



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М. В. ЛОМОНОСОВА  
Факультет вычислительной математики и кибернетики  
Кафедра системного анализа

Отчёт по компьютерному практикуму:  
стохастический анализ и моделирование

*Студент 415 группы*  
А. А. Анашкина

*Руководитель практикума*  
С. Н. Смирнов

# Содержание

<b>1 Задание 1</b>	<b>2</b>
1.1 Реализация датчика биномиального распределения . . . . .	2
1.2 Реализация генератора геометрического распределения . . . . .	4
1.3 Игра в орлянку . . . . .	6
<b>2 Задание 2</b>	<b>7</b>
2.1 Построение датчика сингулярного распределения . . . . .	7
2.2 Использование критерия Смирнова . . . . .	9
2.3 Рассчёт математического ожидания и дисперсии . . . . .	11
<b>3 Задание 3</b>	<b>12</b>
3.1 Построение датчика экспоненциального распределения . . . . .	12
3.2 Распределение случайной величины $Y = \min(X_1, \dots, X_n)$ . . . . .	14
3.3 Построение датчика пуассоновского распределения 1 . . . . .	15
3.4 Построение датчика пуассоновского распределения 2 . . . . .	16
3.5 Построение датчика стандартного нормального распределения . . . . .	17
<b>4 Задание 4</b>	<b>19</b>
4.1 Построение датчика распределения Коши . . . . .	19
4.2 Применение метода фон Неймана . . . . .	20
4.3 Сравнение скорости моделирования стандартного нормального распределения	22
<b>5 Задание 5</b>	<b>23</b>
5.1 Справедливость ЗБЧ и ЦПТ . . . . .	23
5.2 Построение доверительных интервалов для среднего и дисперсии . . . . .	25
5.3 Распределение суммы элементов из распределения Коши . . . . .	27
<b>6 Задание 6</b>	<b>28</b>
6.1 Вычисление интеграла . . . . .	28
<b>7 Задание 7</b>	<b>30</b>
7.1 Метод случайного поиска . . . . .	30
7.2 Метод имитации отжига . . . . .	32
7.3 Сравнение со стандартными методами оптимизации . . . . .	33
<b>8 Задание 8</b>	<b>34</b>
8.1 Метод Монте-Карло для уравнения Лапласа . . . . .	34
8.2 Применение метода . . . . .	35
<b>9 Задание 9</b>	<b>36</b>
9.1 Винеровский процесс . . . . .	36
9.2 Процесс Орнштейна-Уленбека . . . . .	38
<b>10 Задание 10</b>	<b>40</b>
<b>11 Задание 11</b>	<b>42</b>
11.1 Система массового обслуживания . . . . .	42
11.2 Система массового обслуживания с циклической интенсивностью . . . . .	47
11.3 Работа страховой компании . . . . .	48

# 1 Задание 1

## 1.1 Реализация датчика биномиального распределения

**Условие.** Реализовать генератор схемы Бернулли с заданной вероятностью успеха  $p$ . На основе генератора схемы Бернулли построить датчик биномиального распределения.

В общем случае у нас есть некоторое распределение  $\mathbb{P}_t(A)$ . Мы хотим научиться семплировать из него. Для начала определим понятие сэмплирования.

Рассматриваем вероятностное пространство  $(\Omega, \mathcal{F}, \mathbb{P})$ . Хотим построить отображение  $\zeta : \Omega \rightarrow \mathbb{R}$ , для которого:

$$\mathbb{P}(\zeta^{-1}(A)) = \mathbb{P}_t(A).$$

Проблема заключается в том, что мы не знаем явный вид вероятностного пространства, а лишь знаем некоторое отображение  $\eta : \Omega \rightarrow \mathbb{R}$ . Поэтому для конструирования  $\zeta(\omega)$  необходимо построить некоторую детерминированную функцию  $\phi : \mathbb{R} \rightarrow \mathbb{R}$ , которая обеспечит композицию

$$\zeta(\omega) = \phi(\eta(\omega)).$$

**Определение 1.** Схемой Бернулли с заданной вероятностью  $p$  называется эксперимент, состоящий из серии испытаний, удовлетворяющих следующим условиям:

1. Отсутствие взаимного влияния в испытаниях;
2. Воспроизводимость. Однородные испытания проводятся в сходных условиях;
3. Существует признак, который реализуется или не реализуется в испытании. Признак может быть отнесен к любому из испытаний.

Итак, целевое распределение задаётся следующим образом:

$$\mathbb{P}_t(A) = \begin{cases} 1 - p, & (0 \in A) \cdot (1 \notin A), \\ p, & (0 \notin A) \cdot (1 \in A), \\ 1, & (0 \in A) \cdot (1 \in A), \\ 0, & (0 \notin A) \cdot (1 \notin A). \end{cases}$$

Возьмём случайную величину  $\eta \sim U[0, 1]$ , для которой справедливо:

$$\mathbb{P}(\eta \in A) = \lambda(A).$$

Заметим, что  $\mathbb{P}_t$  "чувствительно" только к точкам 0 и 1, поэтому искомая функция  $\phi$  должна принимать значения только 0 и 1. Для такой функции можем записать:

$$\mathbb{P}(\phi(\eta) \in A) = \mathbb{P}(\eta \in \phi^{-1}(A)) = \lambda(\phi^{-1}(A)).$$

Таким образом,  $\lambda(\phi^{-1}(0)) = 1 - p$ ,  $\lambda(\phi^{-1}(1)) = p$  Поэтому в качестве такого отображения можем взять

$$\phi(x) = \begin{cases} 0, & x \in [0, 1 - p], \\ 1, & x \in (1 - p, 1] \end{cases}$$

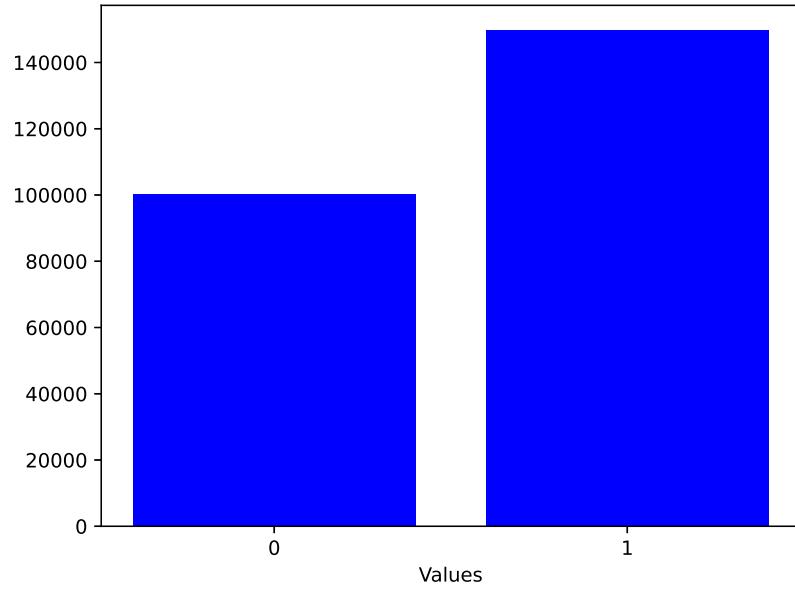


Рис. 1: Результат работы генератора схемы Бернулли с вероятностью успеха  $p = 0.6$  и размерами выходной матрицы  $(500, 500)$

Биномиальное распределение - это распределение числа появления события, имеющего вероятность  $p$  в  $n$  независимых испытаниях. На основе генератора схемы Бернулли можем легко построить датчик этого распределения. Биномиальное распределение можем интерпретировать как распределение количества успехов в схеме Бернулли.

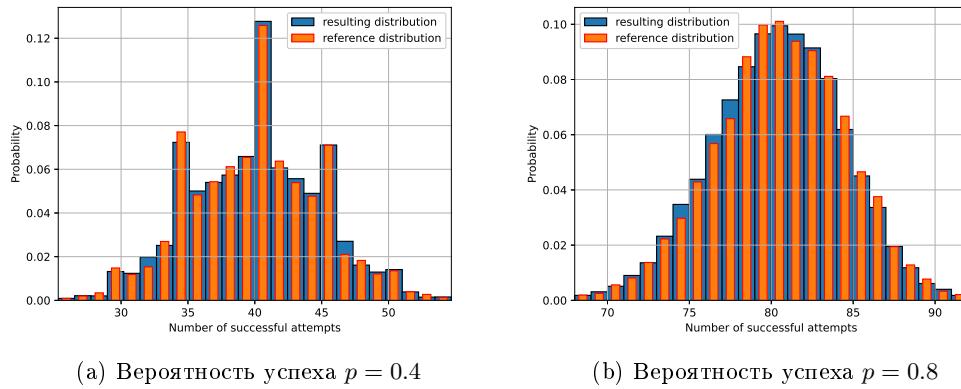


Рис. 2: Результат работы датчика биномиального распределения с объёмом выборки  $n = 100$

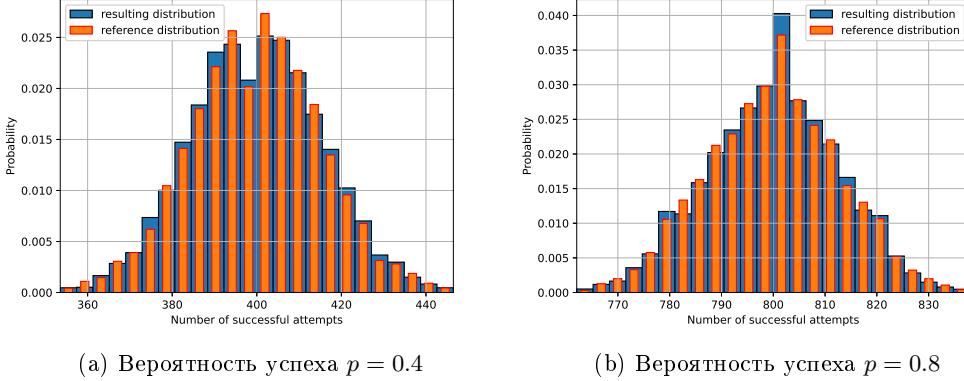


Рис. 3: Результат работы датчика биномиального распределения с объёмом выборки  $n = 1000$

## 1.2 Реализация генератора геометрического распределения

**Условие.** Реализовать генератор геометрического распределения; проверить для данного распределения свойство отсутствия памяти.

**Определение 2.** Случайная величина  $\eta$  имеет геометрическое распределение с параметром  $p \in (0, 1]$ , если

$$\mathbb{P}(\eta = k) = p \cdot (1 - p)^k, \quad k = 0, 1, \dots$$

То есть, геометрическое распределение можно интерпретировать как вероятность того, что в схеме Бернулли спротивило к неудаче до первого успеха. Тогда случайная величина  $\eta$  — число неудач до появления первого успеха.

Воспользуемся методом обратных функций. Введём обозначение:

$$s_{n+1} = p_0 + \dots + p_n = \sum_{i=0}^n p (1 - p)^i = 1 - (1 - p)^{n+1}.$$

Тогда событие — число неудач до появления первого успеха равно  $n$ , можно интерпретировать как событие — значение случайной величины из равномерного распределения на  $[0, 1]$  находится между  $s_n$  и  $s_{n+1}$ . Запишем это на математическом языке:

$$\begin{aligned} \{\eta = n\} &= \{s_n < \alpha \leq s_{n+1}\} = \{1 - (1 - p)^n < \alpha \leq 1 - (1 - p)^{n+1}\} = \\ &= \left\{ (1 - p)^{n+1} \leq 1 - \alpha < (1 - p)^n \right\} = \left\{ (n+1) \ln(1 - p) \leq \ln(1 - \alpha) < n \ln(1 - p) \right\} = \left\{ n < \frac{\ln(1 - \alpha)}{\ln(1 - p)} \leq n+1 \right\}. \end{aligned}$$

Ввиду того, что  $\alpha \sim U[0, 1]$ :

$$\eta = \left\lfloor \frac{\ln(\alpha)}{\ln(1 - p)} \right\rfloor$$

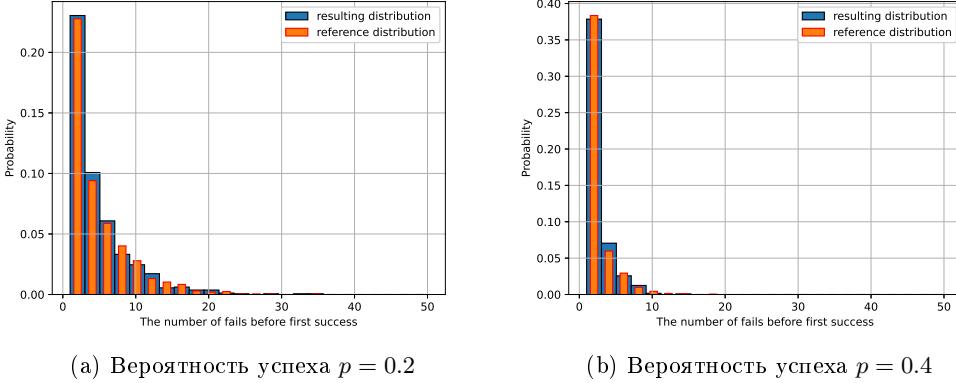


Рис. 4: Результат работы генератора геометрического распределения с объёмом выборки  $n = 1000$

Для геометрического распределения известно, что

$$F_\eta(x) = \mathbb{P}(\eta < x) = \sum_{k=0}^{x-1} \mathbb{P}(\eta = k) = p \sum_{k=0}^{x-1} (1-p)^k = 1 - (1-p)^x.$$

**Свойство 1.** Если  $\eta \sim Geom(p)$ , то  $\forall t, \Delta t \in \mathbb{Z}_+$

$$\mathbb{P}(\eta > t + \Delta t | \eta \geq t) = \mathbb{P}(\eta > \Delta t).$$

*Доказательство.* Воспользуемся формулой условной вероятности и монотонностью меры  $\mathbb{P}$  по вложенным событиям.

$$\begin{aligned} \mathbb{P}(\eta > t + \Delta t | \eta \geq t) &= \frac{\mathbb{P}(\eta > t + \Delta t)}{\mathbb{P}(\eta \geq t)} = \\ &= \left\{ \mathbb{P}(\eta \leq x) = \mathbb{P}(\eta < x) + \mathbb{P}(\eta = x) = 1 - (1-p)^x + p(1-p)^x = 1 - (1-p)^{x+1} \right\} = \\ &= \frac{(1-p)^{t+\Delta t+1}}{(1-p)^t} = (1-p)^{\Delta t+1} = \mathbb{P}(\eta > \Delta t). \end{aligned}$$

□

Это свойство распределений называется свойством отсутствия памяти. Причём геометрическое распределение — единственное дискретное распределение, обладающее этим свойством.

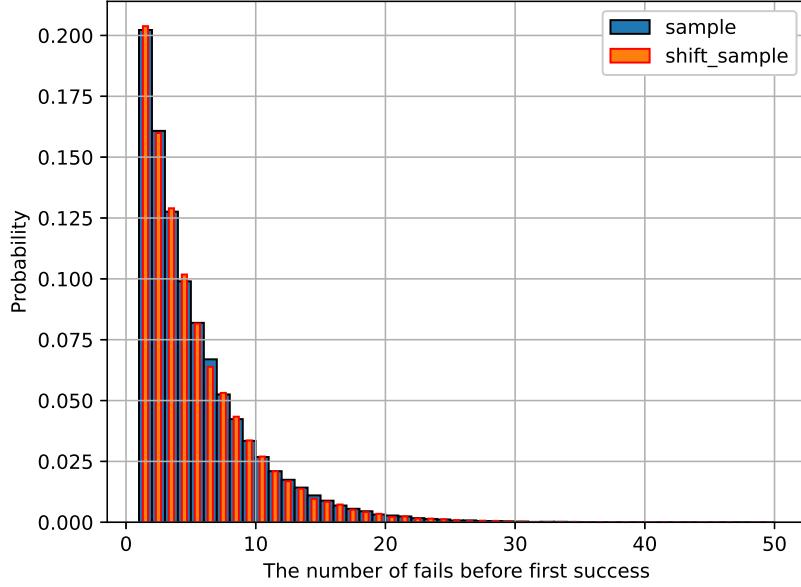


Рис. 5: Свойство отсутствия памяти с вероятностью успеха  $p = 0.6$  и размером выборки 100000

### 1.3 Игра в орлянку

**Условие.** Промоделировать игру в орлянку: бесконечную последовательность независимых испытаний Бернулли с бросанием "правильной" монеты. Величина выигрыша  $S_n$  определяется как сумма по  $n$  испытаниям значений 1 и -1 в зависимости от выпавшей стороны монеты. Проиллюстрировать в виде ломаной поведение нормированной суммы  $Y(i) = \frac{S_i}{\sqrt{n}}$  как функцию от номера испытания  $i$  для отдельно взятой траектории. Дать теоретическую оценку для значения  $Y(n)$  при  $n \rightarrow \infty$ .

