157	71.22691836	87.42272972
	70.65481629	78.64359717	31.68338555	87.82597627	60.87764432	45.11600058	54.73750413	112.63372637	43.69459227	86.31405843	29.41232196
	120.18588737	57.36111914	104.83647622	73.89867626	28.74449979	74.68770462]					
[80.06878138	29.79055134	77.32997009	66.08783283	71.50790638	46.52501072	85.64858249	111.23605137	51.54826991	56.14929096	96.92309659
	92.33745164	49.74309705	60.74139166	92.75120687	122.83342977	56.19568078	51.62092991	42.52077848	123.77749327	104.41931775	46.62227635
	99.16385361	49.1307378	70.96292664	59.01428526	38.56491269	48.2887716	55.91185916	45.61385203	70.10933927	64.07954756	82.86135718
	97.07808754	29.89577838	67.84373465	92.60208117	38.93969299	49.16153319	59.08877185	64.79028152	94.46655911	113.98259734	124.77903597
	59.03412237	32.54610393	45.4285082	41.80801477	60.23057705	45.16890527]					

[59.18572211	62.17112975	115.20237715	65.46842695	44.63343769	124.24702844	41.32296507	144.48266832	81.1148997	49.67868982	86.52815205
68.59824198	47.68019749	48.44166066	74.77235182	103.71449301	129.47628461	98.51282552	58.05665678	50.41401132	55.97184515	120.83376369
97.94735353	56.37532501	55.82436968	88.33867001	50.33294677	24.28564591	55.94468787	92.16368773	40.33708735	24.0392167	57.49317453
70.06100684	63.60877444	60.85075208	83.08818992	66.27246546	44.11082537	49.87660266	94.23070317	60.51340165	84.57968364	52.82804252
43.56085241	106.9999983	92.77716163	45.92364368	31.96977632	83.01598153]					
[44.10157484	29.21643831	54.03333722	39.96193763	78.42609931	65.58705834	92.87450171	39.18479482	70.5303553	94.46209077	87.20379156
57.3483222	36.39224848	112.83615044	77.93883275	99.37928276	56.28725639	59.76295071	58.18235757	71.95841078	78.76696535	61.63411894
49.04911285	44.24378249	75.42676857	78.86439193	90.29609852	46.02743965	45.4185368	66.76464066	65.6557086	28.25144616	58.83449338
34.25869844	48.51485702	31.43043678	79.40227894	44.52627416	70.08369482	43.69709012	72.73641532	44.3277416	42.86917066	69.03647268
40.06786181	33.36884261	90.5341419	101.29478289	82.25168453	50.64531969]					
[33.08643454	54.61294718	56.14490164	76.18864773	41.65805787	44.19657224	57.45047087	57.42971921	60.89293107	63.80506564	89.06392356
37.44555583	51.95973631	70.41211367	42.19176928	52.76509128	65.64804843	30.33017375	68.94098782	63.88860582	25.69646272	72.35878731
78.40682796	40.57166458	30.22710014	41.82089746	85.55036315	85.33355273	81.06281095	58.29040082	65.6403599	66.46393951	78.30275016
105.95849347	48.04986521	47.98719836	53.14696915	45.25806433	115.39972349	53.60309352	51.95313686	61.3096715	100.42533947	88.34080762
56.37990949	49.87926025	53.07130183	51.76997849	55.49856065	61.15298469]					
[82.62141892	58.12852308	46.34648728	38.35506793	39.73844276	48.84995316	33.25640662	59.35470233	61.29199202	73.36277735	89.09967986
30.97849875	56.16107469	128.80905804	45.98997632	56.35898441	74.47088639	80.77751192	63.99074125	47.32123441	78.29142237	28.48366151
61.97537586	33.25021768	42.40982388	64.58189644	62.95921384	42.3808543	110.76515169	76.20932784	51.60942047	85.97898092	26.18268824
69.76580939	74.55397123	47.38100924	67.6406104	81.09488631	50.80036653	54.38788954	46.71965251	81.16116712	32.43373034	119.9807004
92.42154115	82.46180294	32.93388846	40.14709374	48.67128854	48.83925009]					
[77.10793976	75.61585312	49.48479594	82.99474181	54.48168783	45.9548638	61.22397062	75.75812141	89.99975348	61.00393979	24.21983259
46.07240587	89.36707853	60.46159905	64.78125158	43.95832528	69.34762184	46.49570554	56.46681101	50.10301443	33.46057959	105.49876251
36.9275941	60.15149329	59.85952177	106.28218889	53.20230299	33.2710955	27.94809912	47.1438666	36.20724439	75.30135845	102.82332976
68.26213697	29.51102723	36.57483351	43.24644142	33.6682818	28.27883825	138.22441626	54.1489985	82.42840448	49.69034382	48.29411118
72.79041442	70.49440731	104.11923294	69.70667029	101.64669468	39.92690912]					
[66.41622758	74.30192456	41.24463091	42.0663233	28.97232538	57.38241927	48.81242341	72.70066531	63.49568551	49.0070539	96.82252995
70.12385723	65.02462712	69.38512454	60.81691379	54.09650635	58.01458031	61.86487191	31.97848374	57.09966615	82.28722201	58.16298852
87.53314951	49.87415183	64.60375979	87.64044367	64.46987018	53.6918092	65.33935668	111.37893385	34.17559847	73.32375092	22.23781224
74.97717833	94.79712803	91.68811535	61.61928561	42.25759147	62.19788373	82.98430408	48.59654943	76.01428452	56.42971467	98.1027645
35.66912227	52.56103572	96.62569169	50.38569359	54.88612047	70.3697751]					
[71.91321136	59.53570242	58.92821967	86.18152028	60.42074244	73.88565995	66.82752774	78.4601672	61.83353884	39.48079291	82.13153019
61.68555824	87.46379955	39.12040569	23.7750678	30.10479239	88.57027274	48.38996275	64.75572198	44.74820711	57.01567085	109.92290418
97.90582515	62.5061842	61.37204394	66.52582419	46.80928503	76.07729931	53.82122244	51.47163495	40.29028452	64.52748909	56.48284694
28.62176682	69.82063426	56.90586354	96.34014556	72.00916473	58.05532037	63.68692904	62.48457936	53.48405992	53.50632398	88.58201705
37.07968332	90.55417223	27.4581169	131.38749559	52.18534656	50.75651567]					
[53.96995959	68.10191477	40.19856524	75.76054	39.25393768	71.76779164	57.69899547	40.17161501	69.52229907	43.28301517	61.20226566
57.07538977	32.40216974	68.80546761	77.31700213	34.22068286	58.37932842	64.27108432	69.60284656	92.15114587	46.53291145	66.48622924
67.35808431	64.02305721	66.19175005	70.68929253	117.30479208	90.62269146	67.74517444	59.03655756	38.16977004	87.81507969	45.66316636
49.15366229	59.21432661	31.40120854	56.82666784	25.09222698	90.18904413	85.91255079	56.92967818	56.75749183	40.87291588	71.72882964
74.08076667	55.12467783	94.7577538	64.80917523	134.17982261	28.00250698]					
[51.10055019	89.94709224	73.47516702	57.1107022	58.79583053	77.92271889	51.33469501	69.54752278	60.3398243	43.90249611	86.71225191
54.62718909	112.35303371	49.05982292	90.28437856	114.11884801	70.28081486	74.20419348	32.06473276	71.05964087	56.26318209	89.01189609
53.81212922	43.43608646	59.37850866	57.54132351	45.14568862	59.77942653	45.9052484	55.41901536	102.27638693	56.99517161	78.14706945
78.39770849	51.05224296	47.76832145	93.8395283	107.1518561	36.48792258	72.56264204	51.83462467	58.08185993	50.14241801	42.99721851
59.13179555	55.00233092	79.84644871	119.7852803	80.26571456	69.6699289]					
[53.79136658	41.10530165	53.78280348	156.81464001	55.7288348	60.51376716	72.23504831	41.46074937	80.15931871	71.61061338	41.86740412
55.78360201	41.11589474	36.95904277	73.17637518	38.07942545	35.69004798	91.4587637	53.09294534	102.80717317	59.02592848	64.34517899
50.45129981	64.75630681	32.66574778	80.80695015	24.89308183	46.22554074	94.28753857	30.93385352	64.79222278	59.59417549	67.03743176
91.77627715	72.06812583	87.48972493	60.63392488	56.30245425	69.49186307	69.57413087	155.64699997	54.24442762	68.42012537	67.6054504
22.23924366	55.92896836	45.87337138	69.90413542	67.8525834	37.3035711]					
[70.08421823	63.40618274	66.50076464	57.45406679	76.51490784	78.19007435	64.8640508	79.03959436	79.24681802	36.99586022	38.86123032
91.57482378	52.92251443	47.20940324	77.42707043	49.6732815	58.99833078	48.83158629	54.0745572	75.74184809	52.53175015	38.36369074
59.77483153	47.93933364	56.49976478	62.79229806	38.19346151	35.57871822	39.29333085	46.00482323	116.60940085	45.16225698	58.58003933
56.92587244	76.44316399	38.70025156	35.24118504	67.48913723	64.92690203	39.12850529	39.16612844	77.58724175	132.62763456	24.31266519
53.14112502	70.31401332	45.30173293	66.30098617	65.79310102	52.9844919]					
[82.02699758	78.45325895	52.36659746	70.81322365	61.16009627	59.18026307	74.60334718	89.3748344	63.26541779	72.99943332	41.42241958
74.21266075	43.36573537	135.74488397	45.14583103	69.02601694	31.58481198	47.03877432	41.68515132	55.73899705	94.16640992	71.39144622
54.42045812	82.14843623	56.19981423	67.68408832	58.04255467	52.81880373	57.43531164	59.99534066	85.14315731	105.6917974	53.44449745
73.53615975	74.85438034	99.90119806	58.5285866	83.58348889	42.33800349	39.62641736	41.25792718	38.64033896	59.08627297	77.1795018
48.00724086	26.86744144	82.04532098	55.51685083	50.42103261	72.13910301]]					

Рис. 2.16: Выборка из распределения Эрланга N = 1000

2. Построение эмпирической функции распределения

Построим графики эмпирических функций для выборок разного объема. Для этого нам потребуется написать функцию, которая будет реализовывать расчет значений для данной функции по формуле

$$\mathbb{F}_n(t) = \frac{\sum_{i=1}^n I(x_i < t)}{n}$$

Листинг 2.1: CDF for both distributions

```

1  def custom_CDF(s, t):
2      summa = 0
3      for i in s:
4          summa += int(i <= t)
5      return summa/len(s)

```

2.2.1. Дискретное распределение

На графиках представлены эмпирическая функция распределения для каждой выборки (синий цвет) и функция биномиального распределения (оранжевый цвет) с параметрами $n = 35$, $\theta = 0.5$.

