



Московский государственный университет имени М. В. Ломоносова  
Факультет вычислительной математики и кибернетики  
Кафедра системного анализа

Практикум

# «Стохастический анализ и моделирование»

*Студент 415 группы*  
К. Ю. Егоров

*Руководитель практикума*  
к.ф.-м.н., доцент С. Н. Смирнов

Москва, 2019

## Содержание

<b>1</b>	<b>Задание №1</b>	<b>4</b>
1.1	Генераор схемы Бернулли. Датчик биномиального распределения . . . . .	4
1.2	Датчик геометрического распределения. Его свойство отсутствия памяти .	5
1.3	Игра в орлянку . . . . .	5
<b>2</b>	<b>Задание №2</b>	<b>9</b>
2.1	Построение датчика «канторовой» случайной величины . . . . .	9
2.2	Проверка корректности работы датчика . . . . .	11
2.3	Симметричность «канторовой» случайной величины относительно $1/2$ . . . .	11
2.4	Самоподобие «канторовой случайной величины относительно деления на 3	12
2.5	Математическое ожидание и дисперсия «канторовой» случайной величины	13
<b>3</b>	<b>Задание №3</b>	<b>18</b>
3.1	Построение датчика экспоненциальной случайной величины . . . . .	18
3.2	Отсутствие памяти у экспоненциального распределения . . . . .	18
3.3	Распределение минимума экспоненциальных случайных величин . . . . .	19
3.4	Построение датчика распределения Пуассона . . . . .	19
3.5	Построение датчика распределения Пуассона как предел биномиального распределения . . . . .	20
3.6	Проверка корректности работы датчика . . . . .	21
3.7	Построение датчика нормального распределения методом моделирования парами с переходом в полярные координаты . . . . .	21
3.8	Проверка равенства математического ожидания и дисперсии . . . . .	22
<b>4</b>	<b>Задание №4</b>	<b>27</b>
4.1	Построение датчика распределения Коши . . . . .	27
4.2	Построение датчика стандартного нормального распределения методом фон Неймана . . . . .	27
4.3	Дополнительно о методе фон Неймана . . . . .	28
<b>5</b>	<b>Задание №5</b>	<b>32</b>
5.1	Закон больших чисел. Центральная предельная теорема. . . . .	32
5.2	Построение доверительных интервалов для среднего и дисперсии нормаль- ной случайной величины. . . . .	32
5.3	Поведение частичных сумм распределения Коши . . . . .	34
<b>6</b>	<b>Задание №6</b>	<b>38</b>
6.1	Метод Монте–Карло . . . . .	38
6.2	Метод квадратур . . . . .	39
<b>7</b>	<b>Задание №7</b>	<b>41</b>
7.1	Метод случайного поиска . . . . .	41
7.2	Метод имитации отжига . . . . .	43
<b>8</b>	<b>Задание №8</b>	<b>46</b>
8.1	Алгоритм решения задачи . . . . .	46

<b>9</b>	<b>Задание №9</b>	<b>50</b>
9.1	Винеровский процесс . . . . .	50
9.1.1	Переходные вероятности . . . . .	51
9.1.2	Алгоритм . . . . .	51
9.2	Процесс Орнштейна–Уленбека . . . . .	53
9.2.1	Алгоритм . . . . .	54
<b>10</b>	<b>Задание №10</b>	<b>56</b>
10.1	О белом шуме . . . . .	56
10.2	Фильтр Калмана для гауссовского шума . . . . .	56
<b>11</b>	<b>Задание №11</b>	<b>59</b>
11.1	Система массового обслуживания . . . . .	59
11.2	Система массового обслуживания с циклической интенсивностью и единичными скачками . . . . .	61
11.3	Работа страховой компании . . . . .	62

## 1 Задание №1

1. Реализовать генератор схемы Бернулли с заданной вероятностью успеха  $p$ . На основе генератора схемы Бернулли построить датчик для биномиального распределения.
2. Реализовать генератор геометрического распределения. Проверить для данного распределения свойство отсутствия памяти.
3. Рассмотреть игру в орлянку — бесконечную последовательность испытаний с бросанием правильной монеты. Выигрыш  $S_n$  определяется как сумма по всем  $n$  испытаниям значений 1 и  $-1$  в зависимости от выпавшей стороны. Проиллюстрировать (в виде ломаной) поведение нормированной суммы  $Y(i) = \frac{S_i}{\sqrt{n}}$  как функции от номера испытания  $i = 1, \dots, n$  для одной отдельно взятой траектории. Дать теоретическую оценку для  $Y(n)$  при  $n \rightarrow \infty$ .

### 1.1 Генератор схемы Бернулли. Датчик биномиального распределения

**Определение 1.1.** *Схемой Бернулли* называется последовательность испытаний, в каждом из которых возможны два исхода — «успех» и «неудача», при этом «успех» в каждом испытании происходит с одной и той же вероятностью  $p \in (0, 1)$ , а «неудача» — с вероятностью  $q \equiv 1 - p$ . На испытания в схеме Бернулли налагаются следующие требования: отсутствие взаимного влияния, воспроизводимость, а также сходные — но не идентичные — условия проведения.

**Определение 1.2.** Будем говорить, что случайная величина  $X$  имеет *распределение Бернулли*, если она принимает всего два значения: 1 и 0 с вероятностями  $p$  и  $q \equiv 1 - p$  соответственно. Таким образом,

$$\mathbb{P}(X = 1) = p \quad \text{и} \quad \mathbb{P}(X = 0) = q,$$

то есть событие  $\{X = 1\}$  соответствует «успеху», а  $\{X = 0\}$  — «неудаче». Будем обозначать такую случайную величину

$$X \sim \text{Bern}(p).$$

Реализуем генератор схемы Бернулли с заданной вероятностью успеха  $p$  следующим образом: пусть нам дана случайная величина  $\xi$ , равномерно распределённая на отрезке  $[0, 1]$ . Тогда случайная величина  $X \sim \text{Bern}(p)$  задаётся следующим образом:

$$X = \mathbb{I}(\xi < p) = \begin{cases} 1, & \text{при } 0 \leq \xi < p, \\ 0, & \text{при } p \leq \xi \leq 1. \end{cases}$$

**Определение 1.3.** Будем говорить, что случайная величина  $X$  имеет *биномиальное распределение* с параметрами  $n$  и  $p$ , если

$$\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}, \quad \text{где } k \in \mathbb{N}_0.$$

В таком случае  $X$  интерпретируют как число «успехов» в серии из  $n$  испытаний схемы Бернулли с вероятностью успеха  $p$ . Будем обозначать такую случайную величину

$$X \sim \text{Bi}(n, p).$$

Пусть теперь  $X \sim \text{Bi}(n, p)$ , а  $Y_i \sim \text{Bern}(p)$ ,  $i = \overline{1, n}$ . Тогда, как видно из интерпретации биномиального распределения, датчик биномиальной случайной величины будет иметь вид:

$$X = \sum_{i=1}^n Y_i.$$

## 1.2 Датчик геометрического распределения. Его свойство отсутствия памяти

**Определение 1.4.** Будем говорить, что случайная величина  $X$  имеет *геометрическое распределение*, если

$$\mathbb{P}(X = k) = (1 - p)^k p = q^k p, \quad \text{где } k \in \mathbb{N}_0.$$

Геометрически распределенная случайная величина интерпретируется как количество «неудач» до первого «успеха» в схеме испытаний Бернулли с вероятностью  $p$ . Будем обозначать такие случайные величины

$$X \sim \text{Geom}(p).$$

Зная интерпретацию, мы легко строим датчик и для геометрического распределения.