Введём последовательность случайных величин  $X_1, X_2, \dots$ . Будем моделировать их с помощью генератора схемы Бернулли с параметром 0.5. По условию задачи значения, которые принимает случайная величина — 1 и -1, поэтому в выборке, сгенерированной с помощью схемы Бернулли заменяем все 0 на -1.

Далее дадим теоретическую оценку для значения  $Y(n)$  при  $n \rightarrow \infty$ .

**Теорема 1** (Центральная Предельная Теорема). Пусть  $X_1, X_2, \dots$  - последовательность независимых одинаково распределённых случайных величин с  $\mathbb{E}X_1^2 < \infty$  и  $S_n = X_1 + \dots + X_n$ . Тогда при  $n \rightarrow \infty$  для  $\forall x \in \mathbb{R}$  имеем

$$\mathbb{P}\left(\frac{S_n - \mu}{\sigma} \leq x\right) \rightarrow N(0, 1).$$

Доказательство можно найти в [1].

В нашем случае случайные величины  $X_1, X_2, \dots$  за счёт реализации через схему Бернулли являются независимыми одинаково распределёнными. Значит можем применить ЦПТ.

$$\mu = \mathbb{E}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{E}X_i = n \cdot \mathbb{E}X_i = n \cdot (1 \cdot 0.5 - 1 \cdot 0.5) = 0$$

$$\begin{aligned} \text{var}(S_n) &= n \cdot \text{var}(X_i) = n \cdot (\mathbb{E}X_i^2 - (\mathbb{E}X_i)^2) = n \cdot (1 - 0) = n \\ \mathbb{P}\left(\frac{S_n - \mu}{\sigma} \leq x\right) &= \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq x\right) \rightarrow N(0, 1). \end{aligned}$$

Таким образом,

$$Y(n) \rightarrow N(0, 1), \quad n \rightarrow \infty.$$

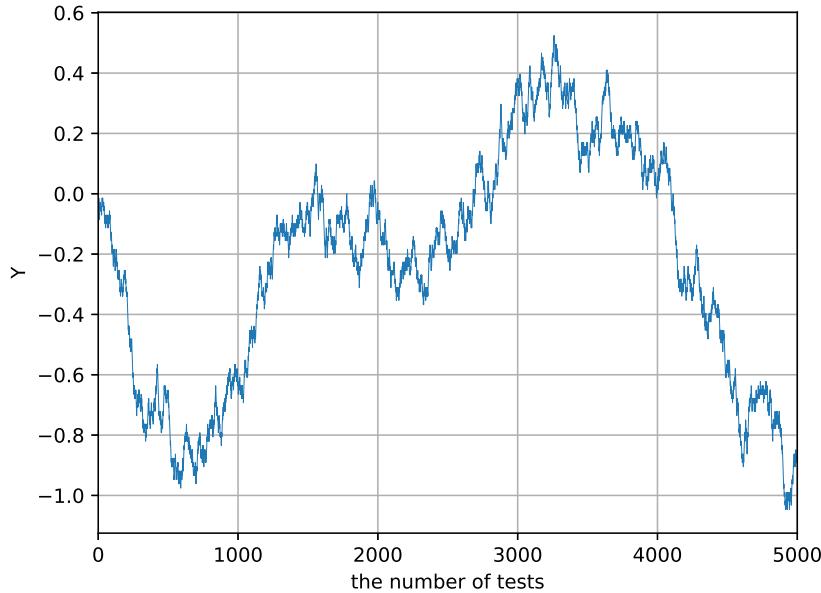


Рис. 6: Поведение нормированной суммы  $Y(i)$

## 2 Задание 2

### 2.1 Построение датчика сингулярного распределения

**Условие.** Построить датчик сингулярного распределения, имеющий в качестве функции распределения канторову лестницу. С помощью критерия Колмогорова убедиться в корректности работы датчика

**Определение 3.** Распределение называется сингулярным, если оно сосредоточено на континуальном множестве нулевой мерой Лебега.

Можем перейти в троичную систему счисления, в которой есть только цифры  $\{0, 1, 2\}$ , и числа из  $[0, 1]$  разложены по отрицательным степеням тройки. Канторово множество содержит числа, в которых нет цифры 1. Таким образом, для построения генератора такого распределения можем использовать генератор схемы Бернулли. Тогда, умножив на 2 вектор случайных величин из схемы Бернулли получим число в троичной записи, останется перевести его обратно в 10-ю систему счисления. Так как мы не можем сгенерировать бесконечное число, необходимо ввести дополнительный параметр  $\epsilon$ .

$$\sum_{n=N+1}^{+\infty} \frac{2 \cdot \eta_n}{3^n} \leq \sum_{n=N+1}^{+\infty} \frac{2}{3^n} = \frac{2 \cdot 3}{2 \cdot 3^N} = \frac{1}{3^N} \leq \epsilon.$$

Таким образом, можем получить длину дробной части числа, достаточную для получения заданной точности:

$$N \geq \frac{\ln \frac{1}{\epsilon}}{\ln 3}.$$

Значение функции распределения можем посчитать по формуле:

$$F(x) = \sum_{n=1}^N \frac{\eta_n}{2^n}, \quad x = \sum_{n=1}^N \frac{2 \cdot \eta_n}{3^n}.$$

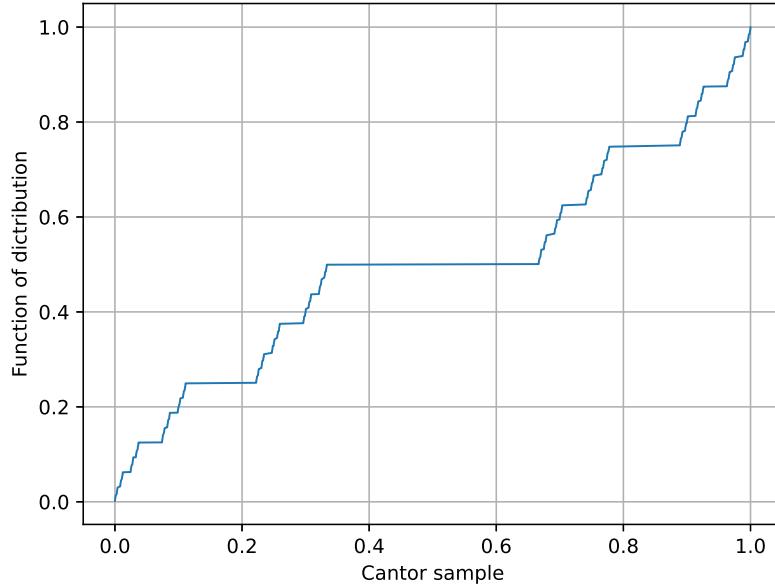


Рис. 7: Пример работы датчика сингулярного распределения с числом точек 1000 и точностью  $1e - 10$

**Определение 4.** Статистикой называется любая измеримая функция от выборки.

**Определение 5.** Под эмпирической функцией распределения  $F_n$ , построенной по выборке  $X = (X_1, \dots, X_n)$  понимают  $F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{X_i \leq x}$ , где

$$I_{X_i \leq x} = \begin{cases} 1, & X_i \leq x, \\ 0, & X_i > x. \end{cases}$$

Статистика критерия Колмогорова для эмпирической функции распределения  $F_n(x)$  определяется следующим образом:

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|.$$

Для проверки простой гипотезы  $H_0$  о соответствии выборки известному закону распределения с заданным уровнем значимости  $\alpha$  и правосторонней критической областью воспользоваться подходом на основе р-значения для статистики  $\sqrt{n}D_n$ . Нулевая гипотеза принимается,

если  $p = 1 - F_K(\sqrt{n}D_n)$  превышает заданный уровень значимости  $\alpha$ . Будем пользоваться следующим представлением функции распределения Колмогорова:

$$F_K(x) = 1 + 2 \sum_{k=1}^{+\infty} (-1)^k \exp^{-2k^2x^2}.$$

Проверка корректности датчика с помощью критерия Колмогорова		
Число испытаний	Уровень значимости	Частота принятия гипотезы
100	2%	0.96
1000	2%	0.983
10000	2%	0.989
100	10%	0.91
1000	10%	0.916

## 2.2 Использование критерия Смирнова

**Условие.** Для Канторовых случайных величин с помощью критерия Смирнова проверить свойство симметричности относительно  $\frac{1}{2}$  и свойство самоподобия относительно деления на  $3$

Будем рассматривать две выборки:

$$X = (X_1, X_2, \dots, X_n),$$

$$Y = (Y_1, Y_2, \dots, Y_m).$$

Нашей целью является проверка гипотезы о принадлежности выборок  $X$  и  $Y$  одному распределению. Критерий Смирнова определяется следующим образом: Статистика критерия для эмпирических функций распределения  $F_n(x)$  и  $G_m(x)$ , построенным по выборкам  $X$  и  $Y$  определяется следующим образом:

$$D_{n,m} = \sup_{x \in \mathbb{R}} |F_n - G_m|.$$

Воспользуемся p-value подходом для статистики  $\sqrt{\frac{nm}{n+m}} D_{n,m}$ . Нулевая гипотеза  $H_0$  принимается, если

$$p = 1 - F_K \left( \sqrt{\frac{nm}{n+m}} D_{n,m} \right) > \alpha.$$

Проверим свойство симметричности относительно  $\frac{1}{2}$ . Иначе это свойство можно переформулировать так: случайные величины  $X$  и  $1 - X$  имеют одинаковое распределение. Докажем это:

$$1 - X = 1 - \sum_{n=1}^{\infty} \frac{2\eta_n}{3^n} = \sum_{n=1}^{\infty} \frac{2}{3^n} - \sum_{n=1}^{\infty} \frac{2\eta_n}{3^n} = \sum_{n=1}^{\infty} \frac{2(1 - \eta_n)}{3^n} = \sum_{n=1}^{\infty} \frac{2\zeta_n}{3^n}.$$

Получили, что  $\zeta_n \sim Bern(0, 5)$ .

Проверка свойства симметричности с помощью критерия Смирнова		
Число испытаний	Уровень значимости	Частота принятия гипотезы
100	2%	0.97
1000	2%	0.95
10000	2%	0.9554
100	10%	0.88
1000	10%	0.875

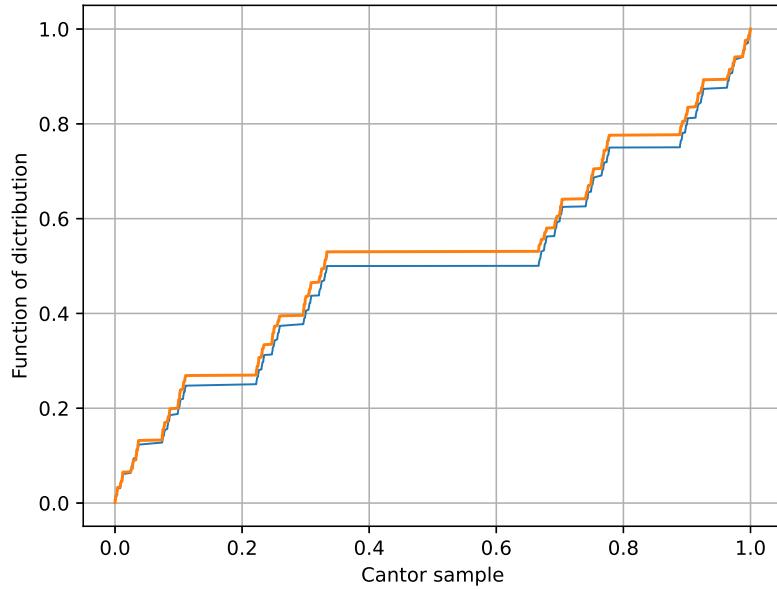


Рис. 8: Графическое подтверждение свойства симметричности

Проверим свойство самоподобия относительно деления на 3. Это свойство можно переформулировать так: условное распределение  $Y$  при условии  $Y \in [0, \frac{1}{3}]$  совпадает с распределением  $\frac{Y}{3}$ . Докажем это:

$$Y = \sum_{n=1}^{\infty} \frac{2\eta_{n+1}}{3^{n+1}} = \frac{1}{3} \sum_{n=1}^{\infty} \frac{2\eta_{n+1}}{3^n} = \frac{1}{3} \sum_{n=1}^{\infty} \frac{2\eta_n}{3^n} = \frac{Y}{3}.$$

Проверка свойства самоподобия с помощью критерия Смирнова		
Число испытаний	Уровень значимости	Частота принятия гипотезы
100	2%	0.99
1000	2%	0.984
10000	2%	0.9837
100	10%	0.95
1000	10%	0.941

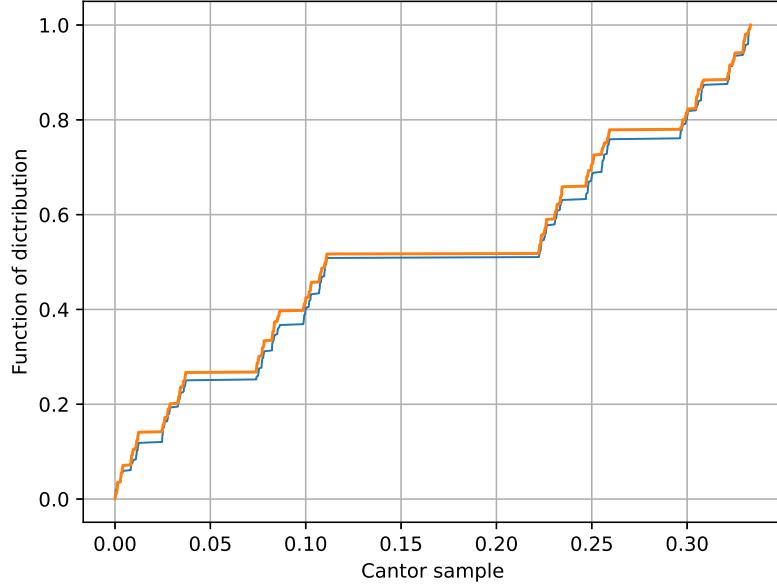


Рис. 9: Графическое подтверждение свойства самоподобия

### 2.3 Рассчёт математического ожидания и дисперсии

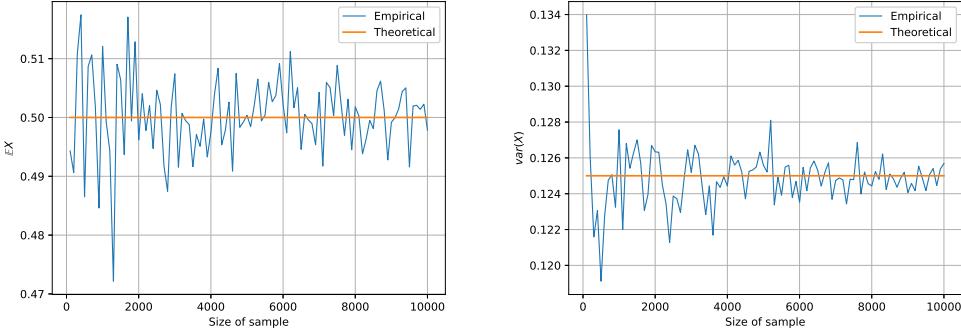
Пусть случайная величина  $X$  имеет сингулярное распределение, в качестве функции распределения - Канторова лестница. Посчитаем математическое ожидание и дисперсию для данного распределения.

$$\mathbb{E}X = \mathbb{E} \sum_{n=1}^{+\infty} \frac{2 \cdot \eta_n}{3^n} = \sum_{n=1}^{+\infty} \frac{2}{3^n} \mathbb{E}\eta_n = \sum_{n=1}^{+\infty} \frac{2}{3^n} (0 \cdot 0.5 + 1 \cdot 0.5) = \frac{1}{2},$$

$$var(X) = var \left( \sum_{n=1}^{+\infty} \frac{2 \cdot \eta_n}{3^n} \right) = \sum_{n=1}^{+\infty} \frac{4}{3^{2n}} \cdot var(\eta_n),$$

$$var(\eta) = \mathbb{E}\eta^2 - (\mathbb{E}\eta)^2 = (0 \cdot 0.5 + 1 \cdot 0.5) - (0 \cdot 0.5 + 1 \cdot 0.5)^2 = \frac{1}{4},$$

$$var(X) = \sum_{n=1}^{+\infty} \frac{4}{3^{2n}} \cdot \frac{1}{4} = \frac{1}{8}.$$



(a) Математическое ожидание для сингулярного распределения

(b) Дисперсия для сингулярного распределения

Рис. 10: Сравнение теоретических значений с эмпирическими

### 3 Задание 3

#### 3.1 Построение датчика экспоненциального распределения

**Условие.** Построить датчик экспоненциального распределения. Проверить для этого распределения свойство отсутствия памяти.