▷ **N = 5:**

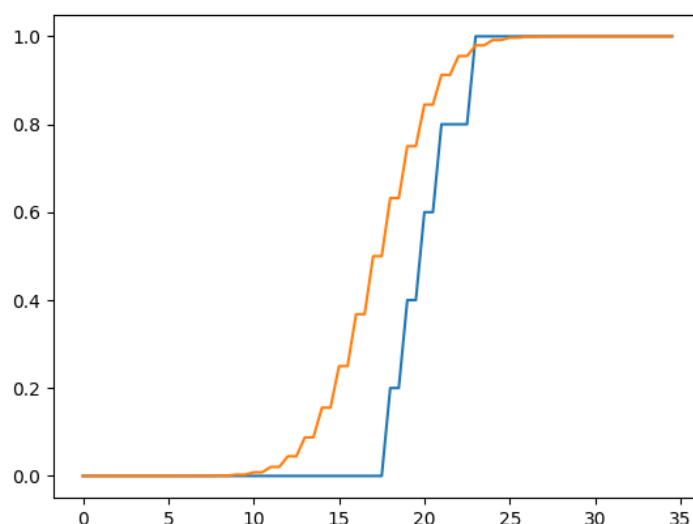


Рис. 2.17: Функция распределения при выборке и $n = 5$

▷ **N = 10:**

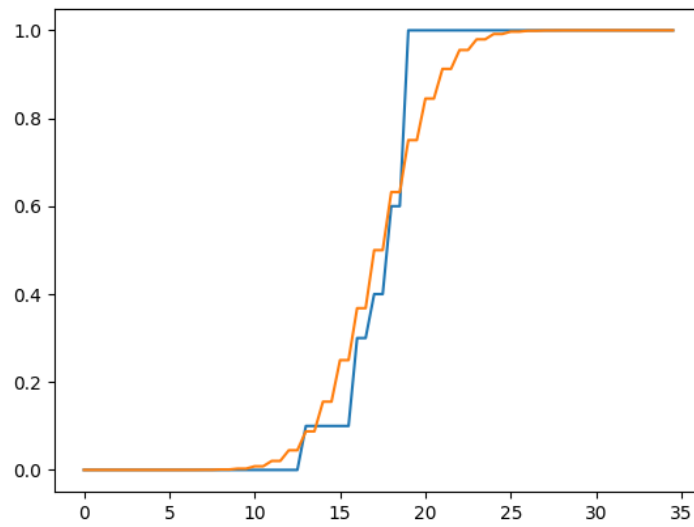


Рис. 2.18: Функция распределения при выборке и $n = 10$

▷ **N = 100:**

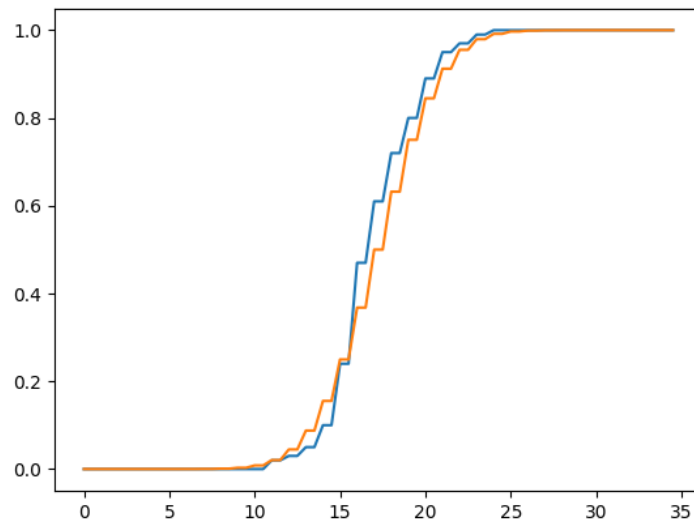


Рис. 2.19: Функция распределения при выборке и $n = 100$

▷ **N = 200:**

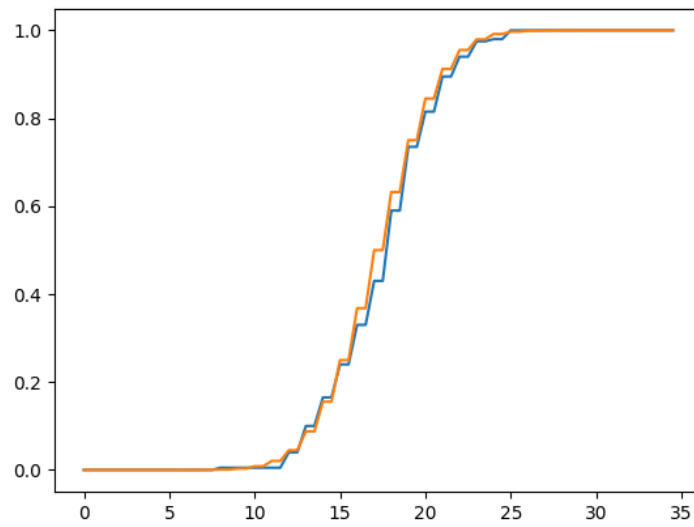


Рис. 2.20: Функция распределения при выборке и $n = 200$

▷ **N = 400:**

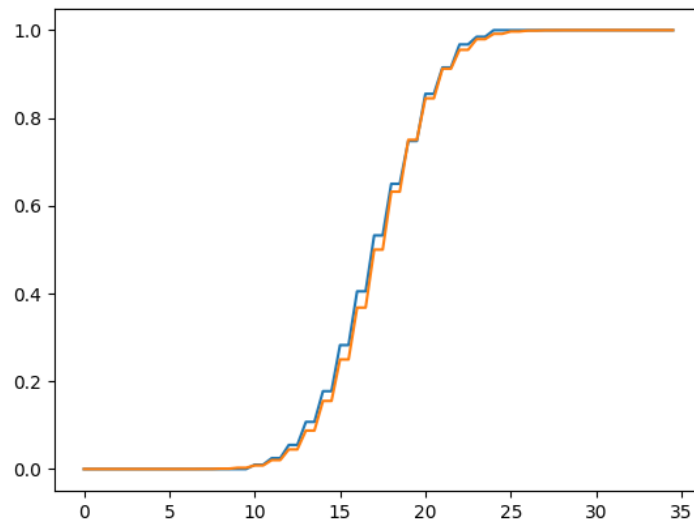


Рис. 2.21: Функция распределения при выборке и $n = 400$

▷ **N = 600:**

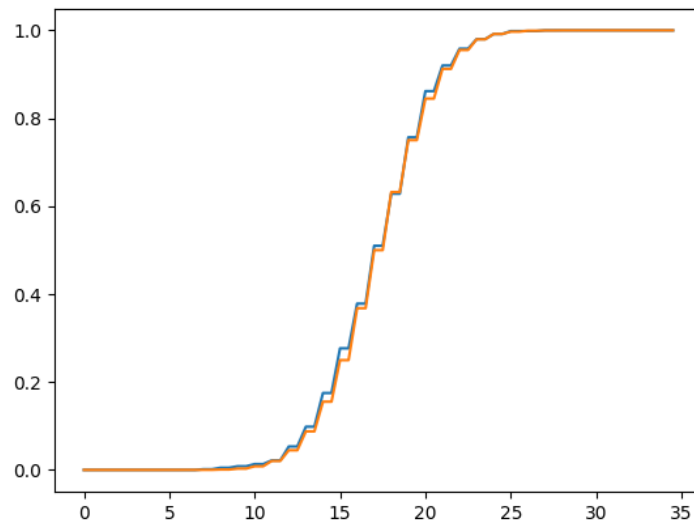


Рис. 2.22: Функция распределения при выборке и $n = 600$

▷ **N = 800:**

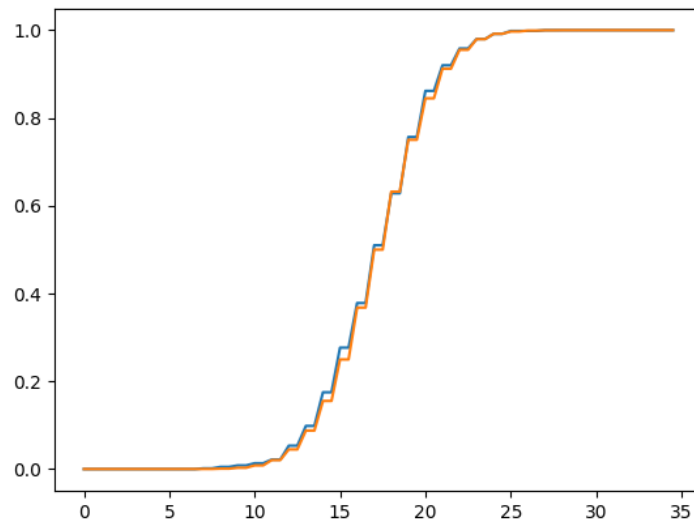


Рис. 2.23: Функция распределения при выборке и $n = 800$

▷ $N = 1000$:

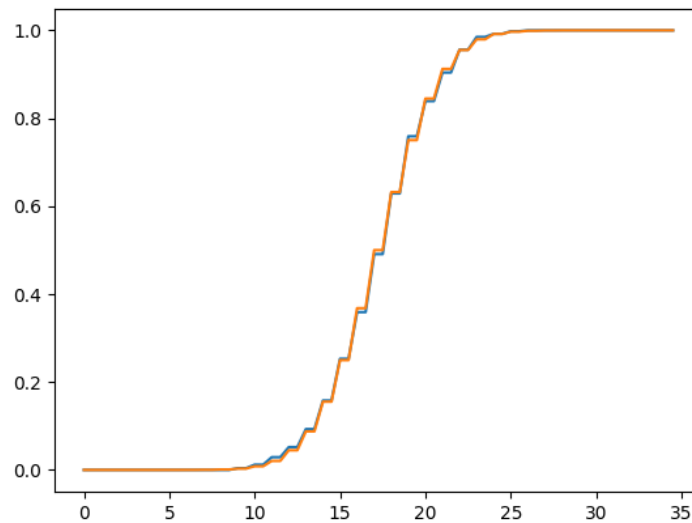


Рис. 2.24: Функция распределения при выборке и $n = 1000$

Проаниализировав приведённые графики, можно сделать вывод, что при увеличении объёма выборки график кумулятивной эмпирической функции распределения всё больше "стремится" к графику функции распределения случайной величины, что подтверждается теоремой:

Теорема: Для $\forall x \in \mathbb{R}$ и для $\forall \epsilon > 0$ при $n \rightarrow \infty$

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

(то есть $\hat{F}_n(x) \xrightarrow{P} F_n(x)$).

Иначе говоря, для произвольного фиксированного $y \in \mathbb{R}$ э.ф.р. $\hat{F}_n(y)$ с увеличением объёма выборки n стремиться к значению функции распределения $F(y)$.

2.2.2. Непрерывное распределение

На графиках представлены эмпирическая функция распределения для каждой выборки (синий цвет) и функция распределения Эрланга (оранжевый цвет) с параметрами $m = 8$, $\theta = 0.125$.

▷ **N = 5:**

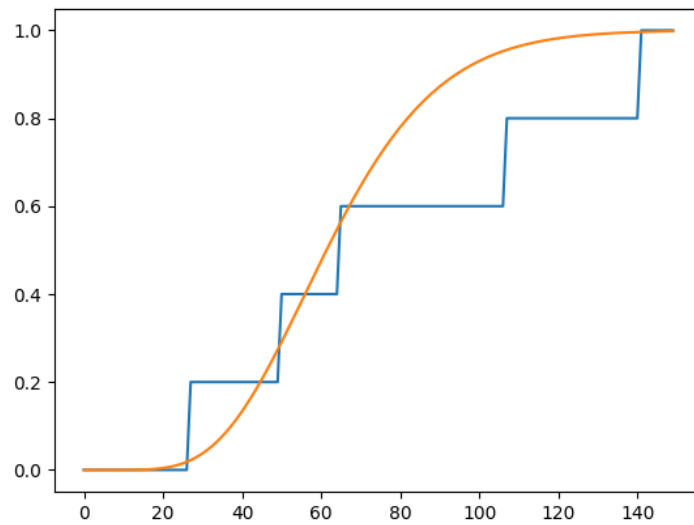


Рис. 2.25: Функция распределения при выборке и $n = 5$

▷ **N = 10:**

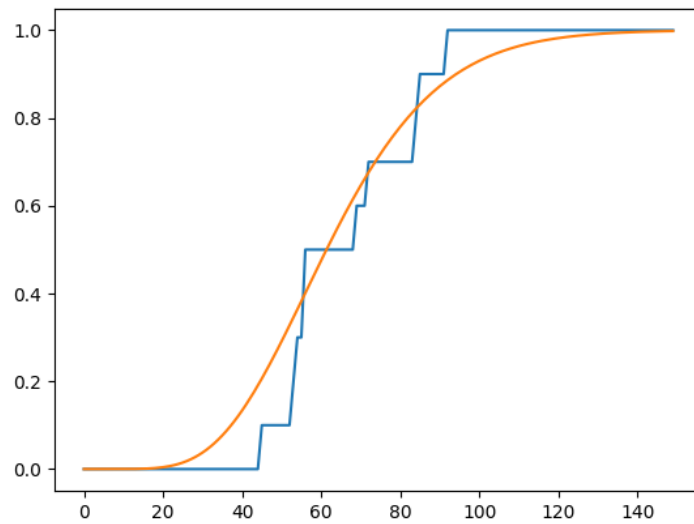


Рис. 2.26: Функция распределения при выборке и $n = 10$

▷ **N = 100:**

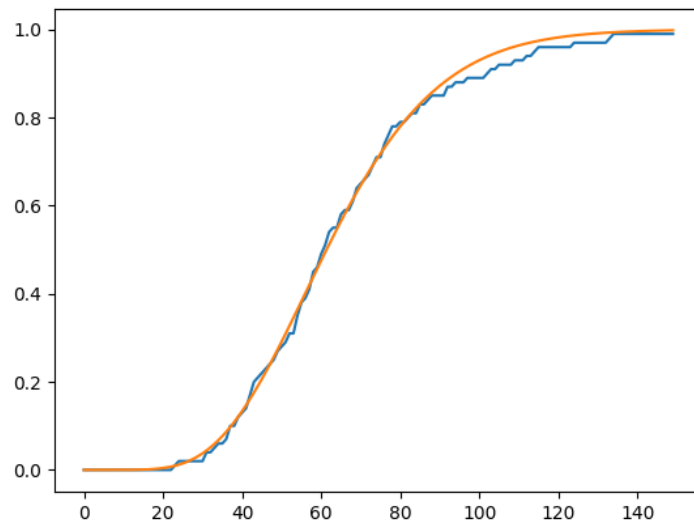


Рис. 2.27: Функция распределения при выборке и $n = 100$

▷ **N = 200:**

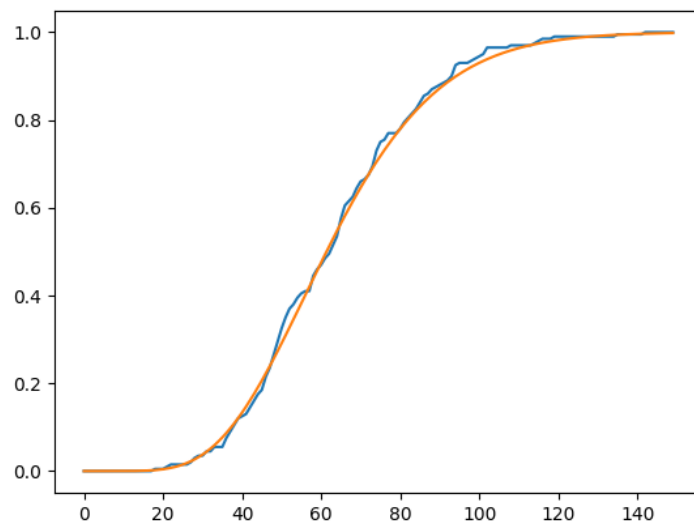


Рис. 2.28: Функция распределения при выборке и $n = 200$

▷ **N = 400:**

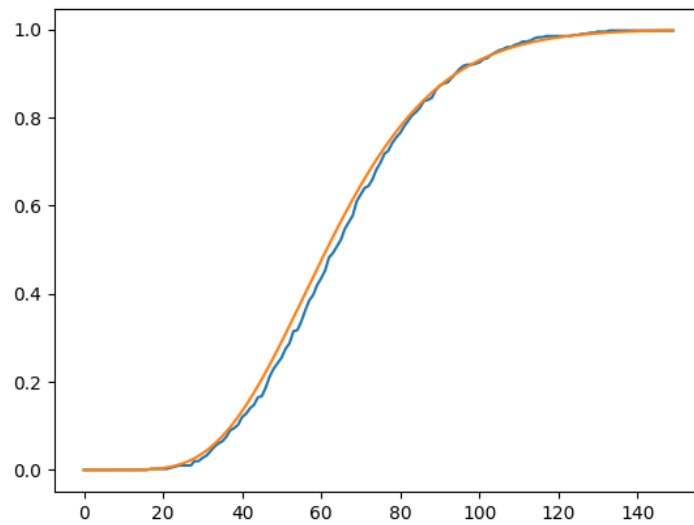


Рис. 2.29: Функция распределения при выборке и $n = 400$

▷ **N = 600:**

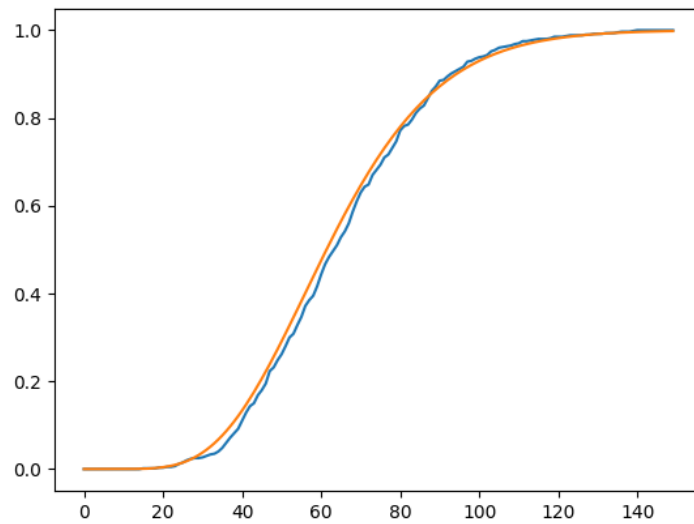


Рис. 2.30: Функция распределения при выборке и $n = 600$

▷ $N = 800$:

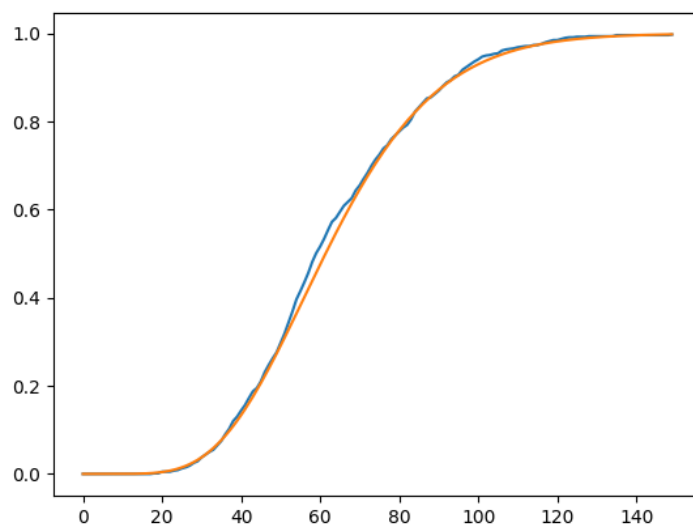


Рис. 2.31: Функция распределения при выборке и $n = 800$

▷ $N = 1000$:

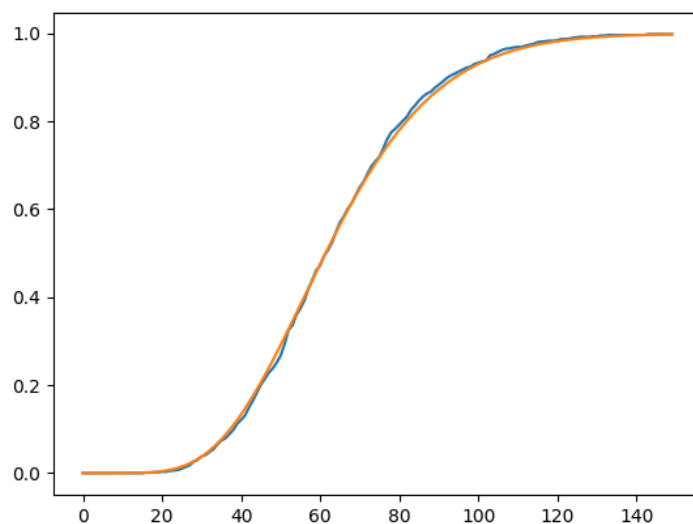


Рис. 2.32: Функция распределения при выборке и $n = 1000$

Проанализировав приведённые графики, можно сделать вывод, что при увеличении объёма выборки график кумулятивной эмпирической функции распределения всё больше "стремится" к графику функции распределения случайной величины, что подтверждается теоремой:

Теорема: Для $\forall x \in \mathbb{R}$ и для $\forall \epsilon > 0$ при $n \rightarrow \infty$

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

(то есть $\hat{F}_n(x) \xrightarrow{P} F_n(x)$).