**Утверждение 1.1** (Свойство отсутствия памяти). Пусть  $X \sim \text{Geom}(p)$ , тогда для любых  $n, m \in \mathbb{N}_0$  справедливо

$$\mathbb{P}(X \geq m + n \mid X \geq m) = \mathbb{P}(X \geq n),$$

то есть количество прошлых «неудач» не влияет на количество будущих «неудач».

**Доказательство.** Рассмотрим левую часть равенства из условия утверждения:

$$\begin{aligned} \mathbb{P}(X \geq m + n \mid X \geq m) &= \frac{\mathbb{P}(X \geq m + n, X \geq m)}{\mathbb{P}(X \geq m)} = \\ &= \frac{\mathbb{P}(X \geq m + n)}{\mathbb{P}(X \geq m)} = \frac{\sum_{i=m+n}^{\infty} q^i p}{\sum_{i=m}^{\infty} q^i p} = \frac{q^{m+n}}{q^m} = q^n. \end{aligned}$$

Теперь рассмотрим правую часть равенства:

$$\mathbb{P}(X \geq n) = \sum_{i=n}^{\infty} q^i p = p \frac{q^n}{1 - q} = q^n.$$

Таким образом, утверждение доказано. ■

## 1.3 Игра в орлянку

Рассмотрим игру Орлянка, правила которой описаны в формулировке задания и построим траекторию заданного процесса  $Y(i)$ .

В данной нормированной сумме  $Y$  фигурируют независимые одинаково распределенные случайные величины  $X_i$ . Посчитаем их математическое ожидание и дисперсию.

$$\mathbb{E} X_i = 1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = 0,$$

$$\text{Var } X_i = \frac{1}{2} \cdot (1 - 0)^2 + \frac{1}{2} \cdot (-1 - 0)^2 = \frac{1}{4}.$$

Теперь можем воспользоваться следующей теоремой.

**Теорема 1.1** (Центральная предельная теорема). Пусть  $X_1, \dots, X_n, \dots$  есть бесконечная последовательность независимых одинаково распределённых случайных величин, имеющих конечное математическое ожидание  $\mu$  и дисперсию  $\sigma^2$ . Пусть также  $S_n = \sum_{i=1}^n X_i$ . Тогда

$$\frac{S_n - \mu n}{\sigma \sqrt{n}} \longrightarrow N(0, 1)$$

по распределению при  $n \rightarrow \infty$ .

Получается, что

$$2 \lim_{n \rightarrow \infty} Y(n) \xrightarrow{\text{dist.}} N(0, 1).$$

Для оценки этого значения воспользуемся «правилом трёх сигм».

**Теорема 1.2.** Практически все значения нормально распределённой случайной величины  $\xi \sim N(\mu, \sigma^2)$  лежат в интервале  $(\mu - 3\sigma, \mu + 3\sigma)$ . Более строго — приблизительно с вероятностью 0,9973 значение нормально распределённой случайной величины лежит в указанном интервале.

Таким образом, приблизительно с вероятностью 0,9973

$$-\frac{3}{2} \leq \lim_{n \rightarrow \infty} Y(n) \leq \frac{3}{2}.$$

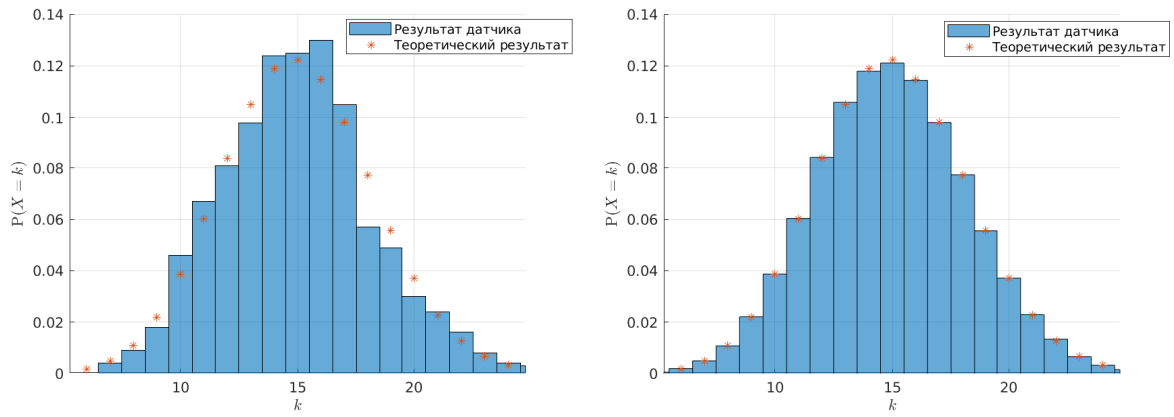


Рис. 1.1: Гистограмма биномиального распределения случайной величины с параметрами  $n = 50$ ,  $p = \frac{3}{10}$  при  $10^3$  (слева) и  $10^5$  (справа) испытаний.

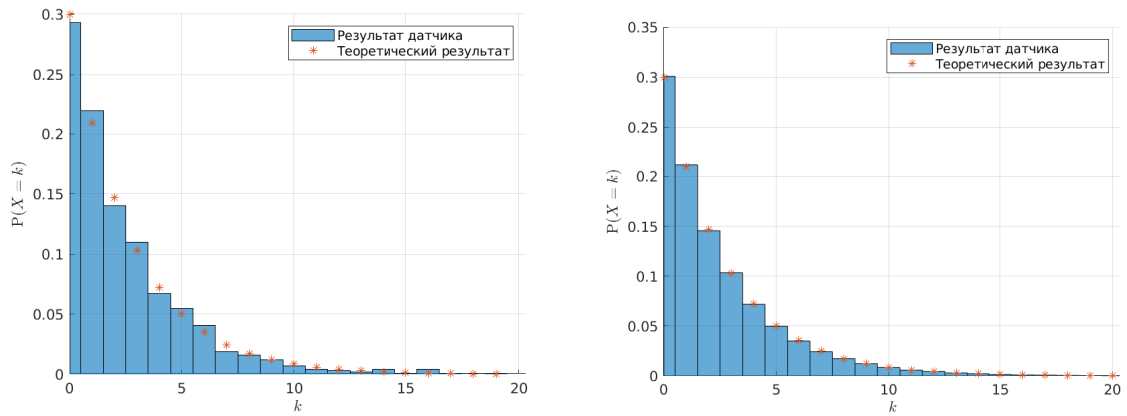


Рис. 1.2: Гистограмма геометрического распределения случайной величины с параметром  $p = \frac{3}{10}$  при  $10^3$  (слева) и  $10^5$  (справа) испытаний.

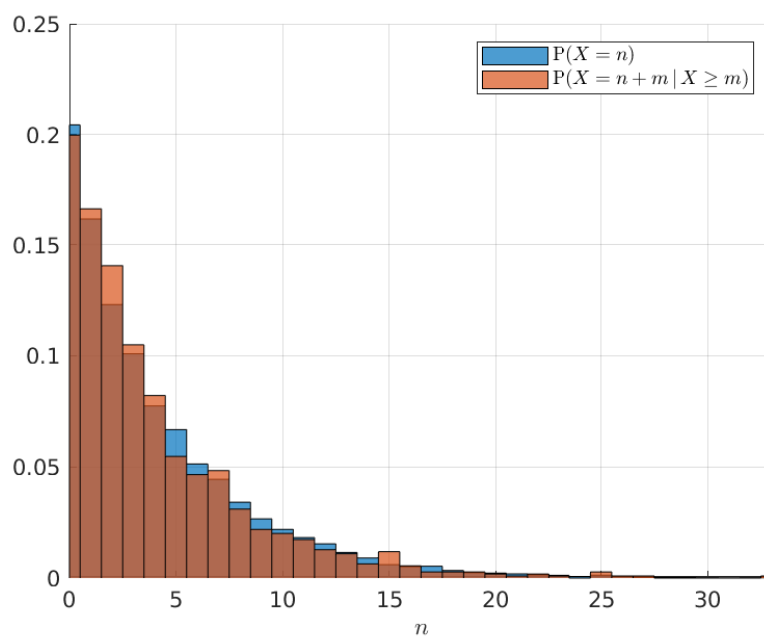


Рис. 1.3: Гистограмма геометрического распределения, демонстрирующая его свойство отсутствия памяти. Здесь задан параметр геометрического распределения  $p = \frac{2}{10}$ , а также «сдвиг»  $m = 10$ .