**Определение 6.** Случайная величина  $X$  имеет экспоненциальное распределение с параметром  $\lambda$ , если её функция распределения имеет вид:

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x > 0. \end{cases}$$

В данном случае можем воспользоваться методом обратных функций.

$$y = 1 - e^{-\lambda x},$$

$$-\lambda x = \ln(1 - y),$$

$$x = -\frac{\ln(1 - y)}{\lambda}.$$

Таким образом, случайная величина  $X$  имеет экспоненциальное распределение с параметром  $\lambda$ , если  $X = -\frac{\ln(1 - Y)}{\lambda}$ , где  $Y \sim U[0, 1]$ .

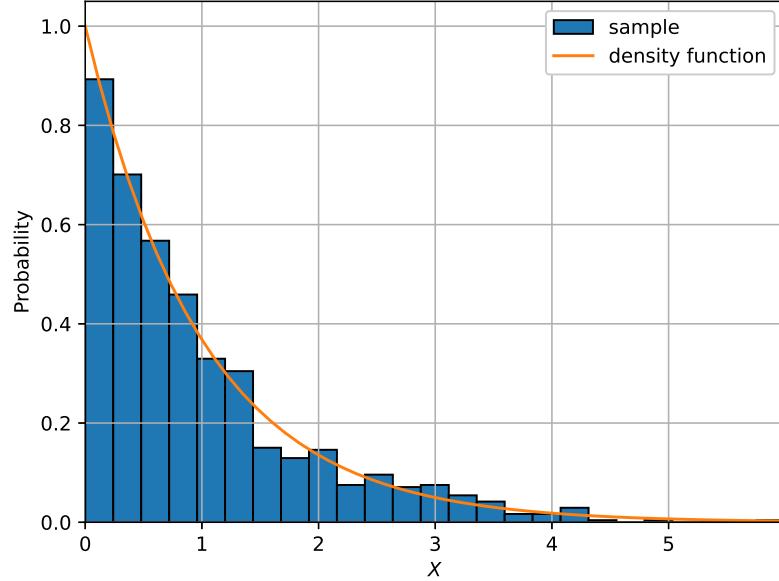


Рис. 11: Результат работы датчика экспоненциального распределения с параметром  $\lambda = 1$  и объёмом выборки 1000

**Свойство 2.** Если  $X \sim \text{Exp}(\lambda)$ , то  $\forall t, \Delta t \in \mathbb{Z}_+$

$$\mathbb{P}(X > t + \Delta t | X \geq t) = \mathbb{P}(X > \Delta t).$$

*Доказательство.* Воспользуемся формулой условной вероятности и монотонностью меры  $\mathbb{P}$  по вложенным событиям.

$$\begin{aligned} \mathbb{P}(X > t + \Delta t | X \geq t) &= \frac{\mathbb{P}(X > t + \Delta t)}{\mathbb{P}(X \geq t)} = \\ &= \left\{ \mathbb{P}(X \leq x) = 1 - \mathbb{P}(X > x) = 1 - F_X(x+0) = 1 - e^{-\lambda x} \right\} = \\ &= \frac{e^{-\lambda(t+\Delta t)}}{e^{-\lambda t}} = e^{-\lambda \Delta t} = \mathbb{P}(X > \Delta t). \end{aligned}$$

□

Это свойство экспоненциального распределения называют свойством отсутствия памяти.

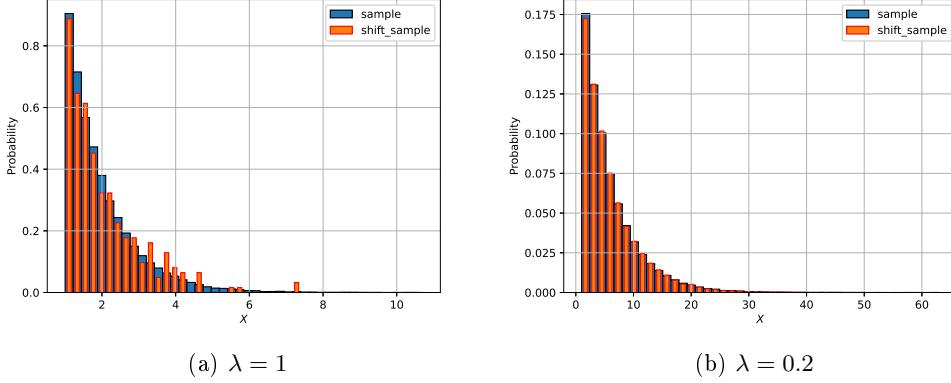


Рис. 12: Свойство отсутствия памяти с объёмом выборки 100000

### 3.2 Распределение случайной величины $Y = \min(X_1, \dots, X_n)$

**Условие.** Пусть  $X_1, X_2, \dots, X_n$  - независимые экспоненциально распределённые случайные величины параметрами  $\lambda_1, \dots, \lambda_n$ . Найти распределение случайной величины  $Y = \min(X_1, \dots, X_n)$ .

По определению:

$$\mathbb{P}(Y < x) = 1 - \mathbb{P}(Y \geq x) = 1 - \mathbb{P}(X_1 \geq x, \dots, X_n \geq x).$$

По условию  $X_1, \dots, X_n$  - независимые случайные величины. Значит можем продолжить равенство:

$$\begin{aligned} \mathbb{P}(Y < x) &= 1 - \mathbb{P}(X_1 \geq x) \cdot \dots \cdot \mathbb{P}(X_n \geq x) = 1 - \prod_{i=1}^n (1 - (1 - e^{-\lambda_i x})) = 1 - \prod_{i=1}^n (e^{-\lambda_i x}) = \\ &= 1 - e^{\sum_{i=1}^n -\lambda_i x} = 1 - e^{-x \cdot \sum_{i=1}^n \lambda_i}. \end{aligned}$$

Таким образом, получили, что случайная величина  $Y$  имеет экспоненциальное распределение с параметром  $\sum_{i=1}^n \lambda_i$ .

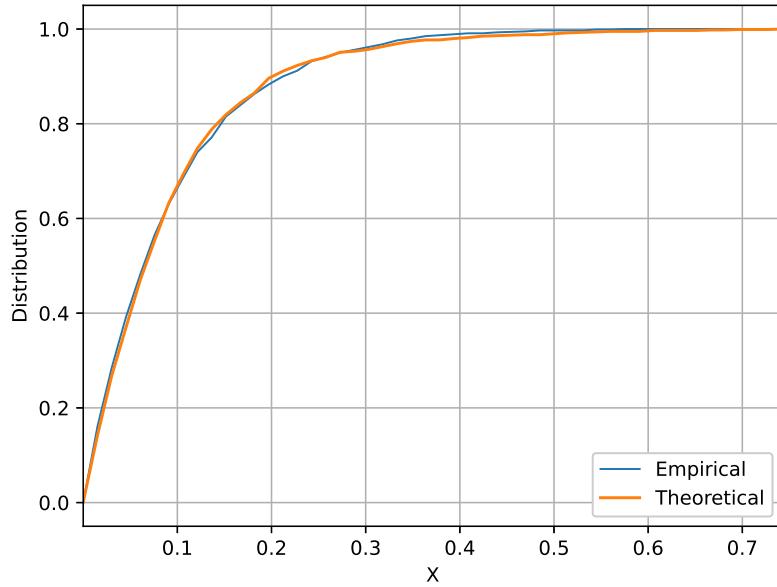


Рис. 13: Графическое подтверждение экспоненциальной распределённости  $Y$  с объёмом выборки 1000

### 3.3 Построение датчика пуассоновского распределения 1

**Условие.** На основе датчика экспоненциального распределения построить датчик пуассоновского распределения

**Определение 7.** Случайная величина  $X$  имеет пуассоновское распределение с параметром  $\lambda$ , если

$$\mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}.$$

**Утверждение 1.** Пусть  $\eta_1, \dots, \eta_n, \dots$  независимые случайные величины, имеющие показательное распределение с параметром  $\lambda$  и

$$X = \max\{n : \eta_1 + \dots + \eta_n \leq 1\}.$$

Тогда случайная величина  $X$  имеет распределение Пуассона с параметром  $\lambda$ .

Доказательство можно найти в [2].

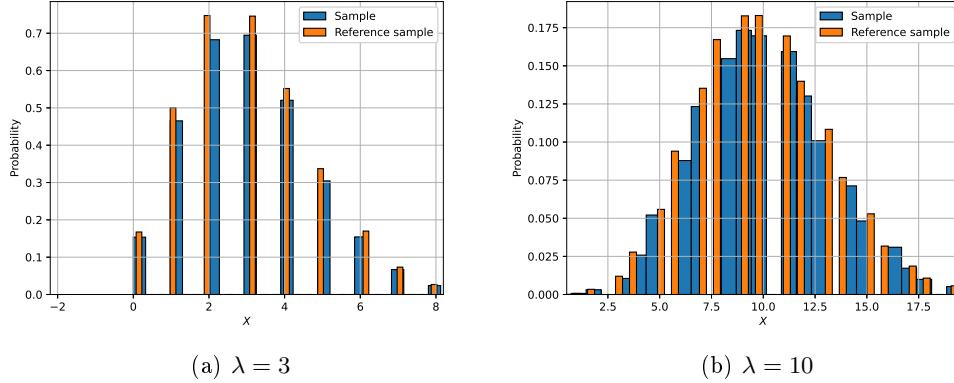


Рис. 14: Распределение Пуассона, построенное на основе экспоненциального распределения, при генерации 100000 величин

### 3.4 Построение датчика пуассоновского распределения 2

**Условие.** Построить датчик пуассоновского распределения как предел биномиального распределения. Убедиться в корректности построенного датчика при помощи критерия  $\chi^2$  Пирсона.

**Утверждение 2.** Введём обозначения:

$P$  - пуассоновское распределение,

$P_n$  - биномиальное распределение.

Пусть параметры биномиального распределение изменяются следующим образом:

$$n \rightarrow +\infty, \quad p \rightarrow 0, \quad np \rightarrow \lambda$$

Тогда

$$\sup_A |P(A) - P_n(A)| \rightarrow 0, \quad n \rightarrow +\infty.$$

Доказательство можно найти в [3].

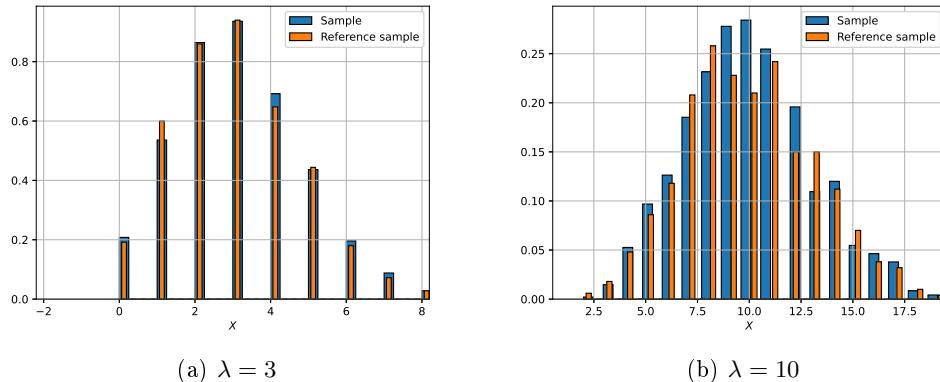


Рис. 15: Распределение Пуассона, построенное на основе биномиального распределения, при генерации 1000 величин

Критерий согласия Пирсона — непараметрический метод, который позволяет оценить значимость различий между фактическим и теоретическим количеством исходов или качественных характеристик выборки. Важно, что этот критерий имеет правостороннюю критическую область. Нулевая гипотеза заключается в том, что частоты согласованы, то есть фактические данные не противоречат ожидаемым. Теоретическая вероятность равна:

$$p_k = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = [a_0, \dots, a_s].$$

Теоретические частоты равны  $n_k^w = n \cdot p_k$ .

Считаем эмпирические частоты  $n_k$  как количество величин в выборке, принимающих значение  $k$ . Статистику  $\chi^2$  определим следующим образом:

$$\chi^2 = \sum_{k=a_0}^{a_s} \frac{(n_k - n_k^w)^2}{n_k^w}.$$

Выберем разбиение  $\delta$ :  $\delta_1 = \{0, 1\}$ ,  $\delta_2 = \{3\}$ ,  $\delta_3 = \{4\}$ ,  $\dots$ ,  $\delta_8 = \{9\}$ ,  $\delta_{10} = \{k > 9\}$ . Таким образом, получим 9 степеней свободы.

Принимаем нулевую гипотезу, если  $\chi^2 \leq q$ , где  $q$  — квантиль распределения  $\chi^2$ .

Проверка гипотезы по критерию Пирсона		
Число испытаний	Уровень значимости	Частота принятия гипотезы
100	20%	0.8
20	20%	0.79
100	2%	0.948
20	2%	0.95

### 3.5 Построение датчика стандартного нормального распределения

**Условие.** Построить датчик стандартного нормального распределения методом моделирования случайных величин парами с переходом в полярные координаты. Проверить при помощи  $t$ -критерия Стьюдента равенство математических ожиданий, а при помощи критерия Фишера — равенство дисперсий.

Рассмотрим пару

$$\eta \sim N(0, 1), \quad \xi \sim N(0, 1).$$

Тогда их совместная плотность распределения:

$$\mathbb{P}(\eta < x, \xi < y) = \frac{1}{2\pi} \int_{-\infty}^x \int_{-\infty}^y e^{-\frac{x_1^2 + x_2^2}{2}} dx_1 dx_2 = \frac{1}{4\pi} \int \int_{\omega \geq 0, \phi \in (0, 2\pi]} e^{-\frac{\omega}{2}} d\omega d\phi.$$

Таким образом,

$$\omega \sim Exp(0.5), \quad \phi \sim U(0, 2\pi),$$

$$\mathbb{P}(\eta < x, \xi < y) = \mathbb{P}(\sqrt{\omega} \cos \phi < x, \sqrt{\omega} \sin \phi < y).$$

То есть можем построить стандартные нормальные величины с помощью уже построенных датчиков.

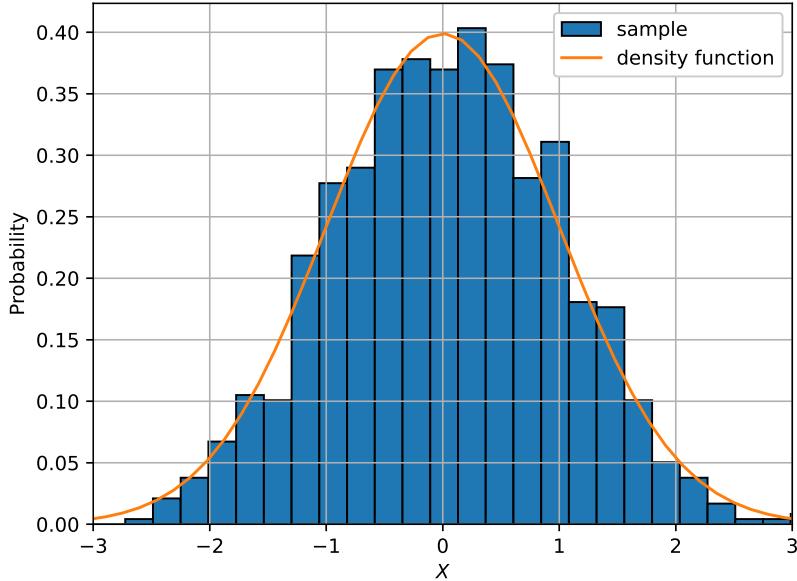


Рис. 16: Датчик нормального распределения с объёмом выборки 1000

Рассматриваем выборку  $X_1, \dots, X_n$  из нормального распределения. Статистику в т-критерии Стьюдента определяем следующим образом:

$$t = \sqrt{n} \frac{\bar{X} - \mu}{S}.$$

Здесь  $S^2$  - несмешённая выборочная дисперсия. Степеней свободы  $n - 1$ . Нулевая гипотеза заключается в том, что математическое ожидание полученной выборки равно единице. Критерий Стьюдента имеет двустороннюю критическую область, следовательно нулевая гипотеза принимается, если  $|t| \leq q$ , где  $q_{1-\frac{\alpha}{2}}$  - квантиль распределения с  $n - 1$  степенью свободы.

Проверка гипотезы по критерию Стьюдента		
Число испытаний	Уровень значимости	Частота принятия гипотезы
100	2%	0.98
1000	2%	0.981
10000	2%	0.9786
100	10%	0.92
1000	10%	0.897

В критерии Фишера можем сравнивать между собой только две выборки, поэтому в качестве одной возьмем выборку, сгенерированную построенным датчиком  $X_1, \dots, X_n$ , а в качестве второй —  $Y_1, \dots, Y_m$ , полученную с помощью встроенной функции. Статистика определяется следующим образом:

$$F = \frac{S_X^2}{S_Y^2}$$

со степенями свободы  $m - 1, n - 1$ . Критерий Фишера имеет двустороннюю критическую область, следовательно нулевая гипотеза принимается  $q_{\frac{\alpha}{2}}^1 \leq F \leq q_{1-\frac{\alpha}{2}}^2$ , где  $q_1, q_2$  - квантили со степенями свободы  $n - 1, m - 1$ .