Иначе говоря, для произвольного фиксированного $y \in \mathbb{R}$ э.ф.р. $\hat{F}_n(y)$ с увеличением объема выборки n стремится к значению функции распределения $F(y)$.

Для того, чтобы вычислить $D_{m,n} = \sqrt{\frac{nm}{m+n}} \sup_{x \in \mathbb{R}} |\mathbb{F}_n(x) - \mathbb{F}_m(x)|$ нужно написать функцию, которая будет реализовывать вычисление supremum по двум выборкам.

Листинг 2.2: $D_{m,n}$

```

1  def supremum(s1 , s2 ):
2      x = np.arange(0 , n , 0.5)
3      f1 = [custom_CDF(s1 , t) for t in x]
4      f2 = [custom_CDF(s2 , t) for t in x]
5      s = 0
6      for i in range(len(x)):
7          s = max(s , abs(f1[i] - f2[i]))
8      return s
9
10 def D(s1 , s2 ):
11     m = len(s1)
12     n = len(s2)
13
14     return math.sqrt((m*n)/(m+n))*supremum(s1 , s2)
```

2.2.3. Дискретное распределение

```

[[0.          1.09544512 1.33112913 0.94971112 1.18333333 1.13567252 1.09505169 1.09517485]
 [1.09544512 0.          0.63317382 0.81780775 0.7886775 0.76315423 0.86816999 0.75832671]
 [1.33112913 0.63317382 0.          1.46969385 0.69318107 0.9258201 1.11958574 1.13462048]
 [0.94971112 0.81780775 1.46969385 0.          1.18356805 0.9797959 0.77475803 0.78750661]
 [1.18333333 0.7886775 0.69318107 1.18356805 0.          0.41311822 0.67360968 0.77754191]
 [1.13567252 0.76315423 0.9258201 0.9797959 0.41311822 0.          0.71751058 0.45830303]
 [1.09505169 0.86816999 1.11958574 0.77475803 0.67360968 0.71751058 0.          0.74313525]
 [1.09517485 0.75832671 1.13462048 0.78750661 0.77754191 0.45830303 0.74313525 0.          ]]
```

Рис. 2.33: $D_{m,n}$ для Биномиального распределения

2.2.4. Непрерывное распределение

```
[[0.          0.73029674 0.82922798 0.76197753 0.71666667 0.79422849 0.79412145 0.73383407]
 [0.73029674 0.          0.42211588 0.52463139 0.40605178 0.45998337 0.56175705 0.42793541]
 [0.82922798 0.42211588 0.          0.57154761 0.55901699 0.57092239 0.57747054 0.59114681]
 [0.76197753 0.52463139 0.57154761 0.          0.57735027 1.08185797 0.63245553 1.00697567]
 [0.71666667 0.40605178 0.55901699 0.57735027 0.          0.59385745 0.95938348 0.60851106]
 [0.79422849 0.45998337 0.57092239 1.08185797 0.59385745 0.          0.9489656 0.54867264]
 [0.79412145 0.56175705 0.57747054 0.63245553 0.95938348 0.9489656 0.          1.30707477]
 [0.73383407 0.42793541 0.59114681 1.00697567 0.60851106 0.54867264 1.30707477 0.          ]]
```

Рис. 2.34: $D_{m,n}$ для Биномиального распределения

По полученным матрицам значений можно отметить, что они симметричны относительно главной диагонали.

3. Построение гистограммы и полигона частот

2.3.1. Дискретное распределение

▷ $N = 5$:

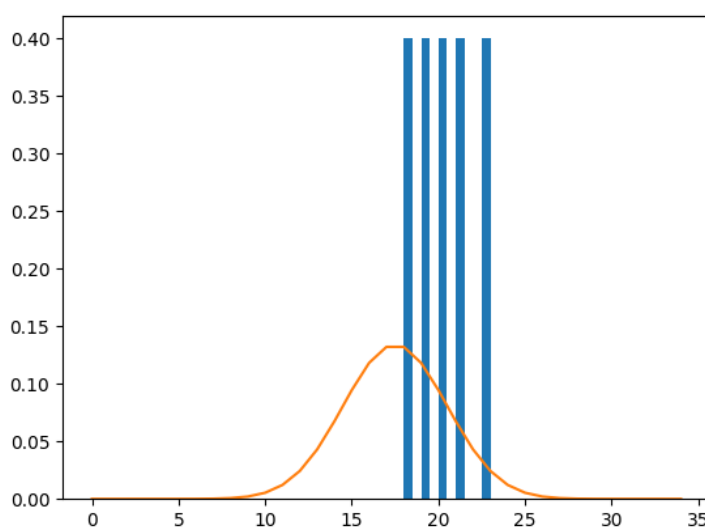


Рис. 2.35: Гистограмма при выборке и $n = 5$

▷ **N = 10:**

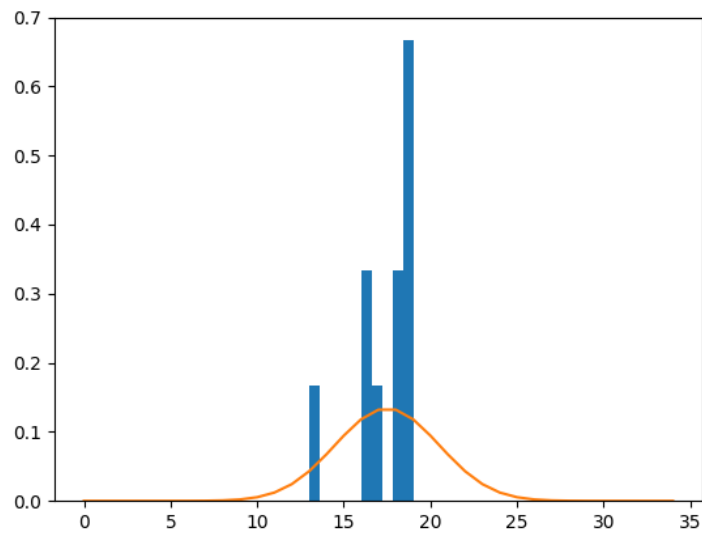


Рис. 2.36: Гистограмма при выборке и $n = 10$

▷ **N = 100:**

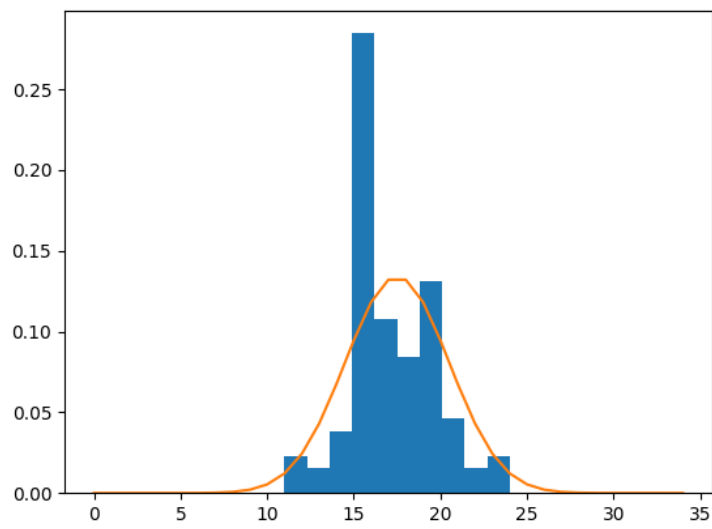


Рис. 2.37: Гистограмма при выборке и $n = 100$

▷ **N = 200:**

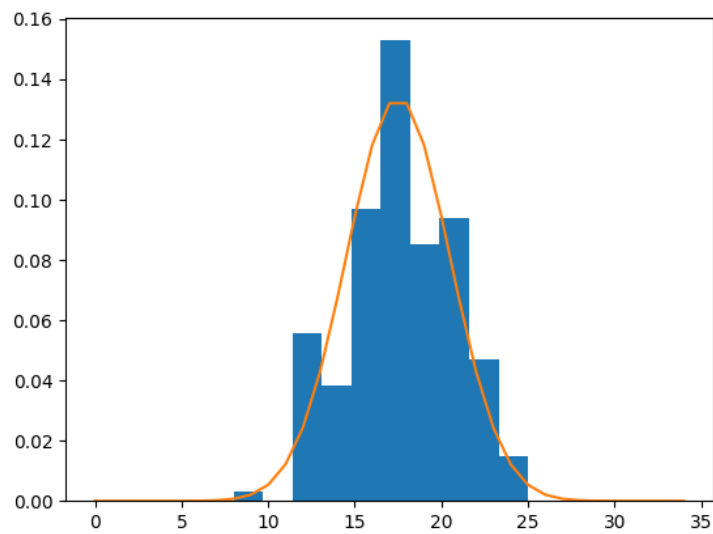


Рис. 2.38: Гистограмма при выборке и $n = 200$

▷ **N = 400:**

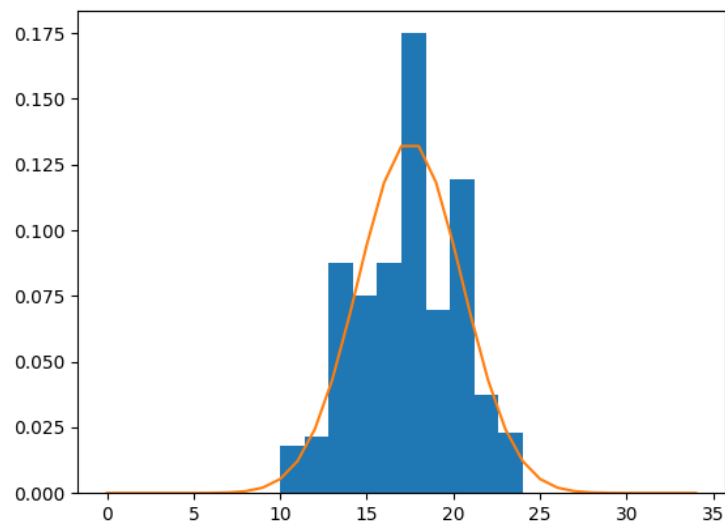


Рис. 2.39: Гистограмма при выборке и $n = 400$

▷ **N = 600:**

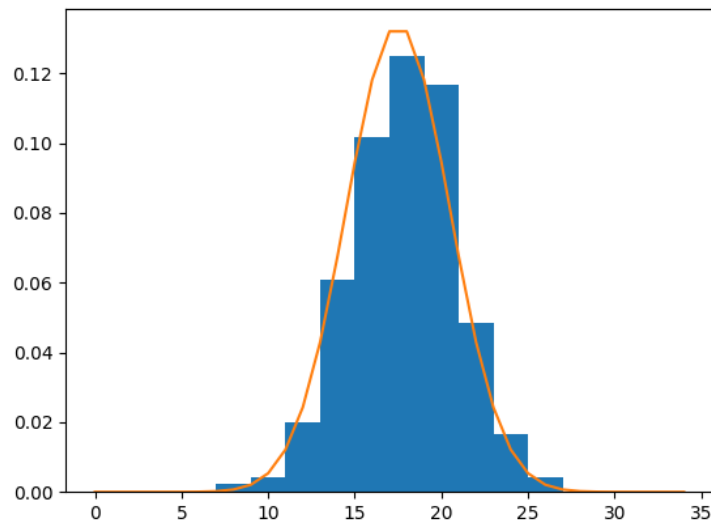


Рис. 2.40: Гистограмма при выборке и $n = 600$

▷ **N = 800:**

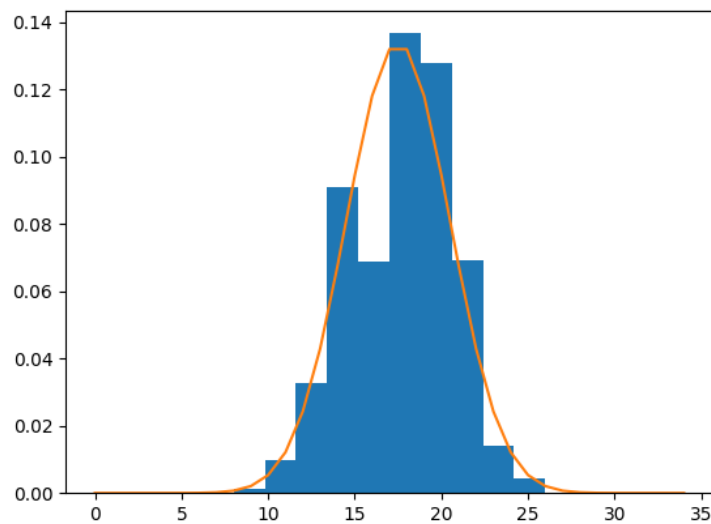


Рис. 2.41: Гистограмма при выборке и $n = 800$

▷ $N = 1000$:

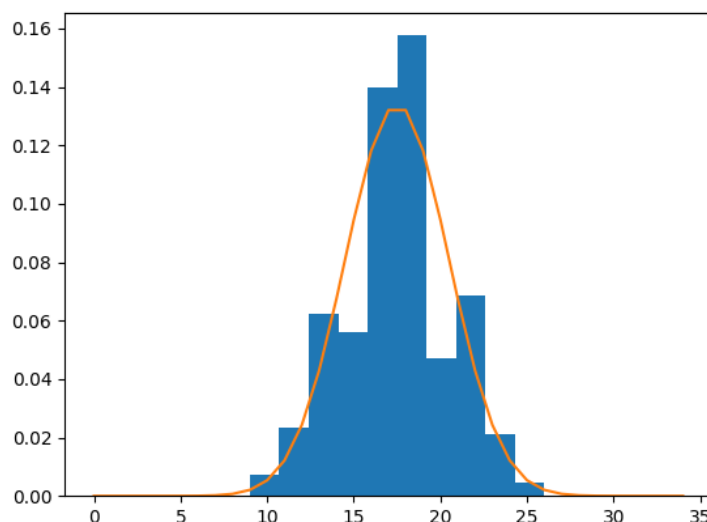


Рис. 2.42: Гистограмма при выборке и $n = 1000$

Из полученных графиков можно сделать вывод, можно сделать вывод, что с увеличением объёма выборки - её гистограмма частот "стремится" к функции вероятности. Это иллюстрирует следующую теорему:

Пусть распределение F абсолютно непрерывно, f - его истинная плотность. Пусть, кроме того, число k интервалов группировки не зависит от n .

Теорема:

При $n \rightarrow \infty \forall j = 1, \dots, k$

$$l_j \cdot f_j = \frac{v_j}{n} \xrightarrow{p} P(X_1 \in A_j) = \int_{A_j} f(x) d(x)$$

Предполагаемую область значений случайной величины ξ делят независимо от выборки на некоторое количество интервалов (не обязательно одинаковых). Пусть A_1, \dots, A_k - интервалы на прямой, называемые интервалами группировки. Обозначим для $j = 1, \dots, k$ через v_j число элементов выборки, попавших в интервал A_j . l_j - длина интервала A_j .¹

Данная теорема утверждает, что площадь столбца гистограммы, построенного над интервалом группировки, с ростом объёма выборки сближается с площадью области под графиком плотности над этим же интервалом.