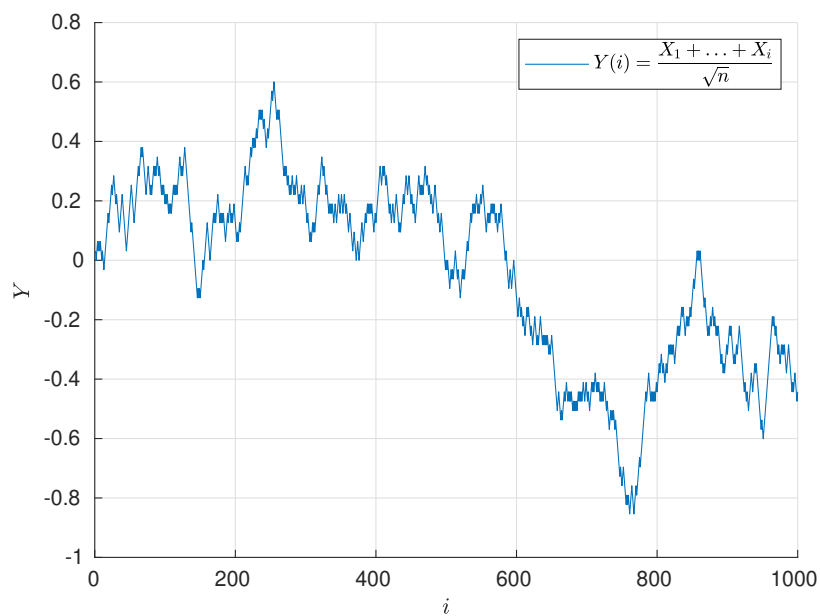


Рис. 1.4: Иллюстрация варианта поведения нормированной суммы  $Y(i)$  игры в орлянку на отрезке  $1 \leq i \leq 10^3$ .



## 2 Задание №2

1. Построить датчик сингулярного распределения, имеющий в качестве функции распределения канторову лестницу. С помощью критерия Колмогорова убедиться в корректности работы датчика.
2. Для канторовых случайных величин проверить свойство симметричности относительно  $\frac{1}{2}$  ( $X$  и  $(1 - X)$  распределены одинаково) и самоподобия относительно деления на 3 (условное распределение  $Y$  при условии  $Y \in [0, \frac{1}{3}]$  совпадает с распределением  $\frac{Y}{3}$ ) с помощью критерия Смирнова.
3. Вычислить значение математического ожидания и дисперсии с эмпирическими для разного объема выборок. Проиллюстрировать сходимость.

### 2.1 Построение датчика «канторовой» случайной величины

**Определение 2.1.** Пусть дано вероятностное пространство  $(\mathbb{R}, \mathcal{F}, \mathbb{P})$ , и на нем определена случайная величина  $\xi$  с распределением  $\mathbb{P}_\xi$ . Тогда *функцией распределения* случайной величины  $X$  называется функция  $F_\xi : \mathbb{R} \rightarrow [0, 1]$ , задаваемая формулой:

$$F_\xi(x) = \mathbb{P}(\xi \leq x) \equiv \mathbb{P}_\xi((-\infty, x]).$$

**Определение 2.2.** Функция распределения некоторой случайной величины называется *сингулярной*, если она непрерывна и ее множество точек роста имеет нулевую меру Лебега.

**Определение 2.3.** Из единичного отрезка  $C_0 = [0, 1]$  удалим интервал  $(\frac{1}{3}, \frac{2}{3})$ . Оставшееся множество обозначим за  $C_1$ . Множество  $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$  состоит из двух отрезков; удалим теперь из каждого отрезка его среднюю часть, и оставшееся множество обозначим за  $C_2$ . Повторив данную процедуру, то есть удаляя средние трети у всех четырех отрезков, получим  $C_3$ . Дальше таким же образом получаем последовательность замкнутых множеств  $C_0 \supset C_1 \supset C_2 \supset \dots \supset C_i \supset \dots$ . Пересечение

$$C = \bigcap_{i=0}^{\infty} C_i$$

называется *канторовым множеством*.

*Замечание 2.1.* Канторово множество так же можно определить как множество всех чисел от нуля до единицы, которые можно представить в троичной записи при помощи только нулей и двоек. То есть

$$C = \{0, \alpha_1 \alpha_2 \dots \alpha_i \dots_3 \mid \alpha_i = 0, 2\}.$$

**Утверждение 2.1.** Канторово множество имеет нулевую меру Лебега. [1]

**Определение 2.4.** Рассмотрим функцию  $K(x)$  такую, что в точках 0 и 1 значение функции принимается равным соответственно 0 и 1. Далее интервал  $(0, 1)$  разбивается на три равные части  $(0, \frac{1}{3})$ ,  $(\frac{1}{3}, \frac{2}{3})$  и  $(\frac{2}{3}, 1)$ . На среднем сегменте полагаем  $K(x) = \frac{1}{2}$ . Оставшиеся два сегмента снова разбиваются на три равные части каждый, и на средних сегментах  $K(x)$  полагается равной  $\frac{1}{4}$  и  $\frac{3}{4}$ . Каждый из оставшихся сегментов снова делится

на три части, и на внутренних сегментах  $K(x)$  определяется как постоянная, равная среднему арифметическому между соседними, уже определенными значениями  $K(x)$ . На остальных точках единичного отрезка определяется по непрерывности. Полученная функция называется *канторовой лестницей*.

*Замечание 2.2.* Из определения канторовой лестницы  $K(x)$  следует, что она действует на точки из канторова множества  $C$  по следующему правилу:

$$K(0, \alpha_1 \alpha_2 \dots \alpha_i \dots 3) = 0, \frac{\alpha_1}{2} \frac{\alpha_2}{2} \dots \frac{\alpha_i}{2} \dots 2.$$

Теперь рассмотрим случайную величину

$$Y = 0, \xi_1 \xi_2 \dots \xi_k \dots 2 = \sum_{k=1}^{\infty} \frac{\xi_k}{2^k}, \quad \text{где } \xi_k \sim \text{Bern}\left(\frac{1}{2}\right).$$

Такая случайная величина имеет равномерное распределение на отрезке  $[0, 1]$ , так как мы равновероятным образом выбираем знаки разложения числа  $\xi_k$  в двоичном представлении. Теперь рассмотрим искомую случайную величину  $X$ , имеющую в качестве функции распределения  $F_X(x)$  канторову лестницу  $K(x)$ . Образ каждой случайной величины  $Y$  для такой функции будет равен

$$K^{-1}(Y) = \sum_{k=1}^{\infty} \frac{2\xi_k}{3^k}.$$

Эта точка лежит в канторовом множестве.