Проверка гипотезы по критерию Фишера		
Число испытаний	Уровень значимости	Частота принятия гипотезы
100	2%	0.98
1000	2%	0.975
10000	2%	0.9802
100	10%	0.92
1000	10%	0.911

## 4 Задание 4

### 4.1 Построение датчика распределения Коши

Случайная величина  $X$  имеет распределение Коши с параметрами  $x_0$  и  $\gamma$ , если её функция распределения имеет вид

$$F_X(x) = \frac{1}{\pi} \arctan \frac{x - x_0}{\gamma} + \frac{1}{2}.$$

В данном случае можем воспользоваться методом обратных функций:

$$y = \frac{1}{\pi} \arctan \frac{x - x_0}{\gamma} + \frac{1}{2},$$

$$\tan \left( \pi \left( y - \frac{1}{2} \right) \right) = \frac{x - x_0}{\gamma},$$

$$x = x_0 + \gamma \tan \left( \pi \left( y - \frac{1}{2} \right) \right).$$

Таким образом, случайная величина  $X$  имеет распределение Коши с параметрами  $x_0, \gamma$ , если  $X = x_0 + \gamma \tan \left( \pi \left( Y - \frac{1}{2} \right) \right)$ , где  $Y \sim U [0, 1]$ .

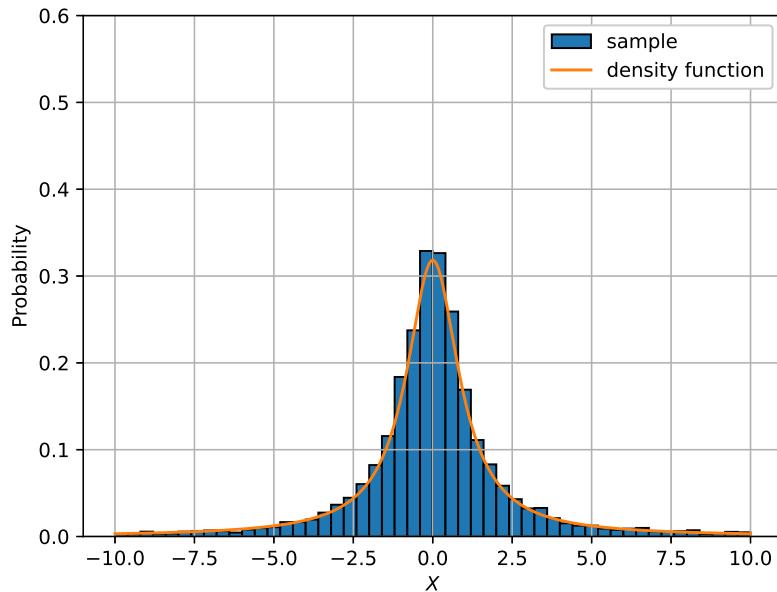


Рис. 17: Датчик распределения Коши с параметрами  $x_0 = 0, \gamma = 1$  и объёмом выборки 10000

## 4.2 Применение метода фон Неймана

**Условие.** На основе датчика распределения Коши с помощью метода фон Неймана построить датчик стандартного нормального распределения. При помощи графика *normal probability plot* убедиться в корректности построенного датчика и обосновать наблюдаемую линейную зависимость.

Метод фон Неймана также известен как метод элиминации. Он применим, если для имеющихся плотностей  $p_1(x)$  и  $p_2(x)$  двух распределений можно подобрать такой константный множитель  $k \geq 1$ , что

$$kp_1(x) \geq p_2(x), \quad \forall x.$$

В нашей задаче имеются две функции плотности распределения:  $p_1$  - плотность распределения Коши,  $p_2$  - плотность стандартного нормального распределения.

$$p_1(x) = \frac{1}{\pi} \cdot \frac{\gamma}{(x - x_0)^2 + \gamma^2}.$$

В нашем случае

$$p_1(x) = \frac{1}{\pi} \cdot \frac{1}{(x)^2 + 1},$$

$$p_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Сформулируем алгоритм построения случайной с плотностью распределения  $p_2$ :

1. Генерируем случайную величину  $X$ , соответствующую распределению с плотностью  $p_1$ ;
2. Вычисляем значение  $R = \frac{p_2(X)}{kp_1(X)}$ ;
3. Генерируем случайную величину  $\eta \sim Be(R)$ . Если  $\eta = 1$ , то  $X$  — величина, соответствующая распределению с плотностью  $p_2$ . Иначе продолжем алгоритм в цикле. Утверждается, что наилучшим выбором константы является:

$$k = \sqrt{\frac{2\pi}{e}}.$$

Полное обоснование можно найти в [4].

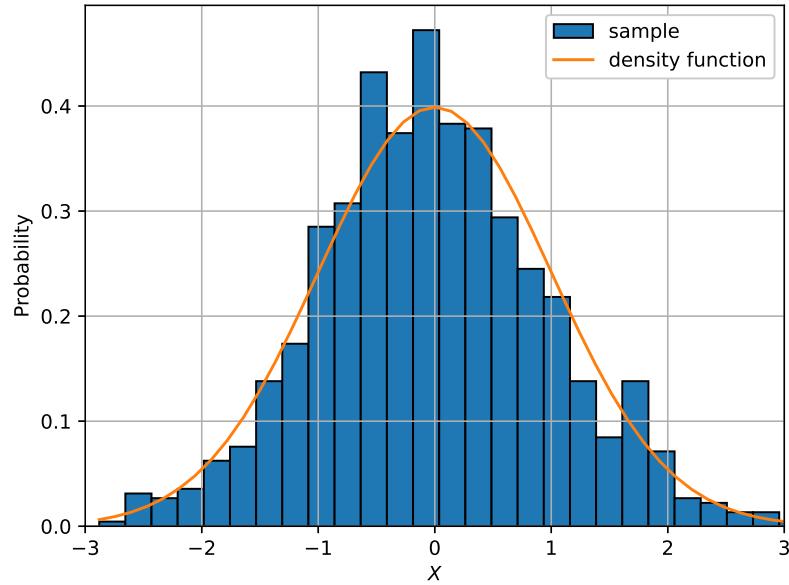


Рис. 18: Датчик стандартного нормального распределения с объёмом выборки 1000

Разберёмся в принципе работы функции `normal probability plot`. Функция сопоставляет квантили поданной на вход выборки с квантилями нормального распределения. Выборка, соответствующая нормальному распределению накладывается на заранее сгенерированные квантили нормального распределения, поэтому график получается линейным для нормального распределения, а для любого другого распределения график линейным не будет. Заметим, что графики будут инвариантны относительно параметров масштаба и сдвига. Пусть  $x_i$  из нормального стандартного распределения, тогда  $y_i = \mu + \sigma x_i$  из нормального распределения с параметрами  $\mu$  и  $\sigma$ . Видно, что при подаче в функцию выборки  $y_i$  происходит пересчёт квантилей и график не изменяется визуально, то есть остаётся линейным.

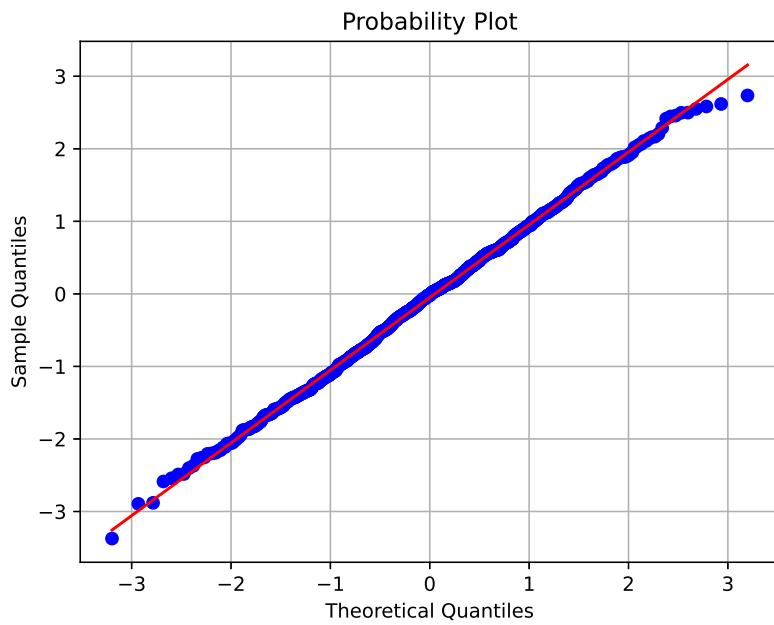


Рис. 19: Подтверждение корректности построенного датчика нормального распределения

#### 4.3 Сравнение скорости моделирования стандартного нормального распределения

**Условие.** Сравнить скорость моделирования стандартного нормального распределения в задании 3.5 и в задании 4.2.

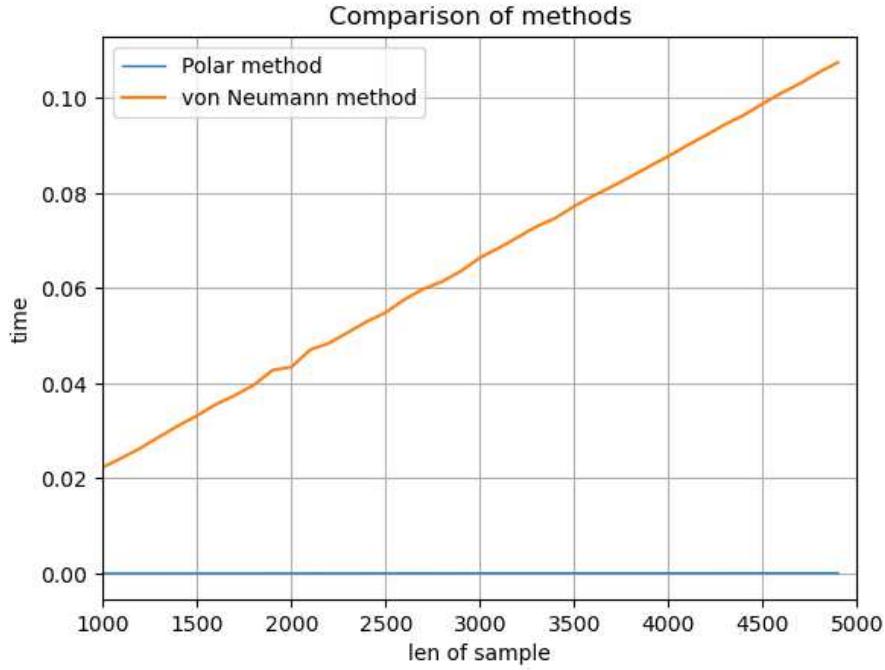


Рис. 20: Сравнение времени работы датчиков нормального распределения

Из Рисунка ?? видим, что генерация с помощью преобразования Бокса-Мллера занимает значительно меньше времени. В дальнейшем, будем использовать только этот датчик для генерации стандартного нормального распределения.

## 5 Задание 5

### 5.1 Справедливость ЗБЧ и ЦПТ

**Условие.** Пусть  $X_i \sim N(\mu, \sigma^2)$ . Убедиться эмпирически в справедливости теоремы о законе больших чисел и центральной предельной теоремы: исследовать поведение суммы  $S_n = \sum_{i=1}^n X_i$  и эмпирического распределения величины  $\sqrt{n} \frac{S_n}{n} - a$ .

**Теорема 2** (Закон больших чисел). Пусть Пусть  $X_1, X_2, \dots$  - последовательность независимых одинаково распределённых случайных величин с  $EX_1^2 < \infty$  и  $S_n = X_1 + \dots + X_n$ . Тогда  $\forall \epsilon > 0$  справедливо

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \left| \frac{S_n}{n} - \mu \right| < \epsilon \right) = 1.$$

Доказательство можно найти в [5].

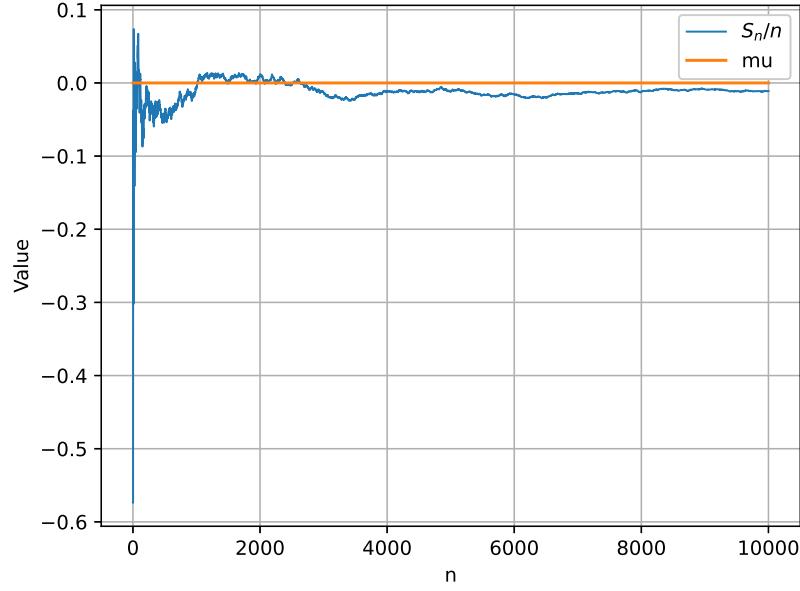


Рис. 21: Справедливость ЗБЧ

На Рис. ?? видим сходимость  $\frac{S_n}{n}$  к  $\mu$ , что подтверждает справедливость ЗБЧ.

Теперь исследуем поведение случайной величины  $\sqrt{n} \frac{S_n}{n} - a$ .

$$\mathbb{E} \frac{S_n}{n} = \frac{1}{n} \mathbb{E} \sum_{i=1}^n X_i = \frac{n \cdot \mu}{n} = \mu,$$

$$var \left( \frac{S_n}{n} \right) = \frac{1}{n^2} var(S_n) = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Из Теоремы 1 получим, что  $\frac{S_n}{n} \sim N \left( \mu, \frac{\sigma^2}{n} \right)$ . Теперь рассмотрим случайную величину  $Y = \sqrt{n} \frac{S_n}{n} - a$ .

$$\mathbb{E} Y = \sqrt{n} \frac{\mathbb{E} S_n}{\sqrt{n}} - \mathbb{E} a = \sqrt{n} \cdot \mu - a,$$

$$var(Y) = n \cdot var \left( \frac{S_n}{n} \right) - var(a) = \frac{n \cdot \sigma^2}{n} = \sigma^2.$$

Таким образом,  $Y \sim N(\sqrt{n} \cdot \mu - a, \sigma^2)$ .

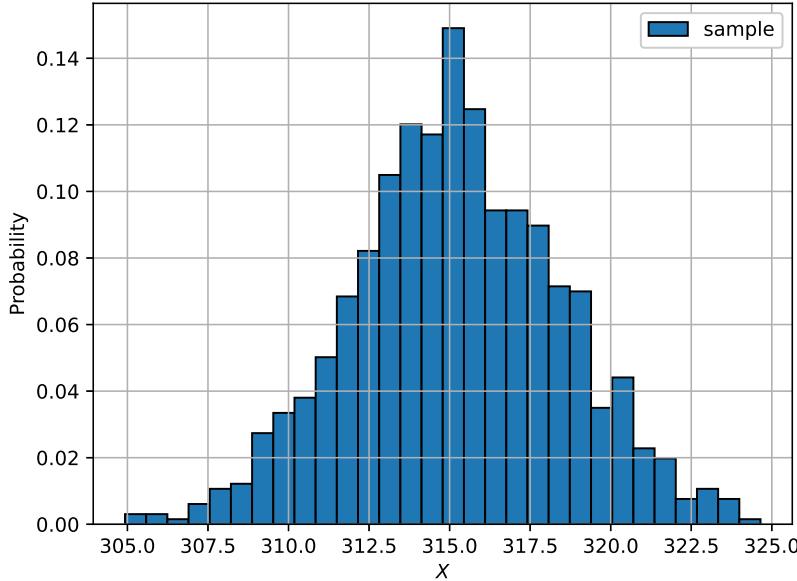


Рис. 22: Справедливость ЦПТ

Из Рис. ?? видим, что случайная величина  $\bar{Y}$  действительно из нормального распределения.

## 5.2 Построение доверительных интервалов для среднего и дисперсии

**Условие.** Считая  $\mu, \sigma$  неизвестными, построить доверительные интервалы для среднего и дисперсии по имеющейся выборке.