¹Н. И. Чернова, "Лекции по математической статистике Нижегородский Государственный Университет: стр. 12

2.3.2. Непрерывное распределение

▷ $N = 5$:

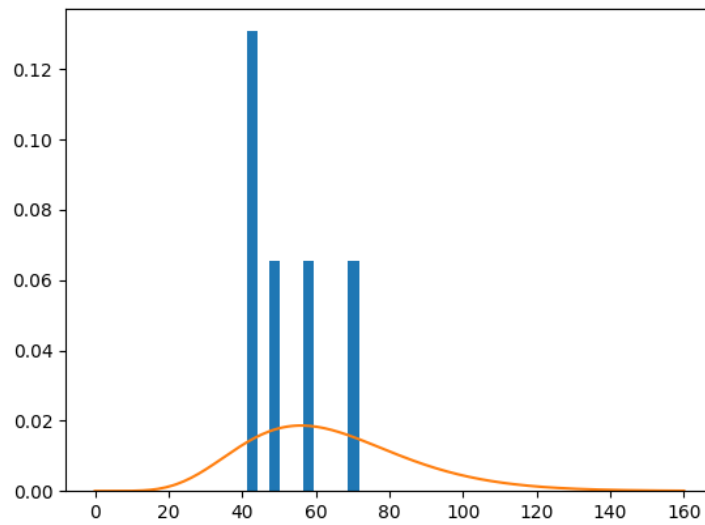


Рис. 2.43: Гистограмма при выборке и $n = 5$

▷ $N = 10$:

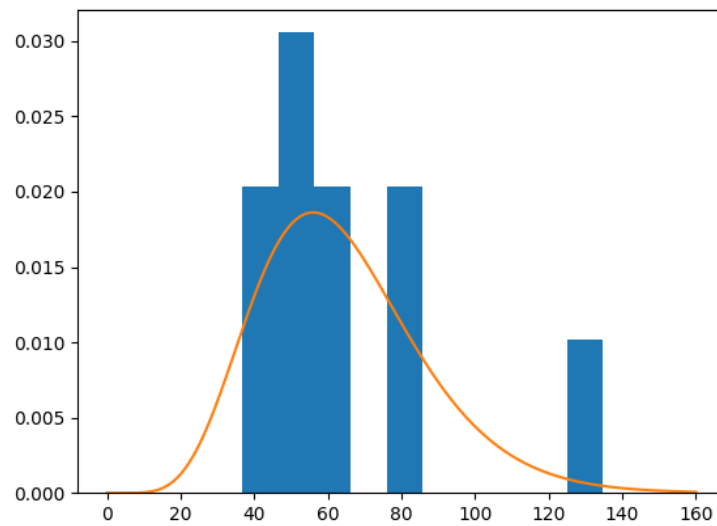


Рис. 2.44: Гистограмма при выборке и $n = 10$

▷ **N = 100:**

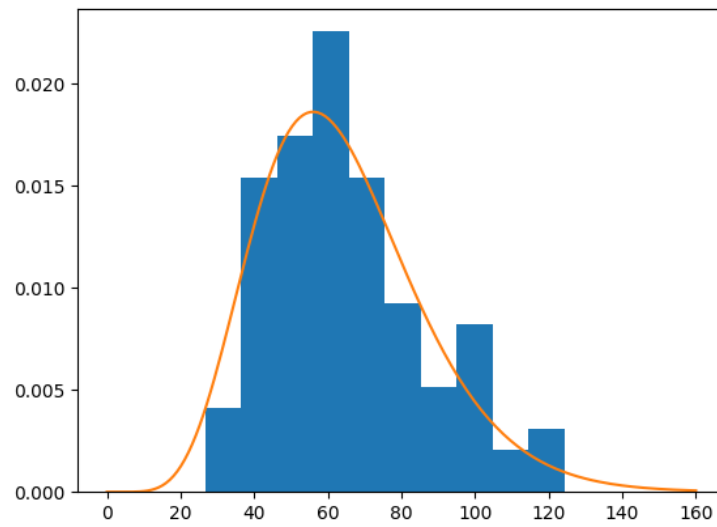


Рис. 2.45: Гистограмма при выборке и $n = 100$

▷ **N = 200:**

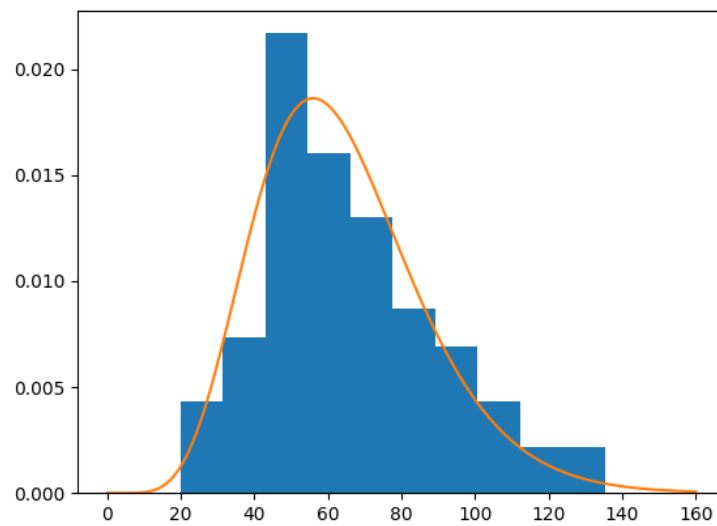


Рис. 2.46: Гистограмма при выборке и $n = 200$

▷ **N = 400:**

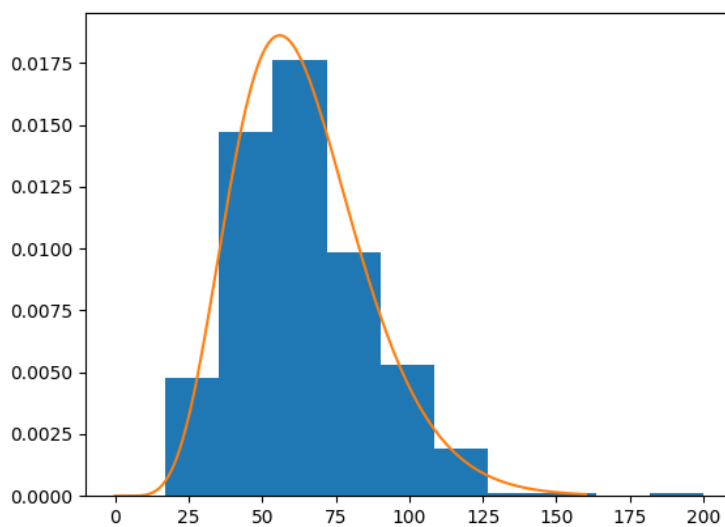


Рис. 2.47: Гистограмма при выборке и $n = 400$

▷ **N = 600:**

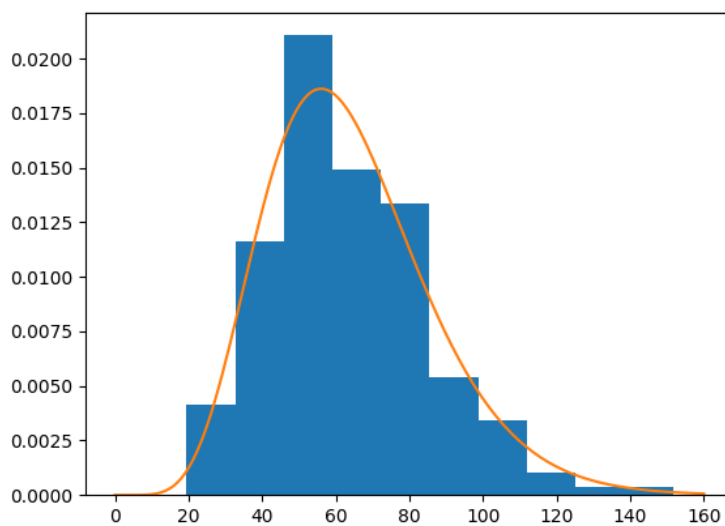


Рис. 2.48: Гистограмма при выборке и $n = 600$

▷ $N = 800$:

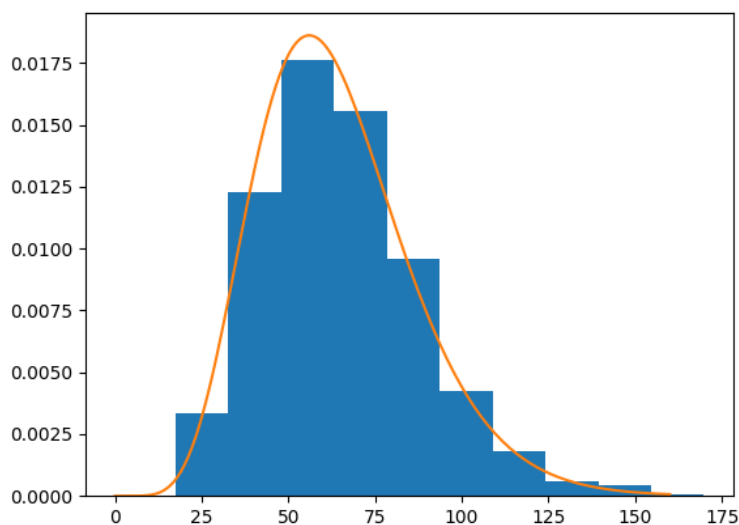


Рис. 2.49: Гистограмма при выборке и $n = 800$

▷ $N = 1000$:

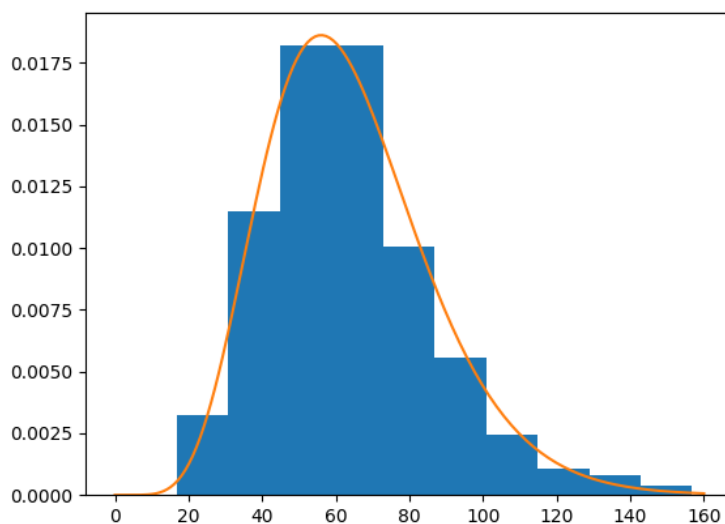


Рис. 2.50: Гистограмма при выборке и $n = 1000$

Из полученных графиков можно сделать вывод, можно сделать вывод, что с увеличением объёма выборки - её гистограмма частот "стремится" к функции вероятности. Это иллюстрирует следующую теорему:
Пусть распределение F абсолютно непрерывно, f - его истинная плотность.

Пусть, кроме того, число k интервалов группировки не зависит от n .

Теорема:

$$\text{При } n \rightarrow \infty \forall j = 1, \dots, k$$

$$l_j \cdot f_j = \frac{v_j}{n} \xrightarrow{p} P(X_1 \in A_j) = \int_{A_j} f(x) d(x)$$

Предполагаемую область значений случайной величины ξ делят независимо от выборки на некоторое количество интервалов (не обязательно одинаковых). Пусть A_1, \dots, A_k - интервалы на прямой, называемые интервалами группировки. Обозначим для $j = 1, \dots, k$ через v_j число элементов выборки, попавших в интервал A_j . l_j - длина интервала A_j .²

Данная теорема утверждает, что площадь столбца гистограммы, построенного над интервалом группировки, с ростом объема выборки сближается с площадью области под графиком плотности над этим же интервалом.

4. Вычисление выборочных моментов

Выборочное среднее является несмещенной оценкой для математического ожидания случайной величины ξ . При $n \rightarrow \infty$ значение выборочной дисперсии хоть и сходится к значению дисперсии случайной величины ξ , но является смещенной оценкой. Докажем это.

Доказательство несмещённости оценки математического ожидания

По определению: Смещением оценки $T(X)$ неизвестного параметра $\tau(\theta)$ будем называть величину

$$b(\theta) = M_\theta T(X) - \tau(\theta)$$

Если смещение равно 0, то оценка несмещенная.

Таким образом, в случае выборочного среднего, имеем:

$$\tau(\theta) = MX_i, i = 1, \dots, n$$

$$\bar{X} = T(X) = \frac{1}{n} \sum_{i=1}^n X_i, \text{ где } n - \text{объем выборки}$$

Посчитаем математическое ожидание выборочного среднего:

$$MT(X) = \frac{1}{n} M \sum_{i=1}^n X_i = \frac{1}{n} \sum_{i=1}^n MX_i = MX_i$$

Таким образом получаем $b(\theta) = M_\theta T(X) - \tau(\theta) = MX_i - MX_i = 0$. Значит оценка является несмещенной.

²Н. И. Чернова, "Лекции по математической статистике Нижегородский Государственный Университет: стр. 12

Доказательство состоятельности оценки математического ожидания

При доказательстве будем использовать утверждение, которое гласит, что для проверки состоятельности несмещенной оценки T_n для неизвестного параметра $\tau(\theta)$ достаточно убедиться, что ее дисперсия стремится к 0 при $n \rightarrow \infty$. А так же, свойство дисперсии о том, что $D(c \cdot \xi) = c^2 D\xi$.

$$DT(X) = D\left(\frac{1}{n} \sum_{i=1}^n MX_i\right) = \frac{1}{n^2} \left(\sum_{i=1}^n nDX_i\right) = \frac{1}{n} DX_i$$

$$\lim_{n \rightarrow \infty} DT(X) = DX_i \lim_{n \rightarrow \infty} \frac{1}{n} = 0, \text{ так как } \frac{1}{n} \rightarrow 0 \text{ при } n \rightarrow \infty$$

Таким образом, доказали несмещенность и состоятельность оценки.

Доказательство смещенности оценки дисперсии

$$\text{Пусть } Y_i = X_i - \alpha_1, \alpha_1 = M(\xi) \quad (1)$$

$$\text{Следовательно } \bar{Y} = \bar{X} - M\bar{X} \quad (2)$$

$$\begin{aligned} MT(X) &= \frac{1}{n} M \sum_{i=1}^n (Y_i - \bar{Y})^2 = /* \text{ по определению } \bar{Y} = \sum_{j=1}^n Y_j, \text{ поэтому } */ = \frac{1}{n} \cdot \\ &= \cdot M \sum_{i=1}^n (Y_i^2 + \bar{Y}^2 - 2\left(\frac{1}{n} \sum_{j=1}^n Y_j\right) Y_i) = \frac{1}{n} \left(\sum_{i=1}^n MY_i^2 + \sum_{i=1}^n M\bar{Y}^2 - \frac{2}{n} \sum_{i,j=1}^n M(Y_j Y_i) \right) \end{aligned}$$

Выделим следующее:

1. $MY_i = MX_i - M(MX_i) = /* \text{ мат ожидание - const, } M(\text{const}) = \text{const} */ = MX_i - MX_i = 0$, опираясь на (1)
2. $MY_i^2 = MY_i^2 - (MY_i)^2 + (MY_i)^2 = DY_i$, так как $MY_i = 0$
3. Если $i \neq j$, то верно следующее: $M(Y_i Y_j) = MY_i \cdot MY_j = 0$. Учитывая пункт 1., получим: $M(Y_i Y_j) = MY_i MY_j = 0$

$$4. M\bar{Y}^2 = \frac{1}{n^2} \sum_{i,j=1}^n M(Y_i Y_j) = \frac{1}{n^2} \left(\sum_{i \neq j} + \sum_{i=1}^n MY_i^2 \right) = /* \text{ учитывая пункты 3.}$$

$$\text{и 2., получим } */ = \frac{1}{n^2} \sum_{i=1}^n DY_i = \frac{1}{n} DY_i$$

Используя данные пункты, получим:

$$\begin{aligned} MT(X) &= \frac{1}{n} \left(\left(1 - \frac{2}{n}\right) \sum_{i=1}^n MY_i^2 + \sum_{i=1}^n nM\bar{Y}^2 \right) = \frac{1}{n} \left(\left(1 - \frac{2}{n}\right) nDY_i + DY_i \right) = \\ &= \frac{n-1}{n} DY_i = \frac{n-1}{n} D(X_i - MX_i) = \frac{n-1}{n} DX_i \end{aligned}$$

Значит оценка дисперсии смещенная.