**Теорема 2.1** (Метод обратной функции распределения). Пусть некоторая функция распределения  $F$  имеет обратную  $F^{-1}$ . Тогда функцией распределения случайной величины

$$\eta = F^{-1}(\xi),$$

где  $\xi$  — равномерно распределенная на отрезке  $[0, 1]$  случайная величина, является  $F$ .

**Доказательство.** Найдем функцию распределения случайной величины  $\eta$ :

$$F_\eta(x) = \mathbb{P}(\eta \leq x) = \mathbb{P}(F^{-1}(\xi) \leq x) = \mathbb{P}(\xi \leq F(x)) = F(x).$$

Таким образом, теорема доказана. ■

Из теоремы вытекает, что при помощи построенного ранее (см. раздел 1) генератора схемы Бернулли мы можем смоделировать случайную величину  $X$ , принимающую с вероятностью 1 значения из канторова множества  $C$  и имеющую канторову лестницу  $K(x)$  в качестве функции распределения  $F_X(x)$ , следующим образом:

$$X = \sum_{k=1}^{\infty} \frac{2\xi_k}{3^k}, \quad \text{где } \xi_k \sim \text{Bern}\left(\frac{1}{2}\right).$$

В программной реализации будем рассматривать частичные суммы. Для этого этого введем погрешность  $\varepsilon$  и найдем такое число  $n$ , при котором частичная сумма будет отличаться от бесконечной не более, чем на заданную погрешность.

$$\sum_{k=n}^{\infty} \frac{2\xi_k}{3^k} \leq 2 \sum_{k=n}^{\infty} \frac{1}{3^k} = \frac{1}{3^{n-1}} \leq \varepsilon,$$

↓

$$n \geq 1 - \lceil \log_3 \varepsilon \rceil \quad \forall \varepsilon < 1.$$

*Замечание 2.3.* Из выведенной формулы также видно, что для столь малой погрешности как  $\varepsilon = 10^{-9}$  достаточно использовать всего  $n = 20$  первых членов ряда.

## 2.2 Проверка корректности работы датчика

**Определение 2.5.** Пусть задана выборка  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ , где  $\xi_i \in \Xi, i = \overline{1, n}$ . Эмпирической (выборочной) функцией распределения, построенной на этой выборке, называется функция  $F_n(x)$ , равная доле таких значений  $\xi_i$ , что  $\xi_i < x$ . Или другими словами

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, x)}(\xi_i), \quad \text{где } \mathbb{I}_{(-\infty, x)}(\xi_i) = \begin{cases} 1, & \text{при } \xi_i < x, \\ 0, & \text{иначе.} \end{cases}$$

**Определение 2.6.** Пусть в некотором эксперименте доступна наблюдению случайная величина  $\xi$ , распределение которой  $\mathbb{P}$  полностью или частично неизвестно. Тогда любое утверждение относительно  $\mathbb{P}$  называется *статистической гипотезой*  $H$ .

**Теорема 2.2** (Критерий согласия Колмогорова). Обозначим нулевую гипотезу  $H_0$  как гипотезу о том, что выборка подчиняется распределению  $F(\xi) \in C^1(\Xi)$ . Введем статистику критерия

$$D_n = \sup_x |F_n(x) - F(x)|.$$

Тогда если гипотеза  $H_0$  верна, то  $\sqrt{n}D_n$  с ростом  $n$  сходится по распределению к случайной величине  $K$  с функцией распределения Колмогорова

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

*Замечание 2.4.* Гипотеза  $H_0$  отвергается, если при большом объеме выборки  $n$  статистика  $\sqrt{n}D_n$  превышает квантиль распределения  $K_\alpha$ , заданного уровня значимости  $\alpha$ , и принимается в противном случае. Здесь  $K_\alpha = F_K^{-1}(1 - \alpha)$ .

В рамках реализации подсчет  $K_\alpha$  в явном виде весьма трудоемкая задача. Поэтому мы будем рассчитывать  $p$ -значение для нашей статистики  $p_{value} = 1 - F_K(\sqrt{n}D_n)$ . Если  $p$ -значение оказалось ниже или равно установленному уровню значимости  $\alpha$ , то наша гипотеза отвергается и применяется альтернативная.

## 2.3 Симметричность «канторовой» случайной величины относительно $1/2$

**Утверждение 2.2** (Свойство симметричности относительно  $1/2$ ). Пусть  $X$  — случайная величина, с канторовой лестницей в качестве функции распределения. Тогда верно

$$F_X(x) = F_{1-X}(x).$$

Число испытаний	Размер выборки	Уровень значимости	Частота принятия гипотезы
$10^3$	$10^3$	0,05	0,958
$10^4$	$10^3$	0,05	0,9536
$10^3$	$10^4$	0,05	0,959
$10^3$	$10^3$	0,1	0,907

Таблица 1: Частота принятия гипотезы  $H_0$  о том, что построенный датчик случайной величины имеет канторову лестницу в качестве функции распределения. Из таблицы видно, что вероятность отклонить гипотезу не превышает допустимый уровень значимости.

**Д о к а з а т е л ь с т в о.** Как мы помним из построения датчика «канторовой» случайной величины, она представима в виде:

$$X = 2 \sum_{k=1}^{\infty} \frac{\xi_k}{3^k}, \quad \text{где } \xi_k \sim \text{Bern}(1/2), k = \overline{1, n}.$$

Тогда

$$1 - X = 1 - 2 \sum_{k=1}^{\infty} \frac{\xi_k}{3^k} = 2 \sum_{k=1}^{\infty} \frac{1}{3^k} - 2 \sum_{k=1}^{\infty} \frac{\xi_k}{3^k} = 2 \sum_{k=1}^{\infty} \frac{1 - \xi_k}{3^k} = 2 \sum_{k=1}^{\infty} \frac{\eta_k}{3^k},$$

при этом случайные величины  $\eta_k = 1 - \xi_k$  также имеют распределение Бернулли с параметром  $p = 1/2$ . Значит, случайные величины  $X$  и  $1 - X$  имеют одинаковое распределение. ■

Теперь эмперическим путем убедимся в выполнении этого свойства. Для этого нам потребуется *критерий Смирнова*. Он используется для проверки гипотезы о принадлежности двух независимых выборок одному и тому же распределению.

**Теорема 2.3** (Критерий однородности Смирнова). *Обозначим за нулевую гипотезу  $H_0$  гипотезу о том, что две исследуемые выборки объемами  $n$  и  $m$  с эмперическими функциями распределения  $F_n(x)$  и  $F_m(x)$  соответственно распределены по одному закону. Введем статистику критерия*

$$D_{n,m} = \sup_x |F_n(x) - F_m(x)|.$$

*Тогда если гипотеза  $H_0$  верна, то при увеличении объемов выборок  $n$  и  $m$  случайная величина  $\sqrt{\frac{nm}{n+m}} D_{n,m}$  будет сходиться по распределению к случайной величине  $K$  с функцией распределения Колмогорова*

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

## 2.4 Самоподобие «канторовой случайной величины относительно деления на 3»

**Утверждение 2.3** (Свойство самоподобия относительно деления на 3). *Пусть  $X$  — случайная величина с канторовой лестницей в качестве функции распределения. Тогда верно*

$$F_{X/3}(x) = F_{X|X \in [0, 1/3]}(x).$$

**Доказательство.** Заметим, что из построения датчика «канторовой» случайной величины вытекает, что случайная величина  $Y = X \mid X \in [0, 1/3]$  задается в виде

$$X = 2 \sum_{k=1}^{\infty} \frac{\xi_k}{3^k}, \quad \text{где } \xi_1 = 0, \xi_k \sim \text{Bern}(1/2), k = \overline{2, n}.$$

Тогда получается, что

$$Y = 2 \sum_{k=2}^{\infty} \frac{\xi_k}{3^k} = 2 \sum_{k=1}^{\infty} \frac{\xi_{k+1}}{3^{k+1}} = \frac{2}{3} \sum_{k=1}^{\infty} \frac{\xi_k}{3^k} = \frac{X}{3}.$$

■

Так же проверим получившийся результат при помощи критерия однородности Смирнова.