Доверительные интервалы для математического ожидания будем строить, используя t-критерий Стьюдента.

$$t = \sqrt{n} \frac{\bar{X} - \mu}{S}.$$

Пусть  $q(\alpha, n - 1)$  - квантили порядка  $\alpha$  для распределения Стьюдента с числом степеней свободы  $n - 1$ . Тогда доверительный интервал можем строить следующим образом:

$$\begin{aligned} \mathbb{P}\left(q\left(\frac{1-\alpha}{2}, n-1\right) \leq t \leq q\left(\frac{1+\alpha}{2}, n-1\right)\right) &= \alpha, \\ \mathbb{P}\left(-q\left(\frac{1+\alpha}{2}, n-1\right) \leq t \leq q\left(\frac{1+\alpha}{2}, n-1\right)\right) &= \alpha, \\ \mathbb{P}\left(-q\left(\frac{1+\alpha}{2}, n-1\right) \leq \sqrt{n} \frac{\bar{X} - \mu}{S} \leq q\left(\frac{1+\alpha}{2}, n-1\right)\right) &= \alpha, \\ \mathbb{P}\left(\bar{X} - \frac{S}{\sqrt{n}} q\left(\frac{1+\alpha}{2}, n-1\right) \leq \mu \leq \bar{X} + \frac{S}{\sqrt{n}} q\left(\frac{1+\alpha}{2}, n-1\right)\right) &= \alpha. \end{aligned}$$

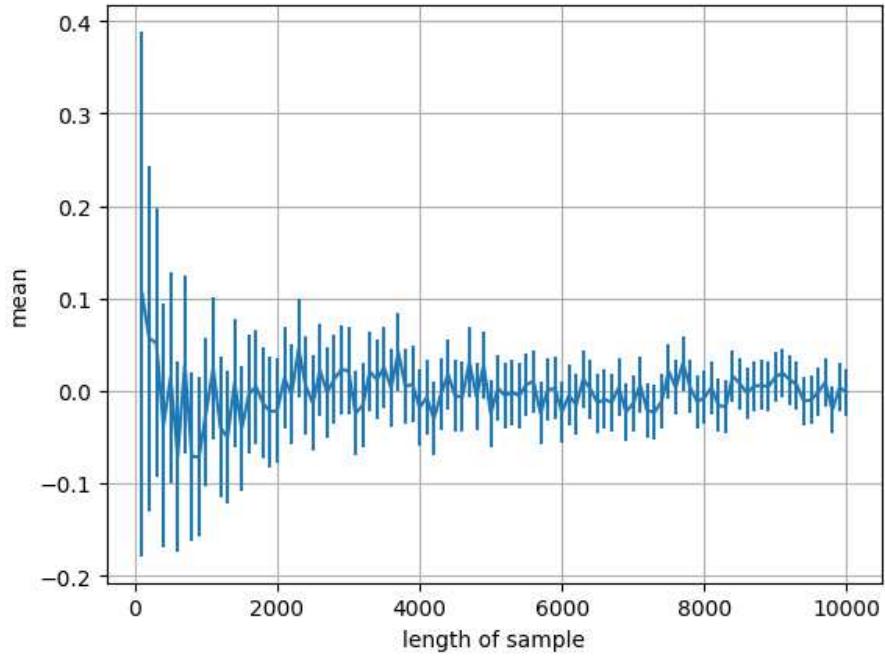


Рис. 23: Доверительный интервал при  $\alpha = 0.8$  и  $\mu = 0$

Аналогичным способом поступим для поиска доверительных интервалов дисперсии. Используем критерий  $\chi^2$ :

$$\mathbb{P} \left( q \left( \frac{1-\alpha}{2}, n-1 \right) \leq \frac{(n-1) \cdot S^2}{\sigma^2} \leq q \left( \frac{1+\alpha}{2}, n-1 \right) \right) = \alpha,$$

$$\mathbb{P} \left( \frac{(n-1) \cdot S^2}{q \left( \frac{1+\alpha}{2}, n-1 \right)} \leq \sigma^2 \leq \frac{(n-1) \cdot S^2}{q \left( \frac{1-\alpha}{2}, n-1 \right)} \right) = \alpha.$$

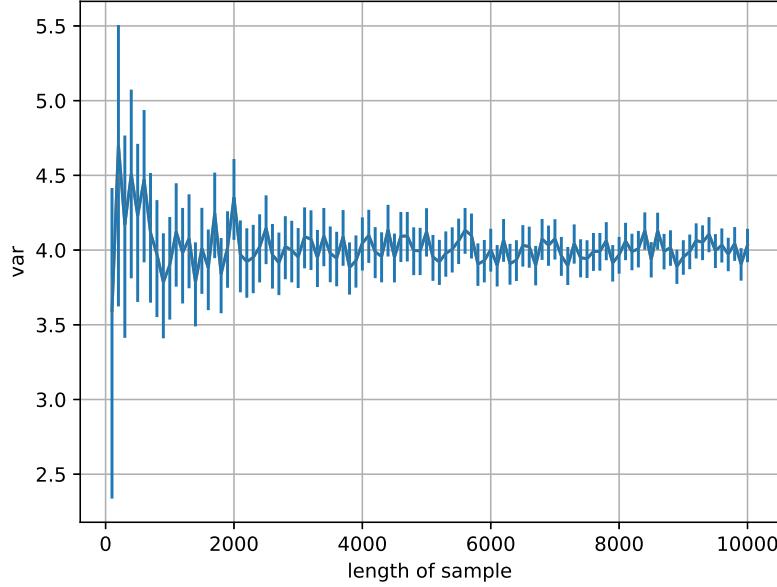


Рис. 24: Доверительный интервал при  $\alpha = 0.8$  и  $\mu = 0$

### 5.3 Распределение суммы элементов из распределения Коши

**Условие.** Пусть  $X_i \sim K(a, b)$  - имеет распределение Коши с параметром сдвига  $a$  и масштаба  $b$ . Изучить эмпирически как ведут себя суммы  $\frac{S_n}{n}$ , объяснить результат и найти закон распределения данных сумм.

Доказательство проведём с помощью характеристических функций. Как известно характеристическая функция случайной величины  $X \sim K(a, b)$  равна:

$$\phi_X(t) = e^{iat - b|t|}.$$

Рассмотрим характеристическую функцию для случайной величины  $Y = \frac{S_n}{n}$ :

$$\phi_Y(t) = \mathbb{E}e^{itY} = \mathbb{E}e^{\frac{it\sum_{i=1}^n X_i}{n}} = \left(\mathbb{E}e^{\frac{it}{n}X}\right)^n = \phi_X^n\left(\frac{t}{n}\right) = e^{iat - b|\frac{t}{n}|}.$$

Таким образом доказано, что  $Y \sim K(a, b)$ . Отсюда видно, что не существует ЗБЧ для любой взвешенной суммы независимых распределений Коши. Это связано с тем, что для распределения Коши не определены конечные первые моменты.

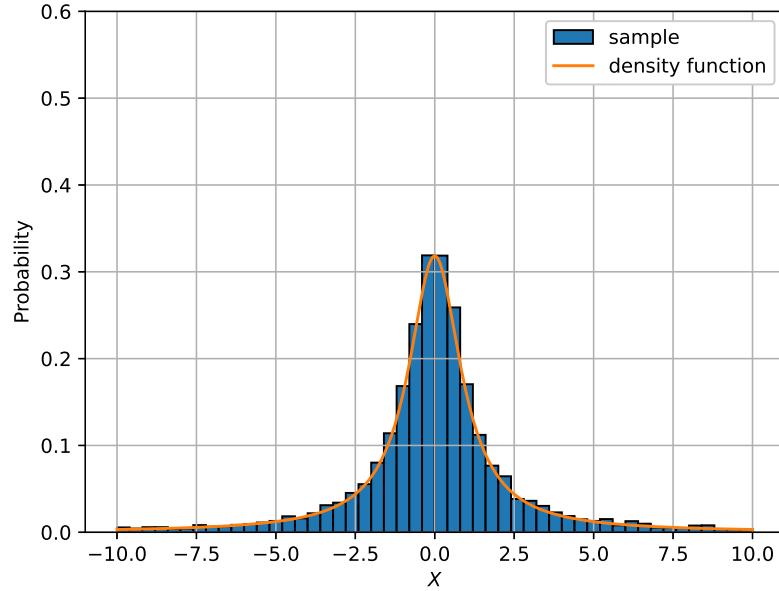


Рис. 25: Распределение суммы элементов из распределения Коши

## 6 Задание 6

### 6.1 Вычисление интеграла

**Условие.** Вычислить следующий интеграл:

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{e^{-\left(x_1^2 + \dots + x_{10}^2 + \frac{1}{27 \cdot x_1^2 \dots x_{10}^2}\right)}}{x_1^2 \dots x_{10}^2} dx_1 dx_2 \dots dx_{10}$$

методом Монте-Карло и методом квадратур. Оценить точность для каждого случая.

Представим интеграл в виде:

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \pi^5 \frac{e^{-\frac{1}{27 \cdot x_1^2 \dots x_{10}^2}}}{x_1^2 \dots x_{10}^2} \cdot \frac{1}{\pi^5} e^{-(x_1^2 + \dots + x_{10}^2)} dx_1 dx_2 \dots dx_{10}.$$

Введём обозначения:

$$f = \pi^5 \frac{e^{-\frac{1}{27 \cdot x_1^2 \dots x_{10}^2}}}{x_1^2 \dots x_{10}^2},$$

$$g = \frac{1}{\pi^5} e^{-(x_1^2 + \dots + x_{10}^2)} = \frac{e^{-\frac{(x_1 - 0)^2}{1} - \dots - \frac{(x_{10} - 0)^2}{1}}}{\sqrt{\pi} \dots \sqrt{\pi}}.$$

Таким образом, функция  $g$  - совместная плотность независимых случайных величин  $x_1, \dots, x_{10}$ , где  $x_i \sim N\left(0, \frac{1}{27}\right)$ . Таким образом, интеграл можно представить в виде:

$$\mathbb{E}f(x_1, \dots, x_{10}), \quad x_i \sim N\left(0, \frac{1}{27}\right).$$

Результаты вычисления интеграла методом Монте-Карло	
Число точек	Результат
$10^3$	129.677
$10^4$	125.179
$10^5$	115.105
$10^6$	126.035

Используя Теорему 2, можем находить интеграл как

$$\frac{S_n}{n}, \quad S_n = \sum_{i_j=1,\dots,n} f(x_{i_1}, \dots, x_{i_1 0}).$$

Заметим, что данный метод является достаточно медленным, но универсальным.

Оценим погрешность метода Монте-Карло, используя Теорему 1.

$$\begin{aligned} \mathbb{P}\left(\left|\frac{S_n}{n} - I\right| < \varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n - I \cdot n}{n}\right| < \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - I \cdot n}{\sigma \cdot \sqrt{n}}\right| < \frac{\sqrt{n}}{\sigma} \varepsilon\right) = \\ &= \mathbb{P}\left(\frac{S_n - I \cdot n}{\sigma \cdot \sqrt{n}} < \frac{\sqrt{n}}{\sigma} \varepsilon\right) - \mathbb{P}\left(\frac{S_n - I \cdot n}{\sigma \cdot \sqrt{n}} \leq -\frac{\sqrt{n}}{\sigma} \varepsilon\right) \xrightarrow{n \rightarrow \infty} \Phi_0\left(\frac{\sqrt{n}}{\sigma} \varepsilon\right) - \Phi_0\left(-\frac{\sqrt{n}}{\sigma} \varepsilon\right) = \\ &= \Phi_0\left(\frac{\sqrt{n}}{\sigma} \varepsilon\right) + \Phi_0\left(\frac{\sqrt{n}}{\sigma} \varepsilon\right) = 2\Phi_0\left(\frac{\sqrt{n}}{\sigma} \varepsilon\right), \quad \text{где } \Phi_0(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^x e^{-\frac{u^2}{2}} du \end{aligned}$$

Посчитаем интеграл методом квадратур. Видно что подинтегральная функция  $h$  чётная, значит  $I = 2^{10} \int_0^{+\infty} \dots \int_0^{+\infty} h$ . Далее сделаем замену

$$x_i = \operatorname{tg} \frac{\pi}{2} t_i, \quad t_i \in (0, 1).$$

Тогда интеграл преобразуется к виду

$$I = \pi^{10} \int_0^1 \dots \int_0^1 \frac{\exp\{-\sum_{i=1}^{10} \operatorname{tg}^2\left(\frac{\pi}{2} t_i\right) - \frac{1}{2^7 \prod_{i=1}^{10} \operatorname{tg}^2\left(\frac{\pi}{2} t_i\right)}\}}{\prod_{i=1}^{10} \sin^2\left(\frac{\pi}{2} t_i\right)} dt_1 \dots dt_{10}.$$

В качестве метода квадратур выберем метод прямоугольников. Для этого необходимо разбить отрезки  $[0, 1]$  на  $N$  равных частей:

$$0 = x_k^0 < x_k^2 < \dots < x_k^{2N} = 1, \quad k = 1, \dots, 10.$$

У каждого отрезка в разбиении вычислим середину:

$$x_k^1, x_k^3, \dots, x_k^{2N-1}, \quad k = 1, \dots, 10.$$

В этих точках будем считать подинтегральную функцию. Тогда искомый интеграл можем вычислять по формуле

$$I = \left(\frac{\pi}{N}\right)^{10} \sum_{i_1=1}^N \dots \sum_{i_{10}=1}^N h(x_1^{2i_1-1}, x_2^{2i_2-1}, \dots, x_{10}^{2i_{10}-1}).$$

Заметим, что значение интеграла не зависит от порядка интегрирования, поэтому для ускорения работы функции интеграл представим в виде:

$$I = \left(\frac{\pi}{N}\right)^{10} \sum_{1 \leq i_0 \leq i_1 \leq \dots \leq i_{10} \leq N}^{m_1, \dots, m_l} h(x_1^{2i_1-1}, x_2^{2i_2-1}, \dots, x_{10}^{2i_{10}-1}).$$

Результаты вычисления интеграла методом квадратур	
Число точек	Результат
$10^3$	270.356
$10^4$	213.600
$10^5$	129.362
$10^6$	116.488

Здесь

$$\frac{m}{m_1, \dots, m_l} = \frac{10!}{m_1!m_2!\dots m_l!},$$

где  $m_k!$  - количество индексов в наборе  $(i_1, i_2, \dots, i_{10})$ , равных  $k$ .

Вид погрешности для метода прямоугольников известен, поэтому можем записать

$$|R| \leq \frac{1}{24N^2} \max_{\eta_i \in [0,1]} |f''(\eta_i)|, \quad \forall i \in 1, \dots, 10.$$

## 7 Задание 7

### 7.1 Метод случайного поиска

**Условие.** Методом случайного поиска найти минимальное значение функции  $f$  на множестве  $A = \{x_1, x_2 : x_1^2 + x_2^2 \leq 1\}$ , где

$$f(x_1, x_2) = x_1^3 \sin \frac{1}{x_1} + 10x_1 x_2^4 \cos \frac{1}{x_2}, \quad x_1, x_2 \neq 0.$$

При  $x_1 = 0$  или  $x_2 = 0$  функции доопределяются непрерывности

В ходе метода случайного поиска произвольно выбираются точки из рассматриваемой области, и в каждой точке рассчитывается значение целевой функции. В качестве решения принимается точка, в которой значение функции минимально.