Доказательство состоятельности оценки дисперсии

$$\begin{aligned} \text{По определению: } \overline{S}^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X})^2 = \frac{1}{n} \sum_{i=1}^n (X_i^2 + (\overline{X})^2 - 2X_i\overline{X}) = \frac{1}{n} \sum_{i=1}^n X_i^2 - \\ &- \frac{2}{n} \overline{X} \sum_{i=1}^n X_i + \frac{1}{n} (\overline{X})^2 \sum_{i=1}^n 1 = \overline{X^2} - 2\overline{X} \cdot \overline{X} + (\overline{X})^2 = \overline{X^2} - (\overline{X})^2 \end{aligned}$$

Статистика \overline{X} представляет из себя сумму независимых одинаково распределенных случайных величин, для которых определено математическое ожидание $M|X_1| < \infty$

Опишем ЗБЧ в форме Хинчина ³:

Пусть на одном пространстве элементарных событий задана последовательность независимых одинаково распределенных случайных величин X_1, X_2, \dots, X_n , для которых определено математическое ожидание. Тогда:

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{X_1 + \dots + X_n}{n} - M(X_1) \right| < \epsilon \right) = 1, \quad \forall \epsilon > 0$$

Тогда, опираясь на вышеописанные записи и Закон Больших Чисел в форме Хинчина, получим, что:

$$\overline{X} \xrightarrow[n \rightarrow \infty]{P} MX_i$$

Тогда:

$$\overline{X^2} - (\overline{X})^2 \xrightarrow[n \rightarrow \infty]{P} MX_i^2 - (MX_i)^2 = DX_i$$

Таким образом, оценка дисперсии состоятельна.

2.4.1. Дискретное распределение

Сначала вычислим истинные значения для математического ожидания - $35 \cdot 0.5 = 17.5$

дисперсии - $35 \cdot 0.5 \cdot (1 - 0.5) = 8.75$

³http://angtu.ru/universitet/kafedry-angtu/math/posobiya/metTV_ch2.pdf, .77

	Выборочное среднее	Выборочная дисперсии
5	20.20000	2.960000
10	17.40000	3.440000
100	17.16000	6.414400
200	17.74500	9.309975
400	17.28500	8.823775
600	17.36500	9.185108
800	17.56625	8.380611
1000	17.48700	8.975831

Рис. 2.51: Посчитанные значения для Биномиального распределения

Теперь рассчитаем абсолютную и относительную погрешности наших расчетов:

	Абс погрешность выборочного среднего	Абс погрешность выборочной дисперсии
5	2.70000	-5.790000
10	-0.10000	-5.310000
100	-0.34000	-2.335600
200	0.24500	0.559975
400	-0.21500	0.073775
600	-0.13500	0.435108
800	0.06625	-0.369389
1000	-0.01300	0.225831

Рис. 2.52: Абсолютная погрешность измерений

	Относ погрешность выборочного среднего (%)	Относ погрешность выборочной дисперсии (%)
5	15.428571	-66.171429
10	-0.571429	-60.685714
100	-1.942857	-26.692571
200	1.400000	6.399714
400	-1.228571	0.843143
600	-0.771429	4.972667
800	0.378571	-4.221589
1000	-0.074286	2.580926

Рис. 2.53: Относительная погрешность измерений

2.4.2. Непрерывное распределение

Сначала вычислим истинные значения для

математического ожидания - $\frac{8}{0.125} = 64$

дисперсии - $\frac{8}{0.125^2} = 512$

	Выборочное среднее	Выборочная дисперсии
5	52.804547	129.131136
10	66.638744	687.091986
100	64.908036	443.176304
200	65.813174	559.474058
400	63.999964	555.556690
600	63.207455	465.865459
800	65.182420	533.367746
1000	63.330723	514.540454

Рис. 2.54: Посчитанные значения для распределения Эрланга

Теперь рассчитаем абсолютную и относительную погрешности наших расчетов:

	Абс погрешность выборочного среднего	Абс погрешность выборочной дисперсии
5	-11.195453	-382.868864
10	2.638744	175.091986
100	0.908036	-68.823696
200	1.813174	47.474058
400	-0.000036	43.556690
600	-0.792545	-46.134541
800	1.182420	21.367746
1000	-0.669277	2.540454

Рис. 2.55: Абсолютная погрешность измерений

	Относ погрешность выборочного среднего (%)	Относ погрешность выборочной дисперсии (%)
5	-17.492895	-74.779075
10	4.123037	34.197653
100	1.418807	-13.442128
200	2.833084	9.272277
400	-0.000056	8.507166
600	-1.238352	-9.010653
800	1.847531	4.173388
1000	-1.045746	0.496182

Рис. 2.56: Относительная погрешность измерений

Домашнее задание 3.

Основные понятия математической статистики

1. Получение оценок

3.1.1. Дискретное распределение

3.1.1.1. Метод моментов

Пусть имеется выборка $X = (X_1, \dots, X_n)$ из распределения $\mathfrak{L}(\xi)$, $\mathfrak{L}(\xi) \in \mathcal{F} = F_\theta$, $\theta \in \Theta$, где $\theta = (\theta_1, \dots, \theta_r) \in \mathbb{R}^r$. Пусть у случайной величины ξ имеются первые r моментов, т.е. $\alpha_k = M_\theta(\xi^k) < \infty$, являющиеся функциями от неизвестного θ : $\alpha_k = \alpha_k(\theta)$, $k = \overline{1, r}$.

Рассмотрим систему:

$$\begin{cases} \alpha_k(\theta) = \hat{\alpha}_k, k = \overline{1, r} \end{cases}$$

в которой r неизвестных $\theta_1, \dots, \theta_r$. Эта система однозначно разрешима и ее решением являются $\hat{\theta}_i = \phi_i(\hat{\alpha}_1, \dots, \hat{\alpha}_r)$, ϕ_i — некоторая функция.

Оценки $\hat{\theta}_i$ будем называть оценками, построенными по методу моментов. Заметим, что если функция ϕ_i является непрерывной, то оценка $\hat{\theta}_i$ является состоятельной.

На практике для получения оценки параметра распределения методом моментов необходимо приравнять теоретические моменты соответствующим эмпирическим моментам того же порядка.

$$M\xi = n\theta, D\xi = n\theta(1 - \theta)$$

$$\overline{X} = \frac{1}{n} \sum_{i=1}^n X_i, \text{ где } X_i - \text{элемент из выборки, } S_X = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X})^2$$

$$M\xi = \overline{X}$$

$$n\theta = \frac{1}{n} \sum_{i=1}^n X_i$$

$$D\xi = S_X$$

$$n\theta(1 - \theta) = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X})^2$$

$$\text{Получаем: } 1 - \theta = \frac{S_X}{\overline{X}}. \text{ Значит } \hat{\theta} = 1 - \frac{S_X}{\overline{X}}$$

	Оценка theta
5	0.853465
10	0.802299
100	0.626200
200	0.475347
400	0.489513
600	0.471056
800	0.522914
1000	0.486714

Рис. 3.1: Оценка методом моментов

3.1.1.2. Метод максимального правдоподобия

Пусть есть выборка $X = (X_1, \dots, X_n)$ из распределения $L(\xi)$. Пусть $f_\theta(x)$ есть функция плотности случайной величины ξ , которая известна с точностью до параметра из распределения (в дискретном случае вместо функции плотности берем функцию вероятности $P_\theta(\xi = x)$). Пусть $\bar{x} = (x_1, \dots, x_n)$

Функцию, заданную равенством $L(\bar{x}, \theta) = \prod_{i=1}^n f_\theta(\xi_i)$ будем называть функцией правдоподобия.

Оценкой максимального правдоподобия называется построенная по реализации выборки \bar{x} значение $\hat{\theta} = \arg \max_{\theta \in \Theta} L(\bar{x}; \theta)$

То есть, чтобы найти оценку максимального правдоподобия (о.м.п.), надо найти такое значение θ , при котором функция правдоподобия принимает максимальное значение.

Если для каждого \bar{x} из выборочного пространства X максимум $L(\bar{x}; \theta)$ достигается в некоторой внутренней точке и $L(\bar{x}; \theta)$ дифференцируема по θ , то $\hat{\theta}(x)$ удовлетворяет условию

$$\frac{dL(\bar{x}; \theta)}{d\theta} = 0$$

Вместо функции правдоподобия для простоты часто рассматривают следующую функцию

$$\ln L(\bar{x}; \theta) = \sum_{i=1}^n \ln f_\theta(x_i)$$

Функцией максимального правдоподобия для биномиального распределения будет функция:

$$\begin{aligned} L(\bar{x}, \theta) &= \prod_{i=1}^N P\{X_1 = x_1\} \cdot \dots \cdot P\{X_N = x_N\} = \\ &= C_n^{x_1} \cdot \dots \cdot C_n^{x_N} \cdot \theta^{x_1 + \dots + x_N} (1 - \theta)^{nN - (x_1 + \dots + x_N)} \end{aligned}$$

$$, P\{X_i = x_i\} = C_n^{x_i} \theta^{x_i} (1 - \theta)^{n - x_i}.$$

Прологарифмируем левую и правую части равенства:

$$\begin{aligned} \ln(L(\bar{x}, \theta)) &= \sum_{i=1}^N \ln(P\{X_i = x_i\}) = \\ &= \sum_{i=1}^N (\ln(C_n^{x_i})) + (x_1 + \dots + x_N) \ln(\theta) + (nN - (x_1 + \dots + x_N)) \ln(1 - \theta) \end{aligned}$$

Чтобы найти оценки, необходимо приравнять производные этой функции по n и по θ к 0 и решить полученную систему уравнений. В случае неизвестного n получаем, что система неразрешима аналитически. Однако при известном n можно получить выражение для θ :

$$L'_\theta(\bar{x}, \theta) = \frac{N\bar{x}}{\theta} - \frac{n - N\bar{x}}{1 - \theta} = 0 \Rightarrow \hat{\theta} = \frac{\bar{x}}{n}$$

Оценка theta	
5	0.577143
10	0.497143
100	0.490286
200	0.507000
400	0.493857
600	0.496143
800	0.501893
1000	0.499629

Рис. 3.2: Оценка методом максимального правдоподобия

3.1.2. Непрерывное распределение

3.1.2.1. Метод моментов

Пусть имеется выборка $X = (X_1, \dots, X_n)$ из распределения $\mathfrak{L}(\xi)$, $\mathfrak{L}(\xi) \in \mathcal{F} = F_\theta$, $\theta \in \Theta$, где $\theta = (\theta_1, \dots, \theta_r) \in \mathbb{R}^r$. Пусть у случайной величины ξ имеются первые r моментов, т.е. $\alpha_k = M_\theta(\xi^k) < \infty$, являющиеся функциями от неизвестного θ : $\alpha_k = \alpha_k(\theta)$, $k = \overline{1, r}$.

Рассмотрим систему:

$$\begin{cases} \alpha_k(\theta) = \hat{\alpha}_k, k = \overline{1, r} \end{cases}$$

в которой r неизвестных $\theta_1, \dots, \theta_r$. Эта система однозначно разрешима и ее решением являются $\hat{\theta}_i = \phi_i(\hat{\alpha}_1, \dots, \hat{\alpha}_r)$, ϕ_i — некоторая функция.

Оценки $\hat{\theta}_i$ будем называть оценками, построенными по методу моментов. Заметим, что если функция ϕ_i является непрерывной, то оценка $\hat{\theta}_i$ является состоятельной.

На практике для получения оценки параметра распределения методом моментов необходимо приравнять теоретические моменты соответствующим эмпирическим моментам того же порядка.

$$M\xi = \frac{m}{\theta}, D\xi = \frac{m}{\theta^2}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \text{ где } X_i - \text{элемент из выборки, } S_X = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$M\xi = \bar{X}$$

$$\frac{m}{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$D\xi = S_X$$

$$\frac{m}{\theta^2} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Получаем: $\theta = \frac{\bar{X}}{S_X}$. Значит $\hat{\theta} = \frac{\bar{X}}{S_X}$

	Оценка theta
5	0.270099
10	0.138715
100	0.146929
200	0.128979
400	0.131239
600	0.123081
800	0.116740
1000	0.116787

Рис. 3.3: Оценка методом моментов

3.1.2.2. Метод максимального правдоподобия

Пусть есть выборка $X = (X_1, \dots, X_n)$ из распределения $L(\xi)$. Пусть $f_\theta(x)$ есть функция плотности случайной величины ξ , которая известна с точностью до параметра из распределения (в дискретном случае вместо функции плотности берем функцию вероятности $P_\theta(\xi = x)$). Пусть $\bar{x} = (x_1, \dots, x_n)$

Функцию, заданную равенством $L(\bar{x}, \theta) = \prod_{i=1}^n f_\theta(\xi_i)$ будем называть функцией правдоподобия.

Оценкой максимального правдоподобия называется построенная по реализации выборки \bar{x} значение $\hat{\theta} = \arg \max_{\theta \in \Theta} L(\bar{x}; \theta)$

То есть, чтобы найти оценку максимального правдоподобия (о.м.п.), надо найти такое значение θ , при котором функция правдоподобия принимает максимальное значение.

Если для каждого \bar{x} из выборочного пространства X максимум $L(\bar{x}; \theta)$ достигается в некоторой внутренней точке и $L(\bar{x}; \theta)$ дифференцируема по θ , то $\hat{\theta}(x)$ удовлетворяет условию

$$\frac{dL(\bar{x}; \theta)}{d\theta} = 0$$

Вместо функции правдоподобия для простоты часто рассматривают следующую функцию

$$\ln L(\bar{x}; \theta) = \sum_{i=1}^n \ln f_\theta(x_i)$$

Функцией максимального правдоподобия для распределения Эрланга будет функция:

$$L(\bar{x}, \theta) = \prod_{i=1}^n \frac{\theta^m}{(m-1)!} x_i^{m-1} \exp(-\theta x_i) = \frac{\theta^m}{(m-1)!} x_1^{m-1} \exp(-\theta x_1) * \\ * \frac{\theta^m}{(m-1)!} x_2^{m-1} \exp(-\theta x_2) * \dots * \frac{\theta^m}{(m-1)!} x_n^{m-1} \exp(-\theta x_n)$$

Прологарифмируем левую и правую части равенства:

$$\ln(L(\bar{x}, \theta)) = \ln\left(\prod_{i=1}^n \frac{\theta^m}{(m-1)!} x_i^{m-1} \exp(-\theta x_i)\right) = \ln\left(\frac{\theta^m}{(m-1)!} x_1^{m-1} \exp(-\theta x_1) * \right. \\ \left. * \frac{\theta^m}{(m-1)!} x_2^{m-1} \exp(-\theta x_2) * \dots * \frac{\theta^m}{(m-1)!} x_n^{m-1} \exp(-\theta x_n)\right)$$

Чтобы получить оценку, приравняем производную по θ к 0:

$$L'_\theta(\bar{x}, \theta) = \frac{nm}{\theta} - (x_1 + x_2 + \dots + x_n) = 0 \Rightarrow \theta = \frac{nm}{(x_1 + \dots + x_n)} = \frac{m}{\bar{x}}$$

Оценка theta	
5	0.127771
10	0.165732
100	0.124064
200	0.124831
400	0.127439
600	0.124318
800	0.124440
1000	0.123426

Рис. 3.4: Оценка методом максимального правдоподобия

2. Поиск оптимальных оценок

3.2.1. Дискретное распределение

Теорема Рао-Блекуэлла-Колмогорова:

Оптимальная оценка, если существует, является функцией от достаточной

статистики.