Число испытаний	$10^3$	$10^3$	$10^3$
Размер первой выборки	$10^2$	$10^3$	$10^3$
Размер второй выборки	$10^3$	$10^3$	$10^3$
Уровень значимости	0,05	0,05	0,01
Частота принятия гипотезы	0,956	0,955	0,912

Таблица 2: Частота принятия гипотезы  $H_0$  о том, что случайные величины  $X$  и  $(1 - X)$  имеют одинаковое распределение.

Число испытаний	$10^3$	$10^3$	$10^3$
Размер первой выборки	$10^2$	$10^3$	$10^3$
Размер второй выборки	$10^3$	$10^3$	$10^3$
Уровень значимости	0,05	0,05	0,01
Частота принятия гипотезы	0,956	0,955	0,915

Таблица 3: Частота принятия гипотезы  $H_0$  о том, что случайные величины  $X \mid X \in [0, 1/3]$  и  $X/3$  имеют одинаковое распределение.

## 2.5 Математическое ожидание и дисперсия «канторовой» случайной величины

Вычислим значение математического ожидания для построенной случайной величины  $X$ :

$$\mathbb{E} X = \mathbb{E} \sum_{k=1}^{\infty} \frac{2\xi_k}{3^k} = \sum_{k=1}^{\infty} \frac{2}{3^k} \mathbb{E} \xi_k = \sum_{k=1}^{\infty} \frac{2}{3^k} \cdot \frac{1}{2} = \frac{\frac{1}{3}}{1 - \frac{1}{3}} = \frac{1}{2}.$$

Теперь, помня о независимости случайных величин  $\xi_k$   $k \in \mathbb{N}$ , вычислим значение дисперсии

$$\mathbb{V}ar X = \mathbb{V}ar \sum_{k=1}^{\infty} \frac{2\xi_k}{3^k} = \sum_{k=1}^{\infty} \left( \frac{2}{3^k} \right)^2 \mathbb{V}ar \xi_k = \sum_{k=1}^{\infty} \frac{4}{9^k} \cdot \frac{1}{4} = \frac{\frac{1}{9}}{1 - \frac{1}{9}} = \frac{1}{8}.$$

*Замечание 2.5.* При подсчете мы использовали известные значения для математического ожидания и дисперсии бернуллиевой случайной величины  $\xi \sim \text{Bern}(p)$ :

$$\mathbb{E} \xi = p \quad \text{и} \quad \text{Var} \xi = p(1 - p).$$

Для сравнения практического и теоретического результатов построим также графики *выборочного среднего*  $\bar{X}$  и *несмещенной выборочной дисперсии*  $S^2$ , задаваемых формулами:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{и} \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

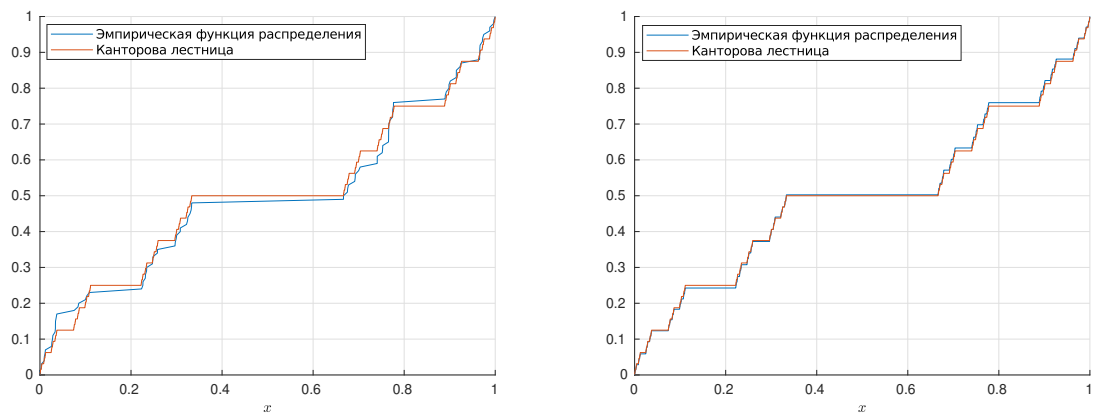


Рис. 2.1: Эмпирическая и теоретическая функции распределения «канторовой» случайной величины  $X$  при выборке из 100 испытаний (слева) и  $10^4$  испытаний (справа).

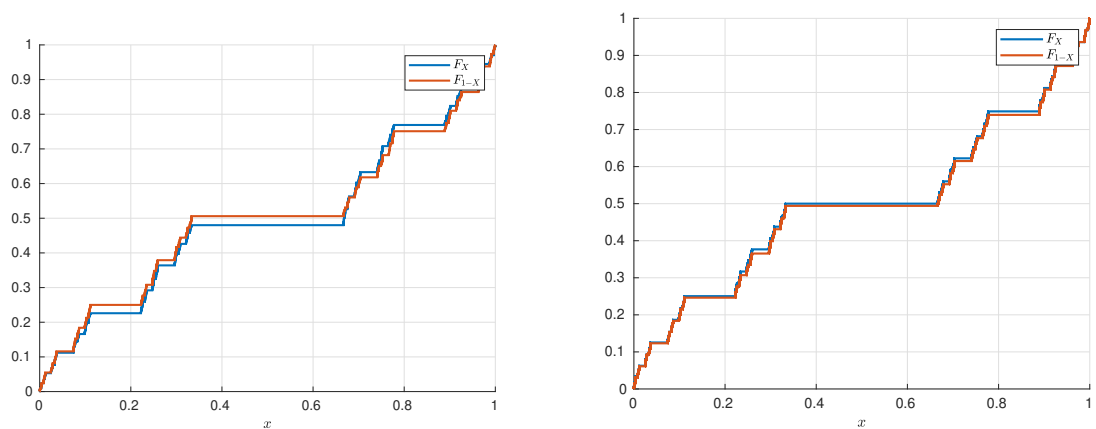


Рис. 2.2: График, иллюстрирующий свойство симметричности относительно  $1/2$  «канторовой» случайной величины. Представлены эмперические функции распределения при объемах выборок  $10^3$  (слева) и  $10^4$  (справа).

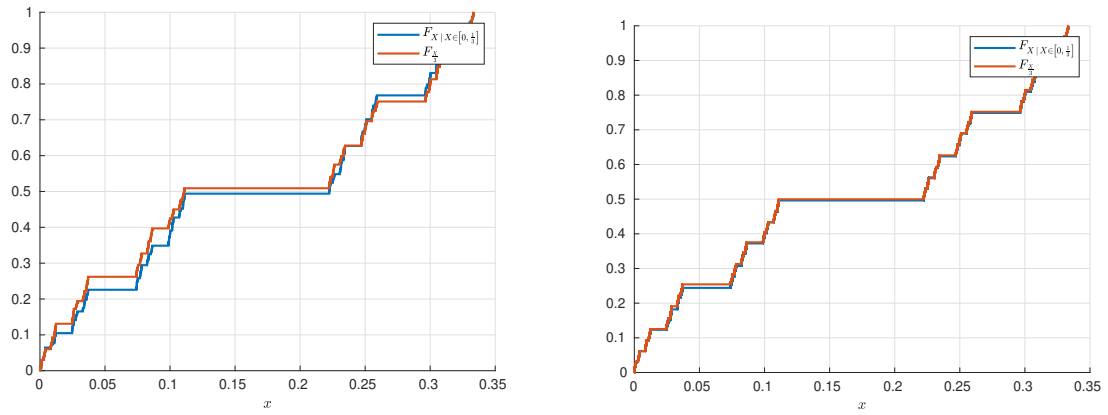


Рис. 2.3: График, иллюстрирующий свойство самоподобия относительно деления на 3 «канторовой» случайной величины. Представлены эмперические функции распределения при объемах выборок  $10^3$  (слева) и  $10^4$  (справа).