Сначала необходимо понять, как выбрать случайную точку в единичной окружности. Далее предложен способ, с помощью которого подходящая точка гарантированно выбирается с первого раза. Перейдём к полярным координатам:

$$\begin{cases} x_1 = r \cdot \cos \phi, \\ x_2 = r \cdot \sin \phi. \end{cases}$$

Если положить  $r \sim U[0, 1]$ ,  $\phi \sim U[0, 2\pi]$ , не получим равномерное распределение (в центре окружности будет содержаться большее число точек). Это обусловлено тем, что каждое кольцо круга будет содержать примерно одинаковое число точек, то есть проблема в распределении  $r$ . Так как длина окружности линейно зависит от радиуса, для сохранения одинаковой плотности точек их количество должно тоже зависеть от радиуса. Тогда пусть функция плотности вероятности  $f(r) = mr$ . Найдем значение  $m$ :

$$\int_0^1 mr dr = 1 \Rightarrow f(r) = 2r.$$

Тогда функция распределения

$$F_X(r) = \int_0^r f(x) dx = \int_0^r 2x dx = r^2 = \mathbb{P}(X \leq r) = \mathbb{P}(U \leq F_X(r)) = \mathbb{P}(F_X^{-1}(U) \leq r).$$

$$X = F_X^{-1}(U), \quad F_X^{-1}(r) = \sqrt{r}.$$

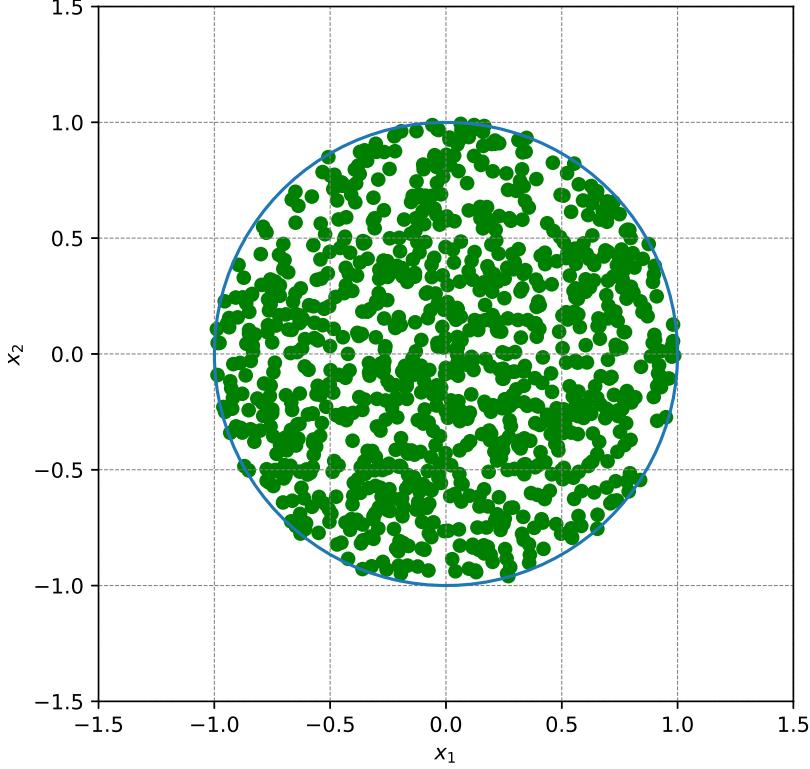


Рис. 26: Равномерное распределение в единичной окружности

Далее оценим точность метода случайного поиска. Сделаем аналитические выводы, которые упрощают задачу минимизации: По условию  $f$  - непрерывная функция и  $f(0, 0) = 0$ . Минимум не может достигаться в точке  $(0, 0)$ . Функция является чётной по  $x_2$ . Из этого следует, что функция имеет не меньше двух точек минимума. Для оценки точности необходимо использовать формулу конечных приращений Лагранжа, для этого рассмотрим частные производные:

$$\frac{\partial f}{\partial x_1} = 3x_1^2 \sin \frac{1}{x_1} - x_1 \cos \frac{1}{x_1} + 10x_2^4 \cos \frac{1}{x_1},$$

$$\frac{\partial f}{\partial x_2} = 40x_1x_2^3 \cos \frac{1}{x_2} - 10x_1x_2^2 \sin \frac{1}{x_2}.$$

Обозначим  $x^* = (x_1^*, x_2^*)$  как точку минимума, полученную теоретически. Пусть

$$\|x^* - x\| = \delta.$$

Тогда справедлива следующая оценка:

$$|f(x_1^*, x_2^*) - f(x_1, x_2)| \leq \sup_A \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2} \cdot \|x^* - x\| \leq \sqrt{14^2 + 50^2} \cdot \delta.$$

Введём дополнительные параметры, необходимые для решения задачи:  $p$  — вероятность попадания хотя бы одной точки в  $\delta$  — окрестность  $x^*$ ,  $n$  — число точек.

$$p = 1 - \mathbb{P}(\forall i : \|x^* - x_i\| > \delta) = 1 - (\mathbb{P}(\|x^* - x_i\| > \delta))^n = 1 - (1 - \mathbb{P}(\|x^* - x_i\| \leq \delta))^n.$$

Через отношение площадей можем посчитать следующую величину:

$$\mathbb{P}(\|x^* - x_i\| \leq \delta) \geq \frac{\pi \delta^2}{\pi \cdot 1^2} = \delta^2,$$

$$p \geq 1 - (1 - \delta^2)^n.$$

Итак,

$$\varepsilon \leq \sqrt{14^2 + 50^2} \delta, \Rightarrow \varepsilon \leq \sqrt{(14^2 + 50^2) \left(1 - \sqrt[n]{1-p}\right)}.$$

Метод случайного поиска при $p = 0.95$			
Число точек	Результат	Точка минимума	Погрешность
$10^3$	-1.2435	$[-0.356, 0.928]$	2.839
$10^4$	-1.2643	$[-0.362, 0.929]$	0.898
$10^5$	-1.287	$[-0.352, -0.935]$	0.284
$10^6$	-1.288	$[-0.359, 0.933]$	0.089

## 7.2 Метод имитации отжига

**Условие.** Методом имитации отжига найти минимальное значение функции Розенброка  $g$  в пространстве  $\mathbb{R}^2$ , где

$$g(x) = (x_1 - 1)^2 + 100(x_2 - x_1^2)^2.$$

За основу метода взят процесс кристаллизации вещества. Цель отжига - привести систему в состояние с наименьшей энергией, поэтому, используя данный метод, можем "хорошо" находить точки глобального минимума. Пусть  $S$  - множество всех состояний системы,  $s_i$  - состояние системы на  $i$ -м шаге,  $t_i$  - температура на  $i$ -м шаге. Для реализации метода необходимо определить следующие функции:

1. функцию, порождающую новое состояние - нормальная случайная величина со средним  $(x_i, y_i)$  и дисперсией  $\sigma^2 \cdot T_i$ ;
2. функцию вероятности перехода  $P = e^{-\Delta F_i / T_i}$ ;
3. функцию изменения температуры. Она обязательно должна уменьшать значение на каждой итерации  $T_{i+1} = kT_i$ ,  $0 < k < 1$ ;

Решение задачи определяется значениями параметров  $\sigma, k, T_0$ . Также знаем, что теоретический минимум функции Розенброка находится в точке  $(1, 1)$ , исходя из этого, оценим оптимальные значения параметров.

1. Изменение температуры:

По сути параметр  $k$  задаёт число итераций, так как условием выхода является  $t_n \leq \varepsilon$ .

$$t_n = k^n \varepsilon$$

Решив это неравенство, получим оценку для  $n$ :

$$n = \lfloor \frac{\ln \varepsilon - \ln T_0}{\ln k} \rfloor + 1.$$

Будем выбирать  $k$  близким к 1, чтобы получить большее число итераций.

2. Переход в новое состояние:

Параметр  $\sigma$  можем выбирать достаточно большим, так как за счёт домножения на  $T_k$ , со временем значение дисперсии окажется в правильном промежутке по правилу трёх сигм (так как нам известен теоретический минимум).

3. Значение  $t_0$  достаточно выбирать большим  $\varepsilon$ . В качестве начального приближения выберем точку  $(0, 0)$ .

Оценим точность графически. Проведём  $m = 50$  тестов и сравним результат в теоретическим.

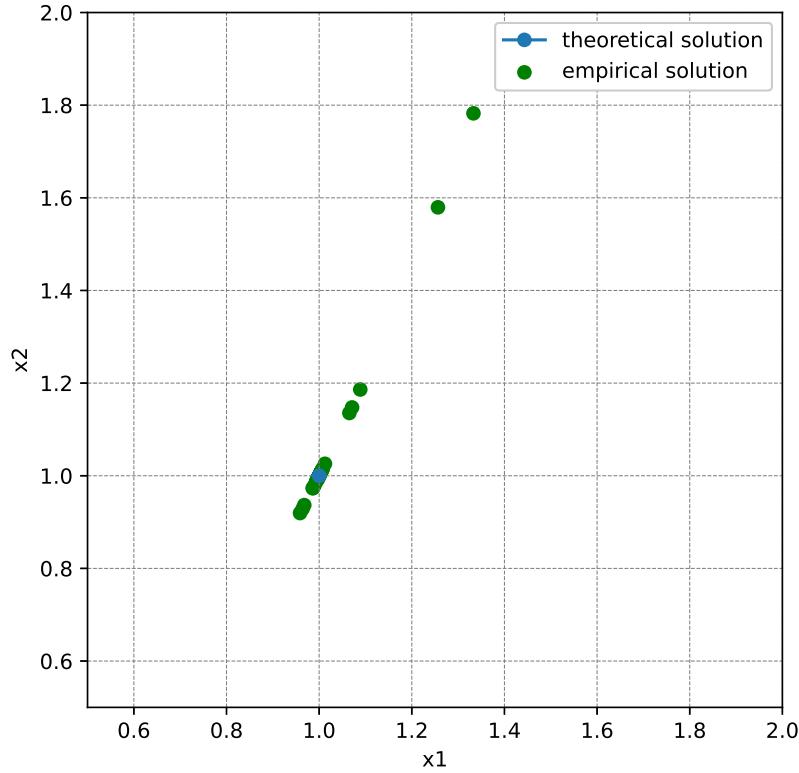


Рис. 27: Решение задачи минимизации методом имитации отжига

### 7.3 Сравнение со стандартными методами оптимизации

**Условие.** Оценить точность и сравнить результаты со стандартными методами оптимизации.

Сначала рассмотрим функцию из задачи 7.1. Вычислим минимум функции  $f(x)$  симплекс-методом Нелдера-Мида. Это алгоритм, в котором вычисляются только значения функции. Симплекс метод является одним из самых простых способов свести достаточно гладкую функцию к минимуму, однако вычисления могут быть слишком долгими.

Метод Нелдера-Мида		
Начальное приближение	Точка минимума	Погрешность
$[-1, 1]$	$[-1.041, 0.764]$	$1e - 8$
$[-0.5, 0.5]$	$[-0.342, 0.642]$	$1e - 8$

Итак, получили точку, близкую к той, которая была получена в 7.1. Правда значение минимума получилось не совсем верное, так как аналитически определили, что значение  $< 0$ .

Теперь рассмотрим функцию из задания 7.2. Функция Розенброка и её матрицы первой и второй производной уже определены в пакете `scipy.optimize`, поэтому будем использовать готовую реализацию.

Минимум функции Розенброка с точностью $1e - 8$		
Метод	Начальное приближение	Точка минимума
Нелдера-Мида	$[0, 0]$	$[1.1]$
Пауэла	$[0, 0]$	$[1.1]$
Сопряжённых градиентов	$[0, 0]$	$[1.000e + 001.000e + 00]$

В случае функции Розенброка все стандартные методы оптимизации показали точный результат.

## 8 Задание 8

### 8.1 Метод Монте-Карло для уравнения Лапласа

**Условие.** Применить метод Монте-Карло к решению первой краевой задачи для двумерного уравнения Лапласа в единичном круге

$$\begin{cases} \Delta u = 0, & (x, y) \in D, \\ u|_{\delta D} = f(x, y), \\ u \in C^2(D), & f \in C(\delta D), \\ D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \end{cases}$$

Использование итерационных методов для решения задачи Дирихле затрачивает очень много времени при вычислениях, поэтому пробуем свести данную задачу к использованию вероятностных процессов. При решении конечноразностной задачи необходимо выбрать шаг  $h$  и построить сетку. Далее рассматриваем узлы, попавшие в область  $D$ , причем можем разделить полученные узлы на две группы. Узлы, для которых четыре соседних узла лежат в области или не её границе, называются внутренними, а узлы, для которых число соседних внутренних узлов меньше 4, называются граничными. В граничных узлах функция  $u$  принимает значения  $u(Q) = f(Q)$ . Во внутренних узлах, в соответствии с разностной схемой, значения функции вычисляются следующим образом:

$$u(P) = \frac{1}{4}(u(P_1) + u(P_2) + u(P_3) + u(P_4)).$$

Здесь  $P_1, P_2, P_3, P_4$  - соседние узлы узла  $P$ . Далее рассмотрим задачу блуждания частицы по некоторой сетке. Предположим, что частица начинает движение из узла  $P$ , и из него с равной вероятностью может попасть в любой из соседних узлов. Аналогично происходит при попадании в любой другой узел по мере движения частицы. Считаем, что выйдя на границу области, частица остается в этом граничном узле. Возникает задача нахождения вероятности  $u(P, Q)$  того, что частица, начав движение в узле  $P$  попадет в граничный узел  $Q$ . Считаем, что в конце движения частица обязательно попадает на границу области. Заметим, что:

$$P \rightarrow Q \Leftrightarrow P \rightarrow P_i \rightarrow Q,$$

где  $P_i \in \{P_1, P_2, P_3, P_4\}$ . Таким образом, пришли к формуле

$$u(P) = \frac{1}{4}(u(P_1) + u(P_2) + u(P_3) + u(P_4)).$$

Также заметим, что  $u$  удовлетворяет следующим краевым условиям:

$$u(Q, Q) = 1, \quad u(Q_1, Q_2) = 0.$$

Предположим дополнительно, что при попадании на границу частица переносит температуру, равную  $f(Q)$ . Значение температуры, перенесённой из узла  $P$  является случайной величиной  $\xi(P)$ . Вероятность переноса температуры  $f(Q_i)$  равна  $u(P, Q_i)$ . Тогда верно следующее выражение:

$$\mathbb{E}\xi(P) = \sum_{i=1}^s f(Q_i) u(P, Q_i).$$

Обозначим эту функцию  $\omega(P)$ . Эта функция также удовлетворяет разностному уравнению:

$$\omega(P) = \frac{1}{4}(\omega(P_1) + \omega(P_2) + \omega(P_3) + \omega(P_4))$$

и краевым условиям

$$\omega(Q) = f(Q) u(Q, Q) = f(Q).$$

Таким образом, функция  $\omega(Q)$  является решением задачи Дирихле.

Полное обоснование можно найти в [6]

## 8.2 Применение метода

**Условие.** Для функции  $f(x, y) = x^2 - y^2$  найти аналитическое решение и сравнить его с полученным по методу Монте-Карло.

Для начала необходимо разделить точки окружности на две группы: граничные и внутренние.

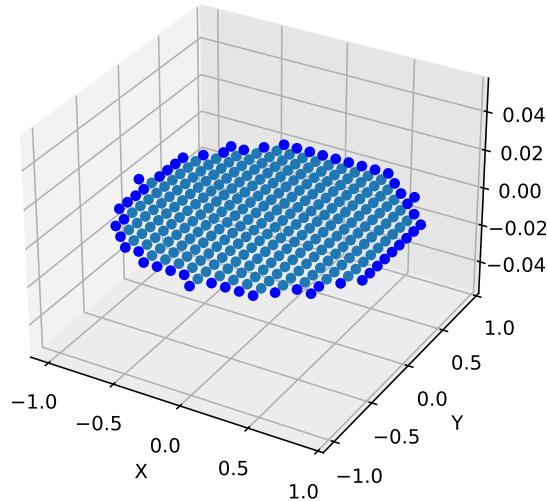
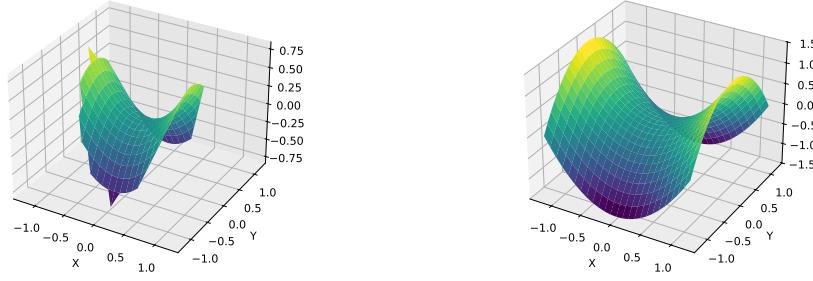


Рис. 28: Разделение точек единичной окружности на группы

Далее по алгоритму, указанному в задании 8.1 построим численное решение.



(a) Численное решение

(b) Аналитическое решение

Рис. 29: Решение первой краевой задачи методом Монте-Карло

Таким образом, получили достаточно хорошее решение.

## 9 Задание 9

**Условие.** Рассмотреть два вида гауссовых процессов: Винеровский процесс, процесс Орнштейна-Уленбека.

1. Найти ковариационную функцию и переходные вероятности
2. Промоделировать независимые траектории процесса с данными переходными вероятностями методом добавления разбиения отрезка
3. Построить график траектории, не соединяя точки ломаной, с целью получения визуально непрерывной линии

### 9.1 Винеровский процесс

**Определение 8.** Винеровским процессом с параметром  $\sigma > 0$  называется случайная функция  $\{W(t), t \geq 0\}$ , удовлетворяющая условиям:

1.  $W(0) = 0$  почти наверное;
2.  $W(t)$  - процесс с независимыми приращениями;
3. Для любых  $t, s \geq 0$  выполнено  $W(t) - W(s) \sim N(0, \sigma^2|t-s|)$ .