Теорема о полноте семейства экспоненциальных функций¹:

Пусть $F = F_\theta, \theta \in \Theta$ - экспоненциальное семейство и ф-ия $A(\theta)$ и параметрическое пространство Θ таковы, что $A(\theta)$ содержит некоторый отрезок, когда θ пробегает множество Θ . Тогда $T(X) = \sum_{i=1}^n B(X_i)$ является полной и достаточной статистикой.

Теорема:

Если \exists полная достаточная статистика, то произвольная функция от нее будет являться оптимальной оценкой своего математического ожидания.

Покажем, что биномиальное распределение принадлежит экспоненциальному семейству.

$$P(\xi = k) = \binom{n}{k} \theta^k (1 - \theta)^{n-k} = \binom{n}{k} \left(\frac{\theta}{1 - \theta} \right)^k (1 - \theta)^n = \\ = \exp \left(\ln \binom{n}{k} + k \ln \frac{\theta}{1 - \theta} + n \ln(1 - \theta) \right)$$

$A(\theta) = \ln \frac{\theta}{1 - \theta}, B(k) = k, C(\theta) = n \ln(1 - \theta), D(k) = \ln \binom{n}{k}$. По теореме о полноте экспоненциального семейства, оценка $T(X) = \sum_{i=1}^N k_i$ является полной и достаточной статистикой. Обозначим за $\chi = \sum_{i=1}^N k_i \sim Bi(nN, \theta)$, где N - объем выборки.

Найдем функцию H , для которой $MH = \theta$.

$$\sum_{i=0}^{nN} H(i) \binom{i}{nN} \theta^i (1 - \theta)^{nN - i} = \theta \\ \sum_{i=0}^{nN} H(i) \binom{i}{nN} \theta^{i-1} (1 - \theta)^{nN - i} = 1 \\ /* (\theta + 1 - \theta)^{nN-1} = \sum_{i=0}^{nN-1} \binom{nN-1}{i} \theta^i (1 - \theta)^{nN-i-1} = 1 */ \quad (3)$$

Пусть $H(i) = 0$ при $i = 0$, тогда

$$\sum_{i=1}^{nN} H(i) \binom{nN}{i} \theta^{i-1} (1 - \theta)^{nN-i} = 1$$

¹<https://teach-in.ru/file/synopsis/pdf/math-statistics-lectures-shabanov-M1.pdf?ysclid=lb9jhpads580207804>, стр. 118

$$\sum_{i=0}^{nN-1} H(i+1) \binom{nN}{i+1} \theta^i (1-\theta)^{nN-i-1} = 1$$

Из уравнения выше и (3) видно, что нужно решить относительно H уравнение

$$H(i+1) \binom{nN}{i+1} = \binom{nN-1}{i}$$

$$H(i+1) \frac{(nN)!}{(i+1)!(nN-i-1)!} = \frac{(nN-1)!}{i!(nN-i-1)!}$$

$$h(i+1) \frac{nN}{i+1} = 1$$

$$h(i+1) = \frac{i+1}{nN}$$

Получаем $H(x) = \frac{x}{nN} = \frac{\bar{x}}{n}$, где $x = \sum_{i=1}^N x_i$

Оценка theta	
5	0.577143
10	0.497143
100	0.490286
200	0.507000
400	0.493857
600	0.496143
800	0.501893
1000	0.499629

Рис. 3.5: Оптимальные оценки

3.2.2. Непрерывное распределение

Теорема Рао-Блекуэлла-Колмогорова:

Оптимальная оценка, если существует, является функцией от достаточной статистики.

Теорема о полноте семейства экспоненциальных функций²:

Пусть $F = F_\theta$, $\theta \in \Theta$ - экспоненциальное семейство и ф-ия $A(\theta)$ и параметрическое пространство Θ таковы, что $A(\theta)$ содержит некоторый отрезок, когда

²<https://teach-in.ru/file/synopsis/pdf/math-statistics-lectures-shabanov-M1.pdf?ysclid=lb9jhpads580207804>, стр. 118

θ пробегает множество Θ . Тогда $T(X) = \sum_{i=1}^n B(X_i)$ является полной и достаточной статистикой.

Теорема:

Если \exists полная достаточная статистика, то произвольная функция от нее будет являться оптимальной оценкой своего математического ожидания.

Покажем, что распределение Эрланга принадлежит экспоненциальному семейству.

$$f(x) = \frac{\theta^m}{(m-1)!} x^{m-1} e^{-\theta x} = \exp\left(-\theta x + (m-1)\ln(x) + m\ln(\theta) - \ln[(m-1)!]\right)$$

$$A(\theta) = -\theta, B(x) = x, C(\theta) = m\ln(\theta), D(x) = (m-1)\ln(x) - \ln[(m-1)!].$$

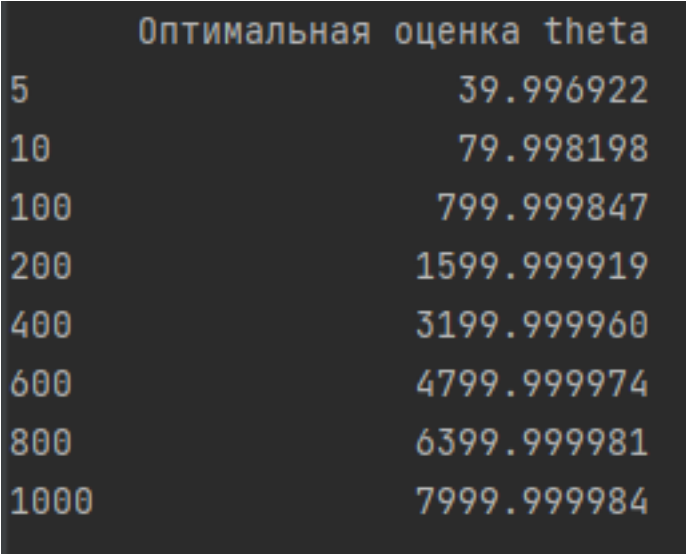
По теореме о полноте экспоненциального семейства, оценка $T(X) = \sum_{i=1}^N x_i$ является полной и достаточной статистикой. Обозначим за $\chi = \sum_{i=1}^N x_i \sim \text{Erlang}(mN, \theta)$, где N - объем выборки.

$$\begin{aligned} \text{Посчитаем математическое ожидание } M\left(\frac{1}{x}\right) &= \int_0^{\infty} \frac{\theta^m x^{m-1} e^{-\theta x}}{(m-1)!} \left(\frac{1}{x}\right) dx = \\ &= \frac{\theta}{m-1} \int_0^{\infty} \frac{\theta^{m-1} x^{m-2} e^{-\theta x}}{(m-2)!} dx = \frac{\theta}{m-1}, \text{ так как } \int_0^{\infty} \frac{\theta^{m-1} x^{m-2} e^{-\theta x}}{(m-2)!} dx = 1 \end{aligned}$$

Пусть оценка параметра $\theta : T(X) = \frac{8N-1}{\chi}$. Проверим ее несмещенность

$$M\left(\frac{8N-1}{\chi}\right) = (8N-1)M\left(\frac{1}{\chi}\right) = (8N-1)\frac{\theta}{Nm-1}$$

Так как нам известен параметр m , равный 8, то подставляя его, получаем что $M\left(\frac{8N-1}{\chi}\right) = \theta$. Значит $T(X) = \frac{8N-1}{\chi}$ - несмещенная оценка. По теореме о полноте экспоненциальных семейств данная оценка является полной и достаточной статистикой. Оценка $T(X) = \frac{8N-1}{\sum_{i=1}^N x_i}$ - оптимальная оценка.



A screenshot of a terminal window with a dark background and light-colored text. The text is organized into a table with two columns. The first column lists sample sizes (5, 10, 100, 200, 400, 600, 800, 1000) and the second column lists the corresponding optimal theta values. The title 'Оптимальная оценка theta' is centered at the top of the table.

Оптимальная оценка theta	
5	39.996922
10	79.998198
100	799.999847
200	1599.999919
400	3199.999960
600	4799.999974
800	6399.999981
1000	7999.999984

Рис. 3.6: Оптимальные оценки

Домашнее задание 4.

Проверка статистических гипотез

1. Проверка гипотезы о виде распределения

4.1.1. Критерий согласия Колмогорова (Смирнова)

Критерий Колмогорова основан на теореме Колмогорова. Статистика критерия определяется формулой:

$$D_n = D_n(X) = \sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F(x) \right|,$$

где D_n — это отклонение эмпирической функции распределения от теоретической функции распределения.

Мы знаем, что \hat{F}_n является оптимальной, несмещенной и состоятельной оценкой для $F(x)$. Отсюда следует, что D_n не должно «сильно» отклоняться от 0. Отсюда следует, что критическую область критерия, основанного на статистике $T = D_n$, следует задавать в виде $\tau_\alpha = \{t \geq t_\alpha\}$, т.е. большие значения D_n надо интерпретировать как свидетельство против проверяемой гипотезы H_0 . Критическая граница t_α при заданном уровне значимости α рассчитывается при этом на основании теоремы Колмогорова. Положив $t_\alpha = \lambda_\alpha / \sqrt{n}$, где $K(\lambda_\alpha) = 1 - \alpha$, будем иметь

$$P(D_n \in \tau_\alpha | H_0) = P(\lambda_\alpha | H_0 \leq \sqrt{n} D_n) \approx 1 - K(\lambda_\alpha) = \alpha$$

Тем самым критерий согласия Колмогорова формулируется следующим образом: при выбранном уровне значимости α число λ_α определено соотношением $K(\lambda_\alpha) = 1 - \alpha$, то H_0

$$H_0 \iff \sqrt{n} D_n < \lambda_\alpha$$

Вместо статистики $\sqrt{n} D_n \geq \lambda_\alpha$ при малых значениях n ($n \geq 20$) рекомендуется использовать статистику

$$S_n = \frac{6nD_n + 1}{6\sqrt{n}}$$

Проверим данные гипотезы для всех имеющихся выборок на уровнях значимости 10%, 5% и 1% (критические значения распределения Колмогорова взяты с сайта http://smc.edu.nstu.ru/krit_kolm.htm и равны 1.22, 1.36 и 1.63 для уровней значимости 10%, 5% и 1% соответственно). Данные уровни значимости будут применены для всех согласий.

4.1.1.1. Дискретное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	0.558525	H ₀ принимается	H ₀ принимается	H ₀ принимается
10	0.475038	H ₀ принимается	H ₀ принимается	H ₀ принимается
100	0.900000	H ₀ принимается	H ₀ принимается	H ₀ принимается
200	0.989949	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	0.745598	H ₀ принимается	H ₀ принимается	H ₀ принимается
600	0.530723	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	0.748687	H ₀ принимается	H ₀ принимается	H ₀ принимается
1000	0.567294	H ₀ принимается	H ₀ принимается	H ₀ принимается

Рис. 4.1: Критерий согласия Колмогорова (Смирнова)

4.1.2. Непрерывное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	0.638489	H ₀ принимается	H ₀ принимается	H ₀ принимается
10	1.059877	H ₀ принимается	H ₀ принимается	H ₀ принимается
100	1.173164	H ₀ принимается	H ₀ принимается	H ₀ принимается
200	0.710442	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	1.423843	H ₀ отвергается	H ₀ отвергается	H ₀ принимается
600	0.880293	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	1.237535	H ₀ отвергается	H ₀ принимается	H ₀ принимается
1000	1.358143	H ₀ отвергается	H ₀ принимается	H ₀ принимается

Рис. 4.2: Критерий согласия Колмогорова (Смирнова)

4.1.3. Критерий согласия хи-квадрат

Пусть случайные величины $\xi_1, \xi_2, \dots, \xi_n$ независимы и имеют стандартное нормальное распределение, тогда случайная величина:

$$\chi_n^2 = \sum_{i=1}^n \xi_i^2$$

имеет распределение, которое называется хи-квадрат с n степенями свободы. Пусть по наблюдению вектора частот $\bar{v} = (v_1, \dots, v_N)$ требуется проверить простую гипотезу H_0 , $\bar{p} = \bar{p}^\circ$ $\bar{p}^\circ = (p_1^\circ, \dots, p_N^\circ)$ - заданный вероятностный вектор ($0 < p_j^\circ < 1$, $j = 1, \dots, N$, $p_1^\circ + \dots + p_N^\circ = 1$). К. Пирсон в 1900 г.

предложил использовать в качестве меры отклонения эмпирических данных от гипотетических значений \bar{p}° меру хи-квадрат

$$\overset{\circ}{X}_n^2 = \overset{\circ}{X}_n^2(\underline{v}) = \sum_{j=1}^N \frac{(v_j - np_j^\circ)^2}{np_j^\circ}$$

Таким образом, классический критерий χ^2 имеет следующий вид: пусть есть выборка объемом n и наблюдавшиеся значения вектора частот $\bar{v} = (v_1, \dots, v_N)$; тогда при заданном уровне значимости α

$$H_0 \iff \overset{\circ}{X}_n^2 \leq \chi_{1-\alpha, N-1}^2$$

4.1.3.1. Дискретное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	10.176205	H ₀ отвергается	H ₀ отвергается	H ₀ принимается
10	5.126385	H ₀ принимается	H ₀ принимается	H ₀ принимается
100	61.588071	H ₀ отвергается	H ₀ отвергается	H ₀ отвергается
200	8.952434	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	15.564610	H ₀ принимается	H ₀ принимается	H ₀ принимается
600	9.490396	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	17.691660	H ₀ принимается	H ₀ принимается	H ₀ принимается
1000	12.017492	H ₀ принимается	H ₀ принимается	H ₀ принимается

Рис. 4.3: Критерий согласия хи-квадрат

4.1.3.2. Непрерывное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	85.954668	H ₀ отвергается	H ₀ отвергается	H ₀ отвергается
10	75.128968	H ₀ отвергается	H ₀ отвергается	H ₀ отвергается
100	41.657762	H ₀ принимается	H ₀ принимается	H ₀ принимается
200	81.830707	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	50.369438	H ₀ принимается	H ₀ принимается	H ₀ принимается
600	44.052301	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	42.352177	H ₀ принимается	H ₀ принимается	H ₀ принимается
1000	62.563067	H ₀ принимается	H ₀ принимается	H ₀ принимается

Рис. 4.4: Критерий согласия хи-квадрат

Из-за того, что для некоторых выборок в Биноминальном распределении и распределении Эрланга гипотеза H_0 отвергается, то можно сделать

несколько общих замечаний по критерию согласия хи-квадрат. Если исходные данные представляют собой выборку из некоторого непрерывного распределения, то, применяя предварительно метод группировки наблюдений, приходят к рассмотрению дискретной схемы, в которых в качестве событий A_j рассматриваются события $\{\xi \in \epsilon_j\}$, где $\epsilon_1, \dots, \epsilon_N$ - интервалы группировки.