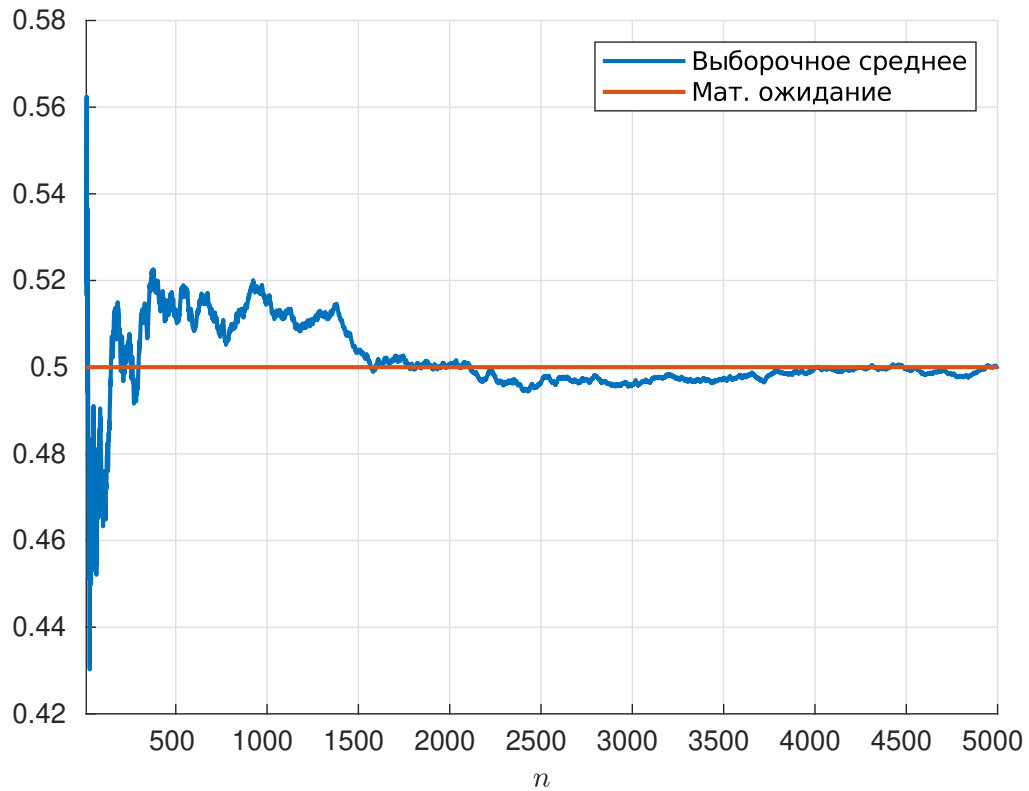


Рис. 2.4: Эмпирическое значение математического ожидания «канторовой» случайной величины  $X$  в зависимости от объема выборки  $10 \leq n \leq 5000$ .



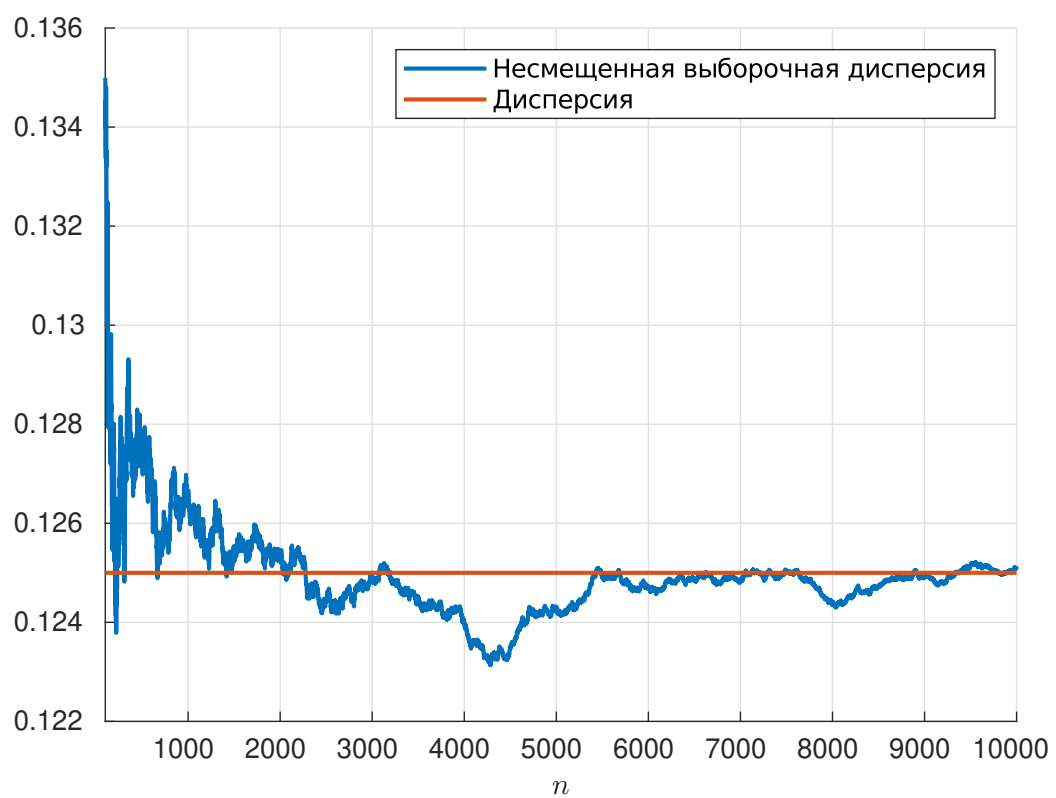


Рис. 2.5: Несмещенная выборочная дисперсия и теоретическая дисперсия «канторовой» случайной величины  $X$  в зависимости от объема выборки  $100 \leq n \leq 10000$ .

### 3 Задание №3

1. Построить датчик экспоненциального распределения. Проверить для данного распределения свойство отсутствия памяти. Пусть  $X_1, X_2, \dots, X_n$  независимо распределенные случайные величины с параметрами  $\lambda_1, \lambda_2, \dots, \lambda_n$  соответственно. Найти распределение случайной величины  $Y = \min\{X_1, X_2, \dots, X_n\}$ .
2. На основе датчика экспоненциального распределения построить датчик пуассоновского распределения.
3. Построить датчик пуассоновского распределения как предел биномиального распределения. С помощью критерия хи-квадрат Пирсона убедиться, что получен датчик распределения Пуассона.
4. Построить датчик стандартного 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.

#### 3.1 Построение датчика экспоненциальной случайной величины

**Определение 3.1.** Будем говорить, что случайная величина  $X$  имеет экспоненциальное распределение с параметром  $\lambda > 0$ , если ее функция распределения имеет вид

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

Будем обозначать такие случайные величины

$$X \sim \text{Exp}(\lambda).$$

Для того чтобы построить датчик экспоненциально распределенной с параметром  $\lambda$  случайной величины  $X$ , воспользуемся методом обратной функции распределения (смотри теорему 2.1). Получается, что такую случайную величину можно представить в виде:

$$X = F_x^{-1}(\xi) = -\frac{1}{\lambda} \ln(1 - \xi),$$

где  $\xi$  — равномерно распределенная на отрезке  $[0, 1]$  случайная величина.