Можно заметить, что

$$W(t) = W(t) - W(0) \sim N(0, \sigma^2 t).$$

Отсюда легко выводится:

$$\mathbb{E}W(t) = 0, \quad \text{var}(W(t)) = \sigma^2 t.$$

Рассмотрим ковариационную функцию. Пусть  $t > s$ , тогда  $W(t) - W(s)$  и  $W(s)$  являются независимыми, значит

$$R(t, s) = \mathbb{E}(W(t) - W(s))\mathbb{E}W(s) + \mathbb{E}W^2(s) = \mathbb{E}W^2(s) = \sigma^2 s.$$

Аналогично, при  $t < s$ ,  $R(t, s) = \sigma^2 t$ . Значит

$$R(t, s) = \sigma^2 \min(t, s).$$

Для моделирования используем метод добавления разбиения отрезка:

1. Знаем, что  $W(0) = 0$ , тогда  $W(1) = W(1) - W(0) \sim N(0, \sigma^2)$ ;
2. Зная значения  $W(t_1), W(t_2)$  вычисляем  $W(t_*)$ , для простоты выбираем  $t_* = \frac{t_1 + t_2}{2}$ ;
3. Повторяем шаг 2 пока не достигнем нужного числа точек, причем после  $n$  итераций получим  $2^n + 1$  точку.

Осталось найти способ вычисления  $W(t_*)$ . Знаем, что  $W(t_*)$  нормально распределена, посчитаем параметры. Для этого рассмотрим условную плотность:

$$f_{W_{t_*}}(\omega_* | W_{t_1} = \omega_1, W_{t_2} = \omega_2) = \frac{f_{W_{t_1} W_{t_*} W_{t_2}}(\omega_1, \omega_*, \omega_2)}{f_{W_{t_1} W_{t_2}}(\omega_1, \omega_2)},$$

причём

$$f_{W_1, \dots, W_n} = \frac{1}{\sqrt{(2\pi)^n \det(R)}} \exp\left\{-\frac{1}{2}(\omega - \mu) R^{-1} (\omega - \mu)^T\right\}.$$

Используя формулы, приведённые выше, можем записать значения  $R$  для формулы в числите и знаменателе:

$$\begin{aligned} f_{W_{t_1} W_{t_*} W_{t_2}} &\rightarrow R_1 = \sigma^2 \cdot \begin{bmatrix} t_1 & t_1 & t_1 \\ t_1 & t_* & t_* \\ t_1 & t_* & t_2 \end{bmatrix} \\ f_{W_{t_1} W_{t_2}} &\rightarrow R_2 = \sigma^2 \cdot \begin{bmatrix} t_1 & t_1 \\ t_1 & t_2 \end{bmatrix} \end{aligned}$$

Для упрощения далее используем символьные вычисления.

Итак, получили

$$\sigma_* = \frac{\sigma^2 (t_2 - t_1)}{4}.$$

Рассмотрим теперь получившуюся степень экспоненты и приведём к нужному виду:

$$\frac{4\omega_*^2 - 4\omega_*\omega_1 - 4\omega_*\omega_2 + \omega_1^2 + 2\omega_1\omega_2 + \omega_2^2}{2\sigma^2(t_1 - t_2)} = -\frac{1}{2} \frac{4\omega_*^2 - 4\omega_* (\omega_1 + \omega_2) + (\omega_1 + \omega_2)^2}{4\sigma_*^2} = -\frac{1}{2} \frac{\left(\omega_* - \frac{\omega_1 + \omega_2}{2}\right)^2}{\sigma_*^2}.$$

Таким образом, подобрали параметры и в алгоритме используем

$$W(t_*) \sim N\left(\frac{\omega_1 + \omega_2}{2}, \frac{\sigma^2 (t_2 - t_1)}{4}\right).$$

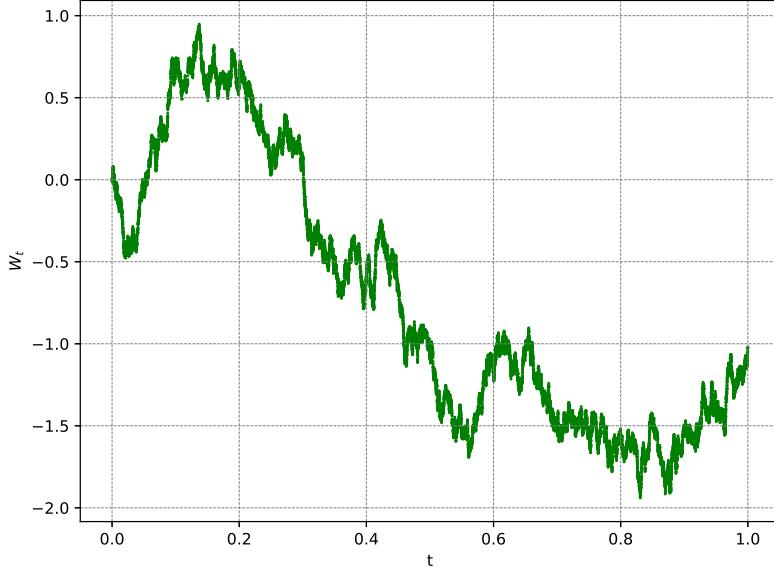


Рис. 30: Траектория Винеровского процесса с  $\sigma = 1.5$

## 9.2 Процесс Орнштейна-Уленбека

Этот процесс вводился как единственный нетривиальный гауссовский процесс, являющийся стационарным и обладающий марковским свойством. Из стационарности следует, что

$$\mathbb{E}X_t = a = \text{const}, \quad \text{var}(X_t) = \sigma^2 = \text{const}, \quad K_X(s, t) = \rho(|s - t|).$$

Из марковости получаем

$$K_X(s, t) = K_X(s, \tau) K_X(\tau, t), \quad \forall 1 \leq s \leq \tau \leq t \leq n.$$

Здесь  $K$  - корреляционная функция. Сделав замену  $s = \tau - s, y = t - \tau$ , получим

$$\rho(|s - t|) = \rho(|s - \tau|) \rho(|\tau - t|) \Leftrightarrow \rho(|-x - y|) = \rho(|-x|) \rho(|-y|) \Leftrightarrow \rho(x + y) = \rho(x) \rho(y).$$

Полученное уравнение называется уравнением Коши и имеет решения:

$$K_X(t) = 0, \quad K_X(t) = e^{-\lambda t}, \quad \lambda > 0.$$

Случай  $K_X(t, s) = 0$  не так интересен для рассмотрения, так как ввиду гауссовой природы получим, что случайные величины  $X_t$  независимы в совокупности.

В другом случае получим

$$R(s, t) = \sigma^2 e^{-\lambda |t-s|}, \quad \lambda > 0.$$

Хотим решить задачу с помощью метода добавления разбиения отрезка. В отличие от винеровского процесса точное значение  $X(0) = X_0$  неизвестно, его нужно выбирать случайным образом для стационарности процесса. Найдём плотность  $f_{X_{t_1}}(x_1 | X_{t_2} = x_2) = \frac{f_{X_{t_1} X_{t_2}}(x_1, x_2)}{f_{X_{t_2}}(x_2)}$ .

$$f_{X_{t_1} X_{t_2}}(x_1, x_2) = \frac{1}{\sqrt{4\pi^2 \det(R_0)}} \exp\left\{-\frac{1}{2}(x - a) R_0^{-1} (x - a)^T\right\},$$

$$f_{X_{t_2}}(x_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_2 - a)^2}{2\sigma^2}\right\}.$$

Воспользуемся символьными вычислениями.

В рассматриваемом случае  $d = 1$ . Тогда

$$\sigma_* = \sigma^2 (1 - e^{-2\lambda}).$$

Преобразовав выражение в степени для экспоненты получим:

$$\mu_* = x_2 e^{-\lambda} + a (1 - e^{-\lambda}).$$

Итак,

$$f_{X_{t_1}}(x_1 | X_{t_2} = x_2) \sim N(x_2 e^{-\lambda} + a (1 - e^{-\lambda}), \sigma^2 (1 - e^{-2\lambda})).$$

Аналогично случаю с винеровским процессом ищем значения по формулам.

$$\begin{aligned} f_{X_{t_*}}(x_* | X_{t_1} = x_1, X_{t_2} = x_2) &= \frac{f_{X_{t_1} X_{t_*} X_{t_2}}(x_1, x_*, x_3)}{f_{X_{t_1} X_{t_2}}(x_1, x_3)}, \\ f_{X_1, \dots, X_n} &= \frac{1}{\sqrt{(2\pi)^n \det(R)}} \exp\left\{-\frac{1}{2} (x - \mu) R^{-1} (x - \mu)^T\right\}. \\ f_{X_{t_1} X_{t_*} X_{t_2}} &\rightarrow R_1 = \sigma^2 \cdot \begin{bmatrix} 1 & e^{-\lambda(t_* - t_1)} & e^{-\lambda(t_2 - t_1)} \\ e^{-\lambda(t_* - t_1)} & 1 & e^{-\lambda(t_2 - t_*)} \\ e^{-\lambda(t_2 - t_1)} & e^{-\lambda(t_2 - t_*)} & 1 \end{bmatrix} \\ f_{X_{t_1} X_{t_2}} &\rightarrow R_2 = \sigma^2 \cdot \begin{bmatrix} 1 & e^{-\lambda(t_2 - t_1)} \\ e^{-\lambda(t_2 - t_1)} & 1 \end{bmatrix} \end{aligned}$$

Введём обозначения

$$d = t - t_1$$

Для упрощения вычислений положим  $a = 0$ . Это возможно, если рассматривать слоучайно величину со сдвигом  $X_t - a$ , а потом переобозначить эту величину за  $X_t$ . Таким образом, упростив выражения, получим

$$\sigma_* = \frac{\sigma^2 (e^{2d\lambda} - 1)}{e^{2d\lambda} + 1},$$

$$\mu = \frac{e^{\lambda d} (x_1 + x_2)}{e^{2\lambda d} + 1},$$

Тогда

$$X_{t_*} \sim N\left(\frac{e^{\lambda d} (x_1 + x_2)}{e^{2\lambda d} + 1}, \frac{\sigma^2 (e^{2d\lambda} - 1)}{e^{2d\lambda} + 1}\right).$$

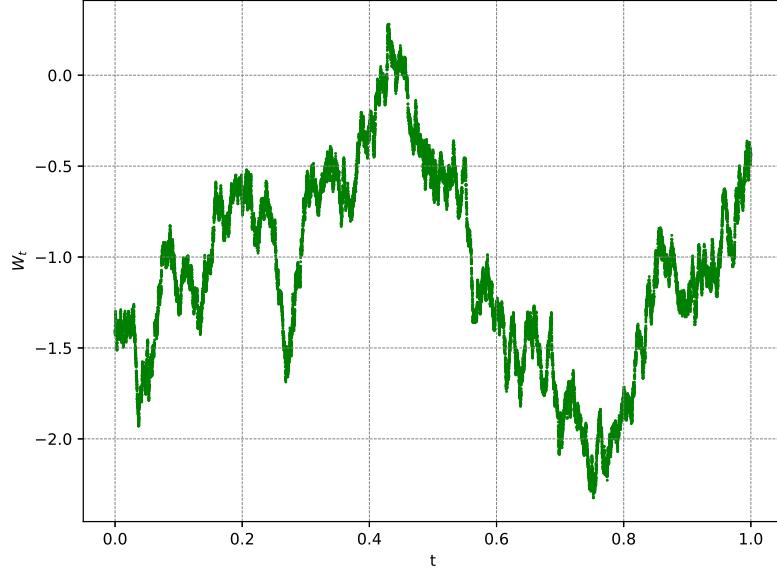


Рис. 31: Траектория процесса Орнштейна-Уленбека с  $\sigma = 1, \lambda = 1$

## 10 Задание 10

**Условие.** Произвести фильтрацию одномерного процесса Орнштейна-Уленбека:

1. Используя генератор белого шума, добавить к реализации процесса Орнштейна-Уленбека случайную ошибку с заранее известной дисперсией;
2. При помощи одномерного фильтра Калмана оценить траекторию процесса по зашумлённому сигналу, считая известными параметры шума и процесса;
3. Рассмотреть следующие виды шума: Гауссов, Коши.

Рассматриваем процесс Орнштейна-Уленбека на сетке, тогда его "координаты" изменяются по следующему закону:

$$x_{k+1} = ax_k + v_k.$$

Считается, что на устройстве принимаем зашумлённый сигнал. Его динамика описывается следующим образом:

$$y_k = x_k + \xi_k.$$

По условию заранее знаем дисперсию ошибки:  $\xi_k \sim N(0, r^2)$ . Считаем, что  $v_k \sim N(0, q^2)$ . Также из предыдущего задания знаем, что  $x_0 \sim N(0, \sigma^2)$ .

Для дальнейшего решения задачи необходимо найти значения параметров  $a, q$ . Зная, что  $x_k$  и  $v_k$  независимы, а также свойства процесса Орнштейна-Уленбека:

$$R(t_k, t_k) = \sigma^2,$$

$$R(t_k, t_{k+1}) = a \cdot \text{var}(x_{t_k}) = \sigma^2 e^{-\lambda(t_{k+1}-t_k)} \Rightarrow a = e^{-\lambda(t_{k+1}-t_k)},$$

$$R(t_{k+1}, t_{k+1}) = a^2 \sigma^2 + q^2 = \sigma^2 \Rightarrow q^2 = \sigma^2 \left(1 - e^{-2\lambda(t_{k+1}-t_k)}\right).$$

Далее необходимо выбрать коэффициент Калмана  $K_{k+1}$  таким, чтобы получившееся оптимальное значение  $x_{k+1}^{opt}$  было бы наиболее близко к истинному значению  $x_{k+1}$ . Чтобы найти точное значение коэффициента Калмана  $K_{k+1}$ , нужно минимизировать ошибку:

$$x_{k+1}^{opt} = K_{k+1}y_{k+1} + (1 - K_{k+1})ax_k^{opt},$$

$$e_{k+1} = x_{k+1} - x_{k+1}^{opt}.$$

Тогда

$$\begin{aligned} e_{k+1} &= x_{k+1} - K_{k+1}y_{k+1} - (1 - K_{k+1})ax_k^{opt} = ax_k + v_k - K(ax_k + v_k + \xi_{k+1}) - (1 - K_{k+1})ax_k^{opt} = \\ &= (1 - K_{k+1})(ae_k + v_k) - K_{k+1}\xi_{k+1}. \end{aligned}$$

Минимизировать ошибку означает  $\mathbb{E}(e_{k+1}^2) \rightarrow \min$ :

$$\mathbb{E}(e_{k+1}^2) = (1 - K_{k+1})^2(a^2\mathbb{E}e_k^2 + q^2) + K_{k+1}^2r^2.$$

Приравнивая производную выражения к 0, получим

$$K_{k+1} = \frac{a^2(\mathbb{E}e_k^2 + q^2)}{a^2(\mathbb{E}e_k^2 + q^2) + r^2}.$$

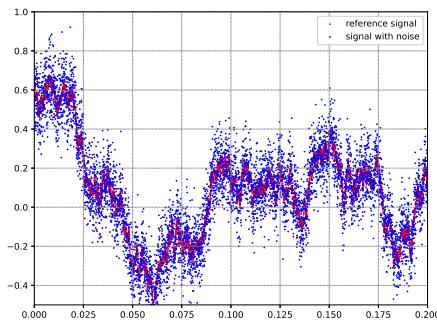
Подставляем это значение в минимизируемую функцию

$$\mathbb{E}(e_{k+1}^2) = \frac{r^2(a^2(\mathbb{E}e_k^2 + q^2))}{a^2(\mathbb{E}e_k^2 + q^2) + r^2}.$$

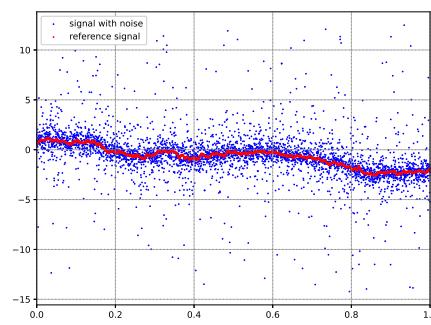
Запишем алгоритм работы программы:

1.  $\mathbb{E}(e_0^2) = r^2$ ,  $x_0^{opt} = y_0$ ;
2. На каждой итерации вычисляем  $\mathbb{E}(e_{k+1}^2)$  по формуле, представленной выше. Далее вычисляем  $K_{k+1} = \frac{\mathbb{E}(e_{k+1}^2)}{r^2}$ . Зная коэффициент Калмана вычисляем  $x_{k+1}^{opt}$  по формуле выше.

Параметры программы:  $\sigma, r, \lambda$ , число точек.