Недостатком метода, является то, что при группировке данных происходит некоторая потеря информации. Кроме того, остается еще вопрос о выборе числа интервалов N и их виде.

4.1.4. Критерий согласия Колмогорова (Смирнова) для сложной гипотезы (в условиях когда неизвестен параметр распределения)

Описанную выше методику проверки гипотезы о виде распределения наблюдаемой случайной величины можно распространить и на случай сложной гипотезы H_0 . В этом случае используют тестовую статистику

$$\hat{D}_n = \sup_{-\infty < x < \infty} \left| \hat{F}_n(x) - F(x; \hat{\theta}_n) \right|$$

где $\hat{\theta}_n$ - оценка максимального правдоподобия параметра θ_n .

Для проверки сложных гипотез будем брать оценки максимального правдоподобия из ДЗ 3.

4.1.4.1. Дискретное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	0.465076	H_0 принимается	H_0 принимается	H_0 принимается
10	0.407976	H_0 принимается	H_0 принимается	H_0 принимается
100	0.672440	H_0 принимается	H_0 принимается	H_0 принимается
200	0.218776	H_0 принимается	H_0 принимается	H_0 принимается
400	0.489914	H_0 принимается	H_0 принимается	H_0 принимается
600	0.340746	H_0 принимается	H_0 принимается	H_0 принимается
800	0.469235	H_0 принимается	H_0 принимается	H_0 принимается
1000	0.532047	H_0 принимается	H_0 принимается	H_0 принимается

Рис. 4.5: Критерий согласия Колмогорова (Смирнова) для сложной гипотезы

4.1.4.2. Непрерывное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	0.571272	H ₀ принимается	H ₀ принимается	H ₀ принимается
10	0.894187	H ₀ принимается	H ₀ принимается	H ₀ принимается
100	0.502443	H ₀ принимается	H ₀ принимается	H ₀ принимается
200	0.898623	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	0.632644	H ₀ принимается	H ₀ принимается	H ₀ принимается
600	0.541355	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	0.482958	H ₀ принимается	H ₀ принимается	H ₀ принимается
1000	0.636729	H ₀ принимается	H ₀ принимается	H ₀ принимается

Рис. 4.6: Критерий согласия Колмогорова (Смирнова) для сложной гипотезы

4.1.5. Критерий согласия хи-квадрат для сложной гипотезы (в условиях когда неизвестен параметр распределения)

В случае проверки сложной гипотезы используют тестовую статистику

$$\hat{X}_n^2 = X_n^2(\hat{\theta}_n) = \sum_{j=1}^N \frac{(v_j - np_j(\hat{\theta}_n))^2}{np_j(\hat{\theta}_n)}$$

где $\hat{\theta}_n$ оценка максимального правдоподобия θ_n .

$$H_0 \iff \overset{\circ}{X}_n^2 \leq \chi_{1-\alpha, N-1-r}^2$$

где r количество параметров предполагаемого распределения.

Для проверки сложных гипотез будем брать оценки максимального правдоподобия из ДЗ 3.

4.1.5.1. Дискретное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	6.001949	H ₀ принимается	H ₀ принимается	H ₀ принимается
10	3.700749	H ₀ принимается	H ₀ принимается	H ₀ принимается
100	40.751239	H ₀ отвергается	H ₀ отвергается	H ₀ отвергается
200	5.210348	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	15.334200	H ₀ принимается	H ₀ принимается	H ₀ принимается
600	9.086341	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	16.046513	H ₀ принимается	H ₀ принимается	H ₀ принимается
1000	11.851024	H ₀ принимается	H ₀ принимается	H ₀ принимается

Рис. 4.7: Критерий согласия хи-квадрат для сложной гипотезы

4.1.5.2. Непрерывное распределение

	Тестовая статистика	alpha=0.1	alpha=0.05	alpha=0.01
5	77.619011	H ₀ отвергается	H ₀ отвергается	H ₀ отвергается
10	74.722337	H ₀ отвергается	H ₀ отвергается	H ₀ отвергается
100	39.223845	H ₀ принимается	H ₀ принимается	H ₀ принимается
200	77.964025	H ₀ принимается	H ₀ принимается	H ₀ принимается
400	51.275138	H ₀ принимается	H ₀ принимается	H ₀ принимается
600	41.985884	H ₀ принимается	H ₀ принимается	H ₀ принимается
800	43.027974	H ₀ принимается	H ₀ принимается	H ₀ принимается
1000	56.797987	H ₀ принимается	H ₀ принимается	H ₀ принимается

Рис. 4.8: Критерий согласия хи-квадрат для сложной гипотезы

2. Проверка гипотезы об однородности выборок

Пусть $X = (X_1, \dots, X_n)$ из распределения $\mathcal{L}(\xi)$ с неизвестной функцией распределения $F_1(x)$ и $Y = (Y_1, \dots, Y_n)$ из распределения $\mathcal{L}(\eta)$ также с неизвестной функцией распределения $F_2(x)$. Гипотеза однородности формулируется следующим образом $H_0 : F_1(x) = F_2(x)$ и заключается в проверке гипотезы о том, что рассматриваются две выборки из одного и того же распределения.

4.2.1. Дискретное распределение

Пусть Y_1, \dots, Y_n - независимые, одинаково распределенные случайные величины, $Y_i \sim \mathcal{R}[0, 1]$, X_1, \dots, X_n - выборка из некоторого распределения (в

нашем случае из Биномиального), функция распределения которого имеет точки разрыва. Построим следующую случайную величину:

$$U_i = F(X_{i-}) + Y_i[F(x_i) - F(X_{i-})],$$

где $F(X_{i-}) = \lim_{z \downarrow 0}$. Тогда случайная величина $U_i \sim \mathcal{R}[0, 1]$. Значит можно применить критерий однородности Смирнова, который описан ниже (пункт 4.2.1.1).

	5	10	100	200	400	600	800	1000
5	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
10	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
100	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
200	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
400	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
600	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
800	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
1000	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается

Рис. 4.9: Проверка гипотезы об однородности выборок

4.2.1.1. Непрерывное распределение

Для проверки будем использовать критерий однородности Смирнова. Его тестовой статистикой является:

$$D_{n,m} = \sup_{-\infty < x < +\infty} |\widehat{F}_n(x) - \widehat{F}_m(x)|$$

Эмпирическая функция распределения является оптимальной оценкой для теоретической функции распределения, и с увеличением объема выборки они сближаются, поэтому в случае справедливости гипотезы однородности H_0 функции $\widehat{F}_n(x)$ и $\widehat{F}_m(x)$ оценивают одну и ту же неизвестную функцию распределения. Тем самым в этом случае (по крайней мере при больших n и m) статистика $D_{n,m}$ не должна существенно отклоняться от нуля. Отсюда следует, что слишком большие значения этой статистики следует расценивать как свидетельство против гипотезы H_0 . Критическую границу $t_\alpha(n, m)$ при заданном уровне значимости α находят на основании известного при гипотезе H_0 предельного распределения статистики $D_{n,m}$. При больших n, m полагают $t_\alpha(n, m) = \sqrt{\frac{1}{n} + \frac{1}{m}} \lambda_\alpha$:

$$P\left(D_{n,m} > \sqrt{\frac{1}{n} + \frac{1}{m}} | H_0\right) = P\left(\sqrt{\frac{nm}{n+m}} > D_{n,m} | H_0\right) \approx 1 - K(\lambda_\alpha) = \alpha,$$

$K(t)$ - функция распределения Колмогорова. Таким образом

$$H_0 \iff D_{n,m} \leq \lambda_\alpha \sqrt{\frac{1}{n} + \frac{1}{m}},$$

α возьмем 0.1

	5	10	100	200	400	600	800	1000
5	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
10	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
100	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
200	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
400	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается
600	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 отвергается
800	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 отвергается
1000	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 принимается	H_0 отвергается	H_0 отвергается	H_0 принимается

Рис. 4.10: Проверка гипотезы об однородности выборок

Домашнее задание 5.

Различение статистических гипотез

1. Теоретическое введение

Пусть имеется выборка $X = (X_1, \dots, X_n)$ из распределения $\mathcal{L}(\xi)$. Рассмотрим две гипотезы:

$$H_0 : \xi \sim \mathcal{L}(\theta_0) \text{ — основная гипотеза}$$

$$H_1 : \xi \sim \mathcal{L}(\theta_1) \text{ — альтернативная гипотеза}$$

Пусть для биномиального распределения гипотезы будут:

$$H_0 : \xi \sim \text{Bin}(35, \theta_0), \theta_0 = 0.5$$

$$H_1 : \xi \sim \text{Bin}(35, \theta_1), \theta_1 = 0.53,$$

а для распределения Эрланга:

$$H_0 : \xi \sim \text{Erlang}(8, \theta_0), \theta_0 = 0.125$$

$$H_1 : \xi \sim \text{Erlang}(8, \theta_1), \theta_1 = 0.115$$

Выбор из двух простых гипотез можно представить в виде параметрической гипотезы

$$\Theta = 0, 1, F_\theta(x) = (1 - \theta)F_0(x) + \theta F_1(x)$$

В случае параметрических гипотез функцию мощности можно переписать в виде:

$$W(\theta) = W(\theta, \mathfrak{X}_{1,\alpha}) = P_\theta(X \in \mathfrak{X}_{1,\alpha})$$

Понятия ошибок

$$P(X \in \mathfrak{X}_1 | H_0) = \alpha \text{ — ошибка первого рода (отвержение истины)}$$

$$P(X \in \mathfrak{X}_0 | H_1) = \beta \text{ — ошибка второго рода (принятие лжи за истину)}$$

Через функцию мощности критерия можно выразить вероятности ошибок первого и второго рода:

$$\alpha = W(F_{0,x}) = W(\theta_0, \mathfrak{X}_{1,\alpha})$$

$$\beta = 1 - W(F_{1,x}) = 1 - W(\theta_1, \mathfrak{X}_{1,\alpha})$$

Наиболее мощный критерий с уровнем значимости α - параметрический критерий минимизирующий ошибку 2 рода при заданной ошибке 1 рода. Критическую область можно построить следующим образом: множество $\mathfrak{X}_{1,\alpha}$ состоит из таких \bar{x} , для которых правдоподобие $L(\bar{x}, \theta_1)$ будет больше $L(\bar{x}, \theta_0)$.

Функция, имеющая вид:

$$l(\bar{x}) = \frac{L(\bar{x}, \theta_1)}{L(\bar{x}, \theta_0)} = \frac{\prod_{i=1}^n f_1(x_i)}{\prod_{i=1}^n f_0(x_i)}$$

называется функцией отношения правдоподобия. Критическим множеством критерия Неймана-Пирсона называется множество $\mathfrak{X}_{1,\alpha}^*$ имеющее вид:

$$\mathfrak{X}_{1,\alpha}^* = \{\bar{x} \in \mathfrak{X} : l(\bar{x}) \geq c_\alpha\},$$

где c_α такое, что ошибка первого рода равна α .

$$W(\theta_0, \mathfrak{X}_{1,\alpha}^*) = P_0(\bar{x} \in \mathfrak{X}_{1,\alpha}^*) = \alpha$$

Лемма Неймана-Пирсона говорит о том, что критическая область $\mathfrak{X}_{1,\alpha}^*$ задает наиболее мощный критерий среди всех критериев с уровнем значимости α . Кроме того критерий Неймана-Пирсона несмещенный.

2. Вычисление функции отношения правдоподобия

5.2.1. Дискретное распределение

$$\begin{aligned} l(\bar{x}) &= \frac{\prod_{i=1}^n f_1(x_i)}{\prod_{i=1}^n f_0(x_i)} = \frac{\prod_{i=1}^n \binom{n}{x_i} \theta_1^{x_i} (1 - \theta_1)^{n-x_i}}{\prod_{i=1}^n \binom{n}{x_i} \theta_0^{x_i} (1 - \theta_0)^{n-x_i}} = \prod_{i=1}^n \left(\frac{\theta_1}{\theta_0} \right)^{x_i} \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^{n-x_i} = \\ &= \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^{n^2} \prod_{i=1}^n \left(\frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} \right)^{x_i} = \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^{n^2} \exp \left[\ln \left(\frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} \right) \sum_{i=1}^n x_i \right] \\ &\left(\frac{1 - \theta_1}{1 - \theta_0} \right)^{n^2} \exp \left[\ln \left(\frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} \right) \sum_{i=1}^n x_i \right] \geq c \\ &\exp \left[\ln \left(\frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} \right) \sum_{i=1}^n x_i \right] \geq c \left(\frac{1 - \theta_1}{1 - \theta_0} \right)^{n^2} \\ &\ln \left(\frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} \right) \sum_{i=1}^n x_i \geq \ln c + n^2 \ln \left(\frac{1 - \theta_1}{1 - \theta_0} \right) \end{aligned}$$

$$\sum_{i=1}^n x_i \geq \frac{\ln c + n^2 \ln \left(\frac{1-\theta_1}{1-\theta_0} \right)}{\ln \left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} \right)}$$

Пусть $c_\alpha = \frac{\ln c + n^2 \ln \left(\frac{1-\theta_1}{1-\theta_0} \right)}{\ln \left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} \right)}$.

Критическая область: $\sum_{i=1}^n x_i \geq c_\alpha$ (Отвергаем H_0).

Тогда $\alpha = P(H_1|H_0) = P_0(\sum_{i=1}^n x_i \geq c_\alpha) = 1 - P_0(\sum_{i=1}^n x_i < c_\alpha)$. Более того,

$$\sum_{i=1}^n x_i = n\bar{X} = \xi_0 \sim \text{Bin}(35n, \theta_0).$$

Значит $\alpha = 1 - P(\xi_0 < c_\alpha)$, а $\beta = P(H_0|H_1) = 1 - P(H_1|H_0) = 1 - P_1(\sum_{i=1}^n x_i \geq c_\alpha) = 1 - (1 - P_1(\sum_{i=1}^n x_i < c_\alpha)) = P(\xi_1 < c_\alpha)$, где $\xi_1 \sim \text{Bin}(35n, \theta_1)$.

5.2.2. Непрерывное распределение

$$\begin{aligned} l(\bar{x}) &= \frac{\prod_{i=1}^n f_1(x_i)}{\prod_{i=0}^n f_0(x_i)} = \frac{\prod_{i=1}^n \frac{\theta_1^m}{(m-1)!} x_i^{m-1} e^{-\theta_1 x_i}}{\prod_{i=1}^n \frac{\theta_0^m}{(m-1)!} x_i^{m-1} e^{-\theta_0 x_i}} = \left(\frac{\theta_1}{\theta_0} \right)^{nm} \prod_{i=1}^n \exp(\theta_0 x_i - \theta_1 x_i) = \\ &= \left(\frac{\theta_1}{\theta_0} \right)^{nm} \prod_{i=1}^n \exp[x_i(\theta_0 - \theta_1)] = \left(\frac{\theta_1}{\theta_0} \right)^{nm} \exp \left[(\theta_0 - \theta_1) \sum_{i=1}^n x_i \right] \end{aligned}$$

$$\left(\frac{\theta_1}{\theta_0} \right)^{nm} \exp \left[(\theta_0 - \theta_1) \sum_{i=1}^n x_i \right] \geq c$$

$$\exp \left[(\theta_0 - \theta_1) \sum_{i=1}^n x_i \right] \geq c \left(\frac{\theta_0}{\theta_1} \right)^{nm}$$

$$(\theta_0 - \theta_1) \sum_{i=1}^n x_i \geq \ln c + nmln \left(\frac{\theta_0}{\theta_1} \right)$$

$$\sum_{i=1}^n x_i \geq \frac{\ln c + nmln \left(\frac{\theta_0}{\theta_1} \right)}{(\theta_0 - \theta_1)}$$

Пусть $c_\alpha = \frac{\ln c + nm \ln(\frac{\theta_0}{\theta_1})}{(\theta_0 - \theta_1)}$.

Критическая область: $\sum_{i=1}^n x_i \geq c_\alpha$ (Отвергаем H_0).

Тогда $\alpha = P(H_1|H_0) = P_0(\sum_{i=1}^n x_i \geq c_\alpha) = 1 - P_0(\sum_{i=1}^n x_i < c_\alpha)$. Более того

$\sum_{i=1}^n x_i = n\bar{X} = \xi_0 \sim \text{Erlang}(nm, \theta_0)$.

Значит $\alpha = 1 - P(\xi_0 < c_\alpha)$, а $\beta = P(H_0|H_1) = 1 - P(H_1|H_0) = 1 - P_1(\sum_{i=1}^n x_i \geq c_\alpha) = 1 - (1 - P_1(\sum_{i=1}^n x_i < c_\alpha)) = P(\xi_1 < c_\alpha)$, где $\xi_1 \sim \text{Erlang}(nm, \theta_1)$.

3. Вычисление критической области

5.3.1. Дискретное распределение

Определим c_α при $\alpha = 0.01$. Для этого потребуется вычислить квантиль уровня 0.99 для $F_{\xi_0}(x)$.

	Sum(x_i)	c_alpha	Итог
5	83	103.0	Принимаем theta = 0.5
10	160	197.0	Принимаем theta = 0.5
100	1760	1819.0	Принимаем theta = 0.5
200	3516	3597.0	Принимаем theta = 0.5
400	7029	7138.0	Принимаем theta = 0.5
600	10466	10669.0	Принимаем theta = 0.5
800	13981	14195.0	Принимаем theta = 0.5
1000	17308	17718.0	Принимаем theta = 0.5

Рис. 5.1: c_α

Теперь вычислим ошибку второго рода для известных $\alpha = 0.01$ и c_α :

	Sum(x_i)	c_alpha	beta
5	83	103.0	9.486609e-01
10	160	197.0	9.007911e-01
100	1760	1819.0	1.146532e-01
200	3516	3597.0	3.542504e-03
400	7029	7138.0	9.501432e-07
600	10466	10669.0	9.923873e-11
800	13981	14195.0	6.217615e-15
1000	17308	17718.0	2.831892e-19

Рис. 5.2: Вычисление ошибки 2-го рода

Листинг 5.1: Binomial distribution

```

1  theta_0 = make_binomial.theta
2  theta_1 = theta_0 + 0.1
3
4  y = [binom.ppf(0.99, i*make_binomial.n, theta_0
5          ) for i in make_binomial.N]
6
7  sums = []
8
9  for i in make_binomial.binomial:
10     sums.append(sum(i))
11
12  data = { 'Sum(x_i)':sums, 'c_alpha': y,
13          '...': [ '..._=' + (f"{theta_0}" if sums[i] < y[i]
14                      else f"{theta_1}") for i in range(len(
15                          make_binomial.N))] }
16
17  table_crit = pd.DataFrame(data, index=make_binomial.N)
18  print(table_crit)
19
20  beta = [binom.cdf(y[i], value*make_binomial.n, theta_1
21                  ) for i,value in enumerate(
22                      make_binomial.N)]
23
24  data_beta = { 'Sum(x_i)':sums, 'c_alpha':y, 'beta':beta}
25
26  table_crit_beta = pd.DataFrame(data=data_beta,
27                                index=make_binomial.N)
28  print(table_crit_beta)

```

5.3.2. Непрерывное распределение

Определим c_α при $\alpha = 0.01$. Для этого потребуется вычислить квантиль уровня 0.99 для $F_{\xi_0}(x)$.

	Sum(x_i)	c_alpha	Итог
5	281.362343	449.315170	Принимаем theta = 0.125
10	539.560920	818.120378	Принимаем theta = 0.125
100	6089.264580	6938.127130	Принимаем theta = 0.125
200	12768.268094	13556.175178	Принимаем theta = 0.125
400	26216.522839	26664.535072	Принимаем theta = 0.125
600	39451.535505	39701.145846	Принимаем theta = 0.125
800	50935.941712	52700.617261	Принимаем theta = 0.125
1000	63482.641857	65676.354774	Принимаем theta = 0.125

Рис. 5.3: c_α

Теперь вычислим ошибку второго рода для известных $\alpha = 0.01$ и c_α :

	Sum(x_i)	c_alpha	beta
5	281.362343	449.315170	9.593271e-01
10	539.560920	818.120378	9.366746e-01
100	6089.264580	6938.127130	4.748562e-01
200	12768.268094	13556.175178	1.523160e-01
400	26216.522839	26664.535072	8.444310e-03
600	39451.535505	39701.145846	2.966450e-04
800	50935.941712	52700.617261	7.940688e-06
1000	63482.641857	65676.354774	1.765450e-07

Рис. 5.4: Вычисление ошибки 2-го рода

Листинг 5.2: Erlang distribution

```

1 theta_0 = make_erlang.theta
2 theta_1 = theta_0 - 0.025
3
4 y = [erlang.ppf(0.99, a=i * make_erlang.m, scale=1/theta_0
5         ) for i in make_erlang.N]
6
7 sums = []
8
9 for i in make_erlang.erlang:
10     sums.append(sum(i))

```

```

11
12 data = { 'Sum(x_i)': sums, 'c_alpha': y,
13           '...': [ '..._=' + (f"{theta_0}" if sums[i] < y[i]
14                       else f"{theta_1}") for i in
15                       range(len(make_erlang.N))] }
16
17 table_crit = pd.DataFrame(data, index=make_erlang.N)
18 print(table_crit)
19
20 beta = [erlang.cdf(y[i], a=value * make_erlang.m,
21                 scale=1/theta_1) for i, value
22          in enumerate(make_erlang.N)]
23
24 data_beta = { 'Sum(x_i)': sums, 'c_alpha': y, 'beta': beta }
25
26 table_crit_beta = pd.DataFrame(data=data_beta,
27                               index=make_erlang.N)
28 print(table_crit_beta)

```

4. Вычисление минимального необходимого количества...

5.4.1. Дискретное распределение

Критическая область критерия Неймана-Пирсона для рассматриваемого случая выглядит следующим образом:

$$\left(\frac{1-\theta_1}{1-\theta_0}\right)^{n^2} \exp \left[\ln \left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} \right) \sum_{i=1}^n x_i \right]$$

Это эквивалентно следующему

$$\sum_{i=1}^n x_i \geq \frac{\ln c + n^2 \ln \left(\frac{1-\theta_1}{1-\theta_0} \right)}{\ln \left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} \right)}$$

Заметим, что

$$\sum_{i=1}^n x_i \sim \text{Bin}(35n, \theta)$$

В силу Ц.П.Т. при $n \rightarrow \infty$:

$$\text{Bin}(35n, \theta) \approx N(35n\theta, 35n\theta(1-\theta))$$

Проведем нормализацию для $\sum_{i=1}^n x_i$ путем сдвига и деления:

$$\frac{\sum_{i=1}^n x_i - 35n\theta_0}{\sqrt{35n\theta_0(1-\theta_0)}} \sim N(0, 1)$$

Обозначим

$$t = \frac{\frac{\ln c + n^2 \ln(\frac{1-\theta_1}{1-\theta_0})}{\ln(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})} - 35n\theta_0}{\sqrt{35n\theta_0(1-\theta_0)}}$$

Тогда ошибка первого рода:

$$\alpha = P_0\left(\frac{\sum_{i=1}^n x_i - 35n\theta_0}{\sqrt{35n\theta_0(1-\theta_0)}} \geq t\right) = P_0(\xi \geq t) = 1 - P_0(\xi \leq t) = \Phi(-t),$$

где Φ - функция стандартного нормального распределения. Пусть t_α - решение данного уравнения от переменной t . Т.е. $\alpha = \Phi(t_\alpha)$ - такое решение найдется всегда в виду непрерывности функции Φ .

Аналогичным образом, при фиксированной ошибке первого рода выпишем мощность критерия:

$$\begin{aligned} 1-\beta &= P_1\left(\frac{\sum_{i=1}^n x_i - 35n\theta_0}{\sqrt{35n\theta_0(1-\theta_0)}} \geq t_\alpha\right) = P_1\left(\sum_{i=1}^n x_i - 35n\theta_0 \geq t_\alpha \sqrt{35n\theta_0(1-\theta_0)}\right) = \\ &= P_1\left(\sum_{i=1}^n x_i - 35n\theta_1 \geq t_\alpha \sqrt{35n\theta_0(1-\theta_0)} + 35n(\theta_0 - \theta_1)\right) = \\ &= P_1\left(\frac{\sum_{i=1}^n x_i - 35n\theta_1}{\sqrt{35n\theta_1(1-\theta_1)}} \geq \frac{t_\alpha \sqrt{35n\theta_0(1-\theta_0)} + 35n(\theta_0 - \theta_1)}{\sqrt{35n\theta_1(1-\theta_1)}}\right) = \\ &= \Phi\left(-\frac{t_\alpha \sqrt{35n\theta_0(1-\theta_0)} + 35n(\theta_0 - \theta_1)}{\sqrt{35n\theta_1(1-\theta_1)}}\right) \end{aligned}$$

Откуда

$$\begin{aligned} \beta &= 1 - \Phi\left(-\frac{t_\alpha \sqrt{35n\theta_0(1-\theta_0)} + 35n(\theta_0 - \theta_1)}{\sqrt{35n\theta_1(1-\theta_1)}}\right) = \\ &= \Phi\left(\frac{t_\alpha \sqrt{35n\theta_0(1-\theta_0)} + 35n(\theta_0 - \theta_1)}{\sqrt{35n\theta_1(1-\theta_1)}}\right) \end{aligned}$$

Пусть $\Phi(t_\beta) = \beta$, тогда

$$t_\beta = \frac{t_\alpha \sqrt{35n\theta_0(1-\theta_0)} + 35n(\theta_0 - \theta_1)}{\sqrt{35n\theta_1(1-\theta_1)}}$$

Получаем $t_\beta = t_\alpha \sqrt{\frac{\theta_0(1-\theta_0)}{\theta_1(1-\theta_1)}} + \sqrt{35n} \left(\frac{\theta_0 - \theta_1}{\sqrt{\theta_1(1-\theta_1)}} \right)$.

Следовательно необходимый объем выборки оценивается по формуле:

$$n = \left\lceil \left(\frac{t_\beta \sqrt{\theta_1(1-\theta_1)} - t_\alpha \sqrt{\theta_0(1-\theta_0)}}{\sqrt{35}(\theta_0 - \theta_1)} \right)^2 \right\rceil$$

alpha	beta	n
0.10	0.10	53
0.05	0.05	86
0.01	0.01	172

Рис. 5.5: Минимальное необходимое n

5.4.2. Непрерывное распределение

Критическая область критерия Неймана-Пирсона для рассматриваемого случая выглядит следующим образом:

$$\left(\frac{\theta_1}{\theta_0} \right)^{nm} \exp \left[(\theta_0 - \theta_1) \sum_{i=1}^n x_i \right]$$

Это эквивалентно следующему

$$\sum_{i=1}^n x_i \geq \frac{\ln c + nm \ln \left(\frac{\theta_0}{\theta_1} \right)}{(\theta_0 - \theta_1)}$$

Заметим, что

$$\sum_{i=1}^n x_i \sim \text{Erlang}(nm, \theta)$$

В силу Ц.П.Т. при $n \rightarrow \infty$:

$$\text{Erlang}(nm, \theta) \approx N(nm\theta, nm\theta^2)$$

Проведем нормализацию для $\sum_{i=1}^n x_i$ путем сдвига и деления:

$$\frac{\sum_{i=1}^n x_i - nm\theta_0}{\sqrt{nm}\theta_0} \sim N(0, 1)$$

Обозначим

$$t = \frac{\frac{\ln c + nml \ln(\frac{\theta_0}{\theta_1})}{(\theta_0 - \theta_1)} - nm\theta_0}{\sqrt{nm}\theta_0}$$

Тогда ошибка первого рода:

$$\alpha = P_0\left(\frac{\sum_{i=1}^n x_i - nm\theta_0}{\sqrt{nm}\theta_0} \geq t\right) = P_0(\xi \geq t) = 1 - P_0(\xi \leq t) = \Phi(-t),$$

где Φ - функция стандартного нормального распределения. Пусть t_α - решение данного уравнения от переменной t . Т.е. $\alpha = \Phi(t_\alpha)$ - такое решение найдется всегда в виду непрерывности функции Φ .

Аналогичным образом, при фиксированной ошибке первого рода выпишем мощность критерия:

$$\begin{aligned} 1 - \beta &= P_1\left(\frac{\sum_{i=1}^n x_i - nm\theta_0}{\sqrt{nm}\theta_0} \geq t_\alpha\right) = P_1\left(\sum_{i=1}^n x_i - nm\theta_0 \geq t_\alpha \sqrt{nm}\theta_0\right) = \\ &= P_1\left(\sum_{i=1}^n x_i - nm\theta_1 \geq t_\alpha \sqrt{nm}\theta_0 + nm(\theta_0 - \theta_1)\right) = P_1\left(\frac{\sum_{i=1}^n x_i - nm\theta_1}{\sqrt{nm}\theta_1} \geq \right. \\ &\quad \left. \frac{t_\alpha \sqrt{nm}\theta_0 + nm(\theta_0 - \theta_1)}{\sqrt{nm}\theta_1}\right) = \Phi\left(-\frac{t_\alpha \sqrt{nm}\theta_0 + nm(\theta_0 - \theta_1)}{\sqrt{nm}\theta_1}\right) \end{aligned}$$

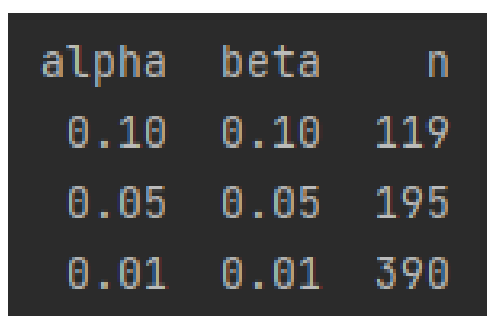
Откуда

$$\beta = 1 - \Phi\left(-\frac{t_\alpha \sqrt{nm}\theta_0 + nm(\theta_0 - \theta_1)}{\sqrt{nm}\theta_1}\right) = \Phi\left(\frac{t_\alpha \sqrt{nm}\theta_0 + nm(\theta_0 - \theta_1)}{\sqrt{nm}\theta_1}\right)$$

Пусть $\Phi(t_\beta) = \beta$, тогда $t_\beta = \frac{t_\alpha \sqrt{nm}\theta_0 + nm(\theta_0 - \theta_1)}{\sqrt{nm}\theta_1}$.

Следовательно необходимый объем выборки оценивается по формуле:

$$n = \left\lceil \left(\frac{t_\alpha \theta_0 - t_\beta \theta_1}{\sqrt{m}(\theta_1 - \theta_0)} \right)^2 \right\rceil$$



A table with three columns: alpha, beta, and n. The first row shows alpha=0.10, beta=0.10, n=119. The second row shows alpha=0.05, beta=0.05, n=195. The third row shows alpha=0.01, beta=0.01, n=390.

alpha	beta	n
0.10	0.10	119
0.05	0.05	195
0.01	0.01	390

Рис. 5.6: Минимальное необходимое n