#### 3.2 Отсутствие памяти у экспоненциального распределения

**Утверждение 3.1** (Свойство отсутствия памяти). Пусть  $X \sim \text{Exp}(\lambda)$ , тогда для любых  $t \neq 0$  и  $s$  справедливо:

$$\mathbb{P}(X \geq s + t \mid X \geq t) = \mathbb{P}(X \geq s).$$

**Доказательство.** Рассмотрим левую часть равенства:

$$\mathbb{P}(X \geq s + t \mid X \geq t) = \frac{\mathbb{P}(X \geq s + t, X \geq t)}{\mathbb{P}(X \geq t)} = \frac{\mathbb{P}(X \geq s + t)}{\mathbb{P}(X \geq t)}.$$

Таким образом получаем, утверждение эквивалентно тому, что

$$\mathbb{P}(X \geq s + t) = \mathbb{P}(X \geq t)\mathbb{P}(X \geq s).$$

Из определения функции распределения  $F_X(t) = \mathbb{P}(X < t) = 1 - \mathbb{P}(X \geq t)$  получаем, что

$$e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t}.$$

Последнее равенство точно верно. Таким образом, утверждение доказано. ■

### 3.3 Распределение минимума экспоненциальных случайных величин

Рассмотрим теперь случайную величину  $Y = \min\{X_1, X_2, \dots, X_n\}$ , где  $X_i, i = \overline{1, n}$  есть независимо распределенные экспоненциальные случайные величины с параметрами  $\lambda_1, \lambda_2, \dots, \lambda_n$  соответственно, и найдем её функцию распределения:

$$\begin{aligned} F_Y(x) &= \mathbb{P}(Y < x) = \mathbb{P}(\min_{i=1, n} X_i < x) = 1 - \mathbb{P}(\min_{i=1, n} X_i \geq x) = \\ &= 1 - \mathbb{P}(X_1 \geq x, X_2 \geq x, \dots, X_n \geq x) = 1 - \prod_{i=1}^n \mathbb{P}(X_i \geq x) = 1 - \prod_{i=1}^n (1 - F_{X_i}(x)). \end{aligned}$$

Таким образом функция распределения случайной величины  $Y$  представима в виде

$$F_Y(x) = 1 - \prod_{i=1}^n (1 - 1 + e^{-\lambda_i x}) = 1 - \prod_{i=1}^n e^{-\lambda_i x} = 1 - e^{-(\lambda_1 + \lambda_2 + \dots + \lambda_n)x}.$$

Получается, что заданная случайная величина  $Y$  имеет экспоненциальное распределение с параметром  $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_n$ .

### 3.4 Построение датчика распределения Пуассона

**Определение 3.2.** Будем говорить, что случайная величина  $X$  имеет распределение Пуассона с параметром  $\lambda > 0$ , если

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{где } k \in \mathbb{N}_0.$$

Будем обозначать такие случайные величины

$$X \sim \text{Pois}(\lambda).$$

Для построения датчика Пуассоновской случайной величины докажем вспомогательное утверждение.

**Утверждение 3.2** (О распределении суммы экспоненциальных случайных величин). Пусть  $S_n = \sum_{i=1}^n \xi_i$ , где  $\xi_i \sim \text{Exp}(\lambda)$ ,  $i = \overline{1, n}$ . Тогда

$$F_{S_n}(x) = 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{\lambda^k x^k}{k!}.$$

**Д о к а з а т е л ь с т в о.** Для доказательства достаточно показать, что случайная величина  $S_n$  имеет плотность распределения, равную

$$\rho_{S_n}(x) = e^{-\lambda x} \frac{\lambda^n x^{n-1}}{(n-1)!}, \quad x \geq 0.$$

Докажем это методом математической индукции. База индукции очевидна. Теперь пусть для шага  $n$  выполнена предыдущая формула. Воспользуемся формулой свертки плотностей распределений для нахождения  $\rho_{S_{n+1}}$

$$\rho_{S_{n+1}}(x) = \int_0^x \rho_{S_1}(x-t) \cdot \rho_{S_n}(t) dt = \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \int_0^x (x-t)^{n-1} dt = \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \frac{x^n}{n}.$$

Теорема доказана. ■

Пусть  $t > 0$ . Рассмотрим независимые случайные величины  $\{\xi_k\}_{k \geq 1}$ , имеющие показательное распределение с параметром  $\lambda$ . Как и в предыдущем утверждении, положим  $S_n = \sum_{k=1}^n \xi_k$ . Наконец, обозначим  $X = \max\{n \geq 0 \mid S_n < t\}$ , полагая  $S_0 = 0$ . Докажем теперь, что  $X \sim \text{Pois}(\lambda t)$ .

Для этого найдем вероятность того, что  $X = n$ . При  $n = 0$

$$\mathbb{P}(X = n) = \mathbb{P}(\xi_1 \geq t) = e^{-\lambda t}.$$

При  $n > 0$ , поскольку  $\xi_k \geq 0$ , то согласно утверждению 3.2

$$\begin{aligned} \mathbb{P}(X = n) &= \mathbb{P}(S_n < t, S_{n+1} \geq t) = \mathbb{P}(S_n < t) - \mathbb{P}(S_n < t, S_{n+1} < t) = \\ &= \mathbb{P}(S_n < t) - \mathbb{P}(S_{n+1} < t) = F_{S_n}(t) - F_{S_{n+1}}(t) = \frac{\lambda^n t^n}{n!} e^{-\lambda t}. \end{aligned}$$

Таким образом мы получали способ построения пуассоновской случайной величины. Нужно просто брать показательные случайные величины с параметром  $\lambda = 1$  и смотреть сумма сколько первых из них меньше параметра пуассоновского распределения.

### 3.5 Построение датчика распределения Пуассона как предел биномиального распределения

Биномиальное распределение сходится к распределению Пуассона, так как число испытаний уходит в бесконечность, в то время как произведение  $np$  остается фиксированным или, по крайней мере,  $p$  стремится к нулю. Поэтому распределение Пуассона с параметром  $\lambda = np$  можно использовать как приближение к  $\text{Bin}(n, p)$  биномиального распределения, если  $n$  достаточно велико, а  $p$  достаточно мало. Подробное доказательство этого факта, а также эмперические правила выбора параметров биномиального распределения и оценки точности приближения можно посмотреть в [4] и [5][Глава 4]. Мы же попытаемся лишь обосновать сам факт этого.

Пусть случайная величина  $\xi$  имеет биномиальное распределение с параметрами  $n$  и  $p$ . И введем обозначение:

$$\mathbb{P}_n(\xi = k) = C_n^k p^k (1-p)^{n-k}.$$

Зафиксируем значение  $\lambda = np$ , которое является математическим ожиданием биномиального распределения и будем устремлять параметр  $n$  к бесконечности:

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P}_n(k) &= \lim_{n \rightarrow \infty} \left[ C_n^k \cdot \left( \frac{\lambda}{n} \right)^k \cdot \left( 1 - \frac{\lambda}{n} \right)^{n-k} \right] = \\ &= \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot \lim_{n \rightarrow \infty} \left[ \frac{n!}{n^k(n-k)!} \left( 1 - \frac{\lambda}{n} \right)^{-k} \right] = \\ &= \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot \lim_{n \rightarrow \infty} \left[ \frac{n(n-1) \dots (n-k+1)}{n^k} \left( 1 - \frac{\lambda}{n} \right)^{-k} \right] = \frac{\lambda^k}{k!} e^{-\lambda}.\end{aligned}$$

### 3.6 Проверка корректности работы датчика

Для проверки правильности работы датчика пуассоновской случайной величины воспользуемся *критерием согласия Пирсона* (или критерием согласия  $\chi$ -квадрат). Это один из наиболее популярных критериев для проверки гипотезы о принадлежности наблюдаемой выборки некому теоретическому закону распределения.

**Теорема 3.1** (Критерий согласия Пирсона). *Обозначим нулевую гипотезу  $H_0$  как гипотезу о том, что выборка  $\xi_1, \xi_2 \dots \xi_n$  подчиняется закону распределения  $\mathbb{P}$ . Обозначим за  $n_k$  количество элементов в выборке, равных  $k$ . За  $r$  обозначим количество различных элементов выборки. А за  $p_k$  — вероятность выпадения значения в теоретическом распределении  $p_k = \mathbb{P}(\xi = k)$ . Введем статистику критерия*

$$X_n^2 = n \sum_{k=1}^r \frac{\left( \frac{n_k}{n} - p_k \right)^2}{p_k}.$$

*Тогда если гипотеза  $H_0$  верна, то статистика  $X_n^2$  имеет  $\chi$ -квадрат распределение с  $r - 1$  степенью свободы.*

**Замечание 3.1.** Функция распределения  $\chi^2$  с  $l$  степеней свободы выглядит следующим образом:

$$F_{\chi_l^2} = \frac{\gamma(l/2, x/2)}{\Gamma(l/2)}, \quad \text{где } \Gamma \text{ и } \gamma \text{ обозначают полную и неполную гамма-функции.}$$

Такое определение делает почти невозможным аналитический поиск квантилей этой функции распределения. По этой причине для данных расчетов мы будем пользоваться встроенной в систему **Matlab** функцией `chi2inv(alpha, 1)`, которая возвращает квантиль  $\chi^2$ -распределения с 1 степенью свободы порядка `alpha`.

### 3.7 Построение датчика нормального распределения методом моделирования парами с переходом в полярные координаты

**Определение 3.3.** Будем говорить, что случайная величина  $X$  имеет нормальное распределение с параметрами  $\mu$  и  $\sigma^2$ , если ее плотность распределения задается формулой

$$\rho_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}.$$

Число испытаний	Размер выборки	Уровень значимости	Частота принятия гипотезы
$10^2$	$10^3$	0,05	0,953
$10^3$	$10^3$	0,05	0,958

Таблица 4: Частота принятия гипотезы  $H_0$  о том, что построенный датчик случайной величины имеет распределение Пуассона, используя критерий согласия Пирсона.

Будем обозначать такие случайные величины

$$X \sim N(\mu, \sigma^2).$$

*Замечание 3.2.* Будем говорить, что случайная величина  $X$  имеет стандартное нормальное распределение, если  $X \sim N(0, 1)$ .

Рассмотрим случайную величину  $Z = \sqrt{2\xi} \sin \eta$ , где  $\xi \sim \text{Exp}(1)$ ,  $\eta \sim U[0, 2\pi] \sim 2\pi U[0, 1]$ . Тогда

$$\begin{aligned} \mathbb{P}(Z < x) &= \mathbb{P}(\sqrt{2\xi} \sin \eta < x) = \iint_{\{(\xi, \eta) \mid \sqrt{2\xi} \sin \eta < x\}} \frac{e^{-\xi}}{2\pi} d\xi d\eta = \left\{ \xi = \frac{\psi^2}{2} \right\} = \\ &= \iint_{\{(\psi, \eta) \mid \psi \sin \eta < x\}} \frac{e^{-\frac{\psi^2}{2}}}{2\pi} \psi d\psi d\eta = \{X = \psi \cos \eta, Y = \psi \sin \eta\} = \\ &= \iint_{\{(X, Y) \mid Y < x\}} \frac{e^{-\frac{X^2}{2}} e^{-\frac{Y^2}{2}}}{2\pi} dX dY = \int_{-\infty}^{+\infty} \frac{e^{-\frac{X^2}{2}}}{\sqrt{2\pi}} dX \int_{-\infty}^x \frac{e^{-\frac{Y^2}{2}}}{\sqrt{2\pi}} dY = \int_{-\infty}^x \frac{e^{-\frac{Y^2}{2}}}{\sqrt{2\pi}} dY. \end{aligned}$$

Таким образом случайная величина  $Z$  имеет стандартное нормальное распределение.

### 3.8 Проверка равенства математического ожидания и дисперсии

В этом разделе мы проверим построенный датчик на то, действительно ли математическое ожидание смоделированной случайной величины равно нулю, а дисперсия — единице. Для первой задачи используют *одновыборочный критерий Стьюдента* (или *t-критерий*).

**Теорема 3.2** (Критерий Стьюдента). *Обозначим нулевую гипотезу  $H_0$  как гипотезу о том, что математическое ожидание выборки  $X = \{X_1, X_2, \dots, X_n\}$  равно некоторому известному значению  $\mu$ . Введем статистику критерия:*

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

*Тогда если гипотеза  $H_0$  верна, то статистика  $t$  имеет распределение Стьюдента с  $n - 1$  степенью свободы.*

*Замечание 3.3.* Как и в случае с распределением хи-квадрат (смотри замечание 3.1), по причине сложности вычисления обратной функции распределения Стьюдента, воспользуемся встроенной в систему **Matlab** функцией `tinv(alpha, r)`, считающей квантиль распределения Стьюдента с **r** степенями свободы порядка **alpha**.

Число испытаний	Размер выборки	Уровень значимости	Частота принятия гипотезы
$10^3$	$10^3$	0,05	0,951
$10^4$	$10^4$	0,05	0,9527

Таблица 5: Частота принятия гипотезы  $H_0$  о том, что построенный датчик случайной величины имеет равное нулю матожидание, используя одновыборочный критерий Стьюдента.

Теперь проверим равенство дисперсии единице. Для этого воспользуемся *двувыборочным критерием Фишера* (или f-критерием).

**Теорема 3.3** (Критерий Фишера). *Обозначим нулевую гипотезу  $H_0$  как гипотезу о том, что дисперсии двух выборок  $X = \{X_k\}_{k=1}^n$  и  $Y = \{Y_k\}_{k=1}^m$  равны друг другу. Введем статистику критерия*

$$F = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}, \quad \text{где } \hat{\sigma}^2 - \text{выборочная дисперсия.}$$

Тогда если гипотеза  $H_0$  верна, то статистика  $F$  имеет распределение Фишера  $F(n-1, m-1)$ .

*Замечание 3.4.* Так же, как и в предыдущем случае, мы будем пользоваться встроенной в систему Matlab функцию `finv(alpha, n, m)`, считающую квантиль распределения Фишера с параметрами `n` и `m` порядка `alpha`.

*Замечание 3.5.* Для того, чтобы использовать критерий Фишера, нам нужна еще одна выборка, которая бы точно имела дисперсию, равную единице. Можно использовать, например, построенный датчик случайной величины игры в орлянку (смотри раздел 1).

Число испытаний	Размер выборки	Уровень значимости	Частота принятия гипотезы
$10^3$	$10^3$	0,05	0,993
$10^4$	$10^4$	0,05	0,9833

Таблица 6: Частота принятия гипотезы  $H_0$  о том, что построенный датчик случайной величины имеет дисперсию, равную единице, используя критерий Фишера.

Важно отметить, что в отличие от рассмотренных ранее критериев, критерии Стьюдента и Фишера имеют двухстороннюю критическую область. Это значит, что если мы знаем квантиль  $K_r$  распределения статистики