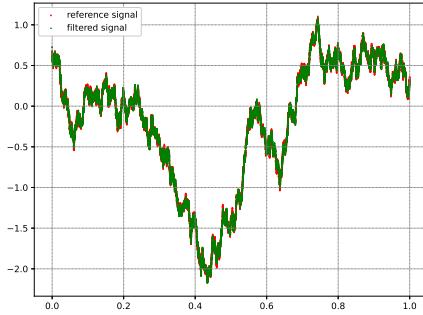


(a) Гауссовский шум

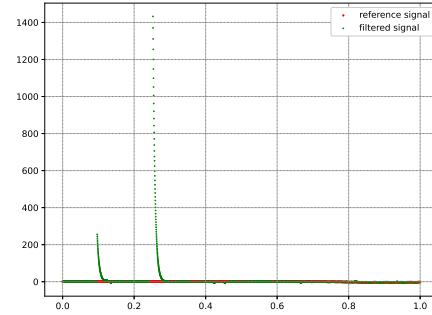


(b) Шум Коши

Рис. 32: Зашумлённые данные



(a) Гауссовский шум



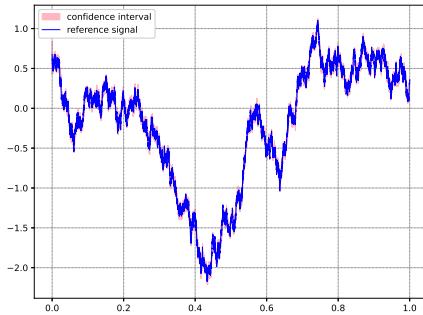
(b) Шум Коши

Рис. 33: Результат фильтрации

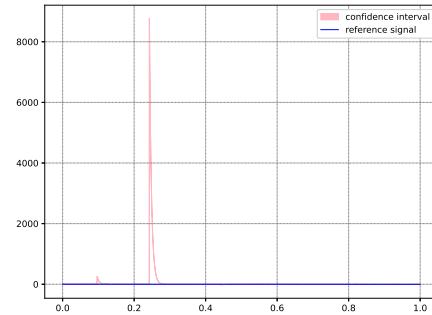
Ошибка фильтрации имеет нормальное распределение и доверительный интервал имеет вид:

$$[x^{opt} - \varepsilon, x^{opt} + \varepsilon], \quad \varepsilon = -\sqrt{\mathbb{E}e} \cdot \Phi\left(\frac{\alpha}{2}\right),$$

где  $\alpha$  - уровень значимости.



(a) Гауссовский шум



(b) Шум Коши

Рис. 34: Доверительный интервал

## 11 Задание 11

**Условие.** Построить двумерное пуассоновское поле, отвечающее сложному пуассоновскому процессу.

### 11.1 Система массового обслуживания

**Условие.** Система массового обслуживания. Первая координата поля - время поступления заявки в СМО (распределённое равномерно), а вторая - время обслуживания заявки (распределение  $\chi^2$  с 10 степенями свободы)

Рассматриваем первую координату поля на отрезке  $[0, T]$ . В теории пуассоновского поля  $k$  - мера Лебега,  $\lambda$  - интенсивность поля (в нашем случае интенсивность поступления заявок =

среднее число новых заявок в единицу времени). Тогда случайная считающая мера, равная числу заявок определяется следующим образом:

$$n \sim Pois(\lambda T).$$

По условию время поступления заявки распределено равномерно. Введём обозначение:

$$t_i \sim U[0, T].$$

После моделирования моментов поступления заявок необходимо упорядочить их по возрастанию. Пусть  $\tau_i$  - время обслуживания заявки, по условию имеет распределение  $\chi^2$  с 10 степенями свободы. Вспомним важное свойство этого распределения:

$$\mathbb{E}\tau_i = 10.$$

Проведём логическое рассуждение. В момент времени  $t_1$  в обработку поступает первая заявка, следовательно, в момент времени  $t_1 + \tau_1$  первая заявка будет обработана. Причём в момент обработки первой заявки могут поступать другие заявки, которые образуют очередь и обрабатываются постепенно. Тогда для последующих заявок выполняется правило: если  $(i-1)$ -я заявка была обработана до момента поступления  $i$ -й заявки  $t_i$ , то она будет обработана в момент времени  $t_i + \tau_i$ , иначе она будет обработана через  $\tau_i$  единиц после завершения обработки  $(i-1)$ -й заявки. Обозначим момент окончания обработки  $(i-1)$ -й заявки как  $t_i^{end}$ . Тогда можем посчитать количество заявок, находящихся в очереди в момент  $t$ :

$$Q(t) = \sum_{i=1}^n \mathbb{1}_{t_i \leq t < t_{i-1}^{end}}.$$

Ввиду свойств рассматриваемого пуассоновского поля:

$$t_{i+1} - t_i \sim Exp(\lambda).$$

Тогда решение задачи можем разбить на 3 случая. Из распределения времени обслуживания заявки знаем, что среднее время обслуживания одной заявки равно числу степеней свободы, то есть 10. Из свойства указанного выше среднее время между поступлением заявок равно  $\frac{1}{\lambda}$ .

Тогда

1. Если  $\frac{1}{\lambda} > 10$ , то есть заявки поступают медленнее, чем обрабатываются СМО справляется с задачей.

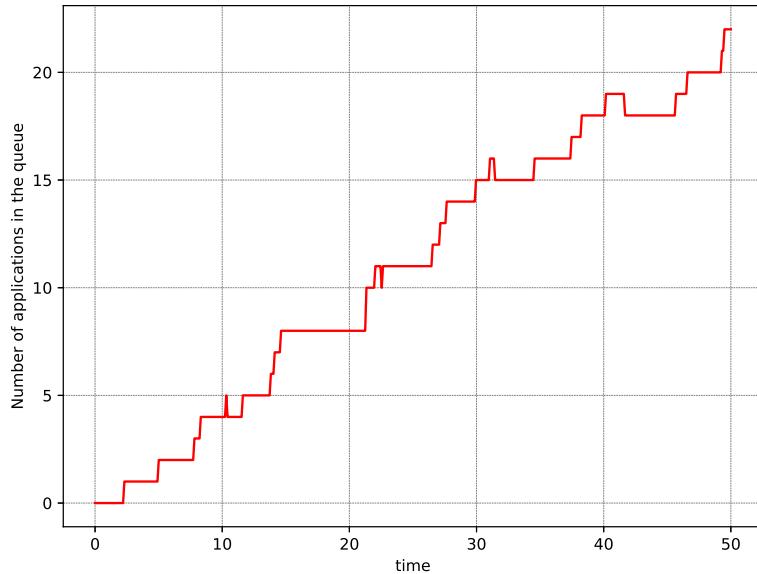


Рис. 35: Система массового обслуживания  $\lambda = 0.5$

2. Если  $\frac{1}{\lambda} < 10$ , то СМО не справляется с обработкой.

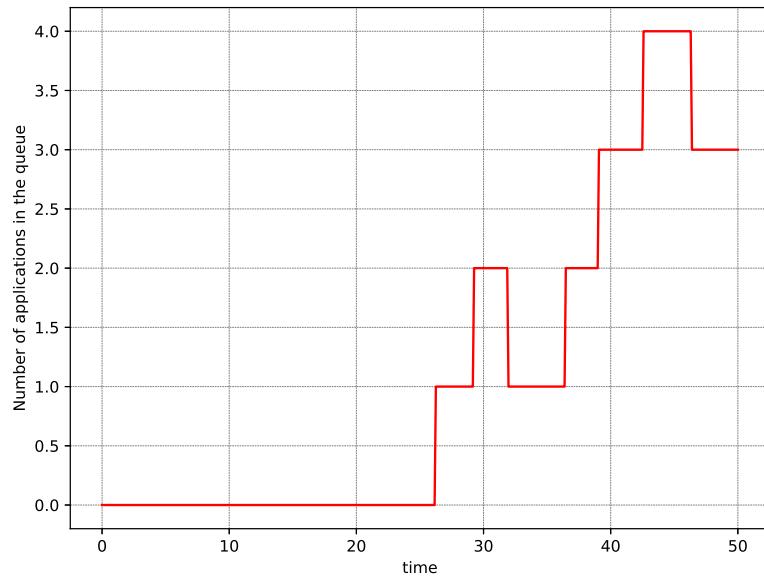


Рис. 36: Система массового обслуживания  $\lambda = 0.07$

3. Если  $\frac{1}{\lambda} = 10$ , СМО находится в равновесии.

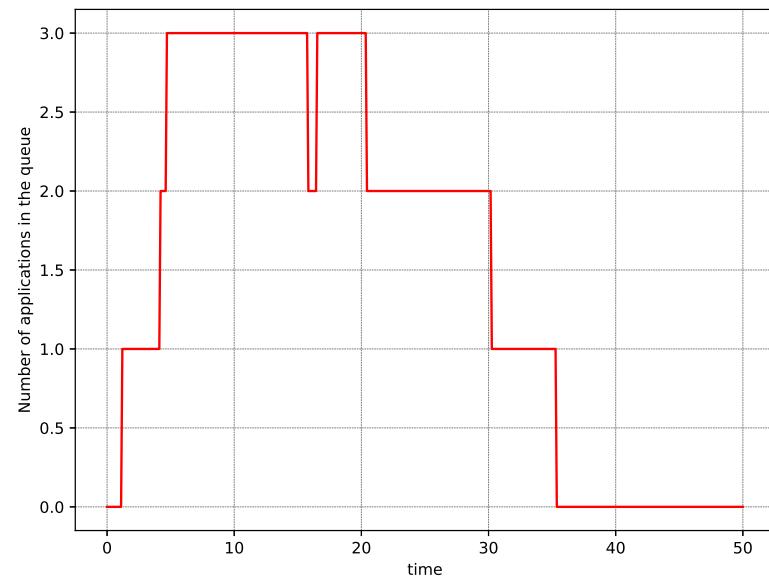


Рис. 37: Система массового обслуживания  $\lambda = 0.1$

## 11.2 Система массового обслуживания с циклической интенсивностью

**Условие.** Система массового обслуживания с циклической интенсивностью  $\lambda(1 + \cos(t))$  и единичными скачками. При помощи метода Льюиса и Шедлеара, свести задачу моделирования пуассоновского процесса к моделированию двумерного пуассоновского поля, где первая координата распределена равномерно, а вторая имеет распределение Бернулли.

Метод Льюиса и Шедлеара позволяет преобразовать неоднородный пуассоновский процесс в однородный. Для этого:

1. Определим верхнюю границу для интенсивности. В нашем случае максимальная интенсивность равна  $2\lambda$ .
2. Генерируем пуассоновский процесс как в предыдущем пункте с постоянной интенсивностью  $2, \lambda$ .
3. Для каждого события, сгенерированного в однородном процессе, определяем, произошло ли оно на основе исходной функции интенсивности  $\lambda(t)$ . То есть удаляем  $t_i$  с вероятностью  $1 - \frac{\lambda(t_i)}{2\lambda}$ .

Таким образом, как раз получим поле, где первая координата распределена равномерно, а вторая имеет распределение Бернулли.

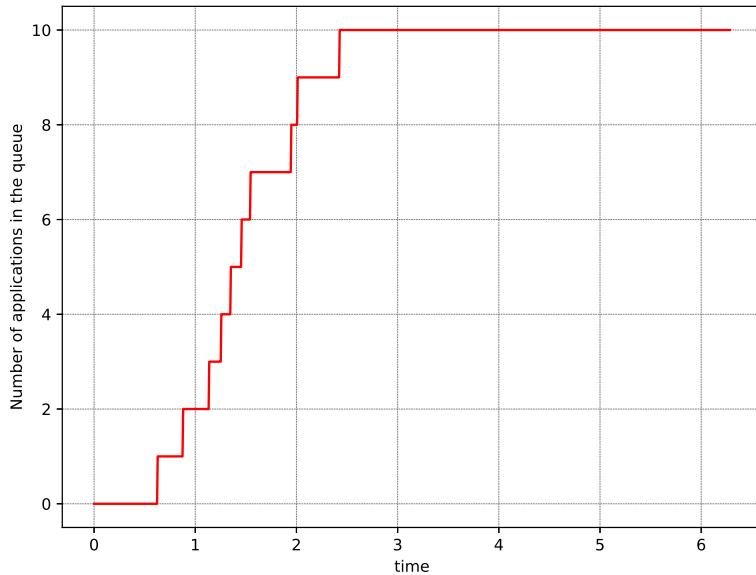


Рис. 38: Система массового обслуживания с циклической интенсивностью  $\lambda = 2$

### 11.3 Работа страховой компании

**Условие.** Работа страховой компании: первая координата — момент наступления страхового случая (равномерное распределение), вторая — величина ущерба (распределение Парето). Поступление капитала считать линейным по времени со скоростью  $c > 0$ , начальный капитал  $W > 0$ .

Величина ущерба  $\tau_i$  имеет распределение Парето с параметрами  $(a, b)$ . Знаем, что

$$\mathbb{E}\tau_i = \begin{cases} \frac{ab}{a-1}, & a > 1, \\ +\infty, & a \leq 1. \end{cases}$$

Будем моделировать случайную величину из этого распределения методом обратных функций:

$$X = \frac{b}{(1 - Y)^{1/a}}, \quad Y \sim U[0, 1].$$

Значит

$$X = \frac{b}{Y^{1/a}}, \quad Y \sim U[0, 1].$$

Капитал страховой компании изменяется по правилу:

$$W(t) = W + ct - \sum_{i=1}^{Q(t)} \tau_i,$$

где  $Q(t)$  — число страховых случаев.

После того, как в какой то момент времени  $W(t) = 0$ , оставшийся вектор заполняется нулями — это случай банкротства.

Алгоритм моделирования такой же как в случаях выше.

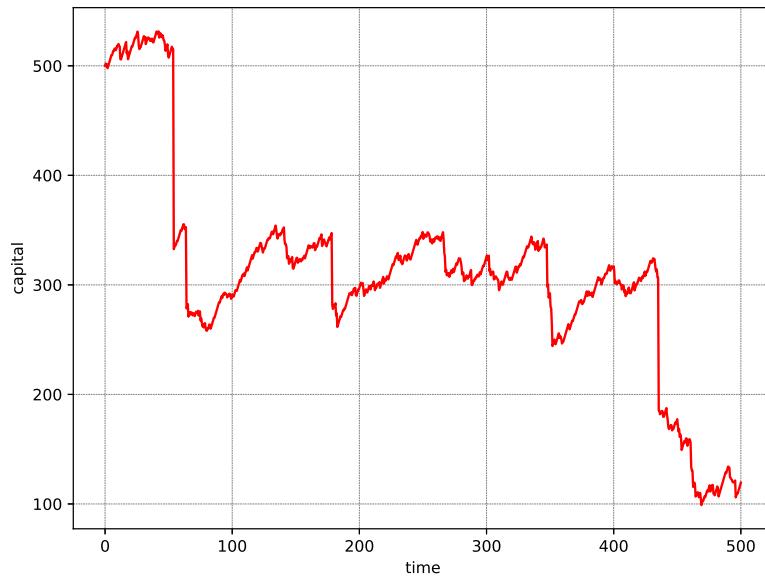


Рис. 39: Работа страховой компании  $\lambda = 1, W = 500, T = 500, a = 2, b = 2, c = 4$

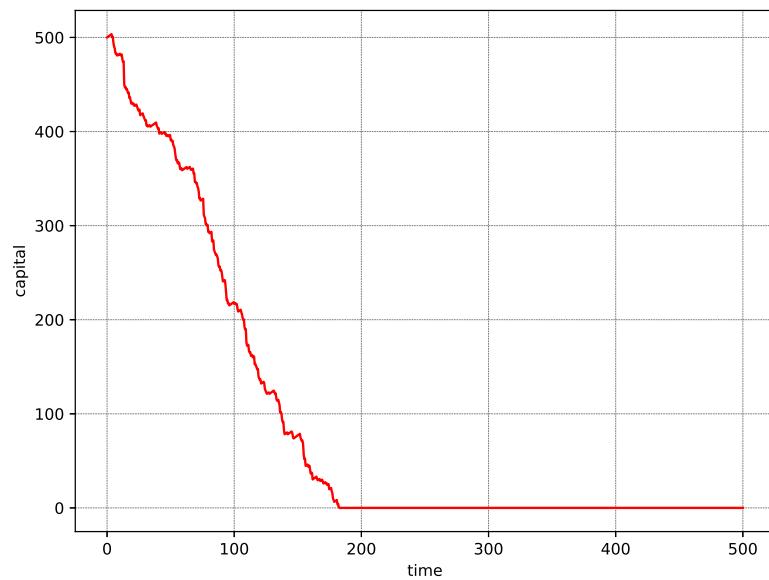


Рис. 40: Случай банкротства  $\lambda = 1, W = 500, T = 500, a = 2, b = 2, c = 1$

## Список литературы

- [1] Ширяев А. Н. *Вероятность — 1.* МЦНМО, 2004.
- [2] Кропчаева Н. Ю., Тихомирова А. С. *Моделирование случайных величин: метод, указания.* НовГУ им. Ярослава Мудрого, 2004.
- [3] Булинский А. В., Ширяев А. Н. *Теория случайных процессов.* ФИЗМАТЛИТ, 2005.
- [4] Devroye L. *Non-uniform random variate generation.* Springer-Verlag, New York 1986.
- [5] Ульянов В. В. *Лекции по теории вероятностей и математической статистике,* 2022.
- [6] Бусленко Н. П., Соболь И. М., Срагович В. Г., Шрейдер Ю. А., *Метод статистических испытаний.* 1961.
- [7] Смирнов С. Н. *Лекции по стохастическому анализу и моделированию,* 2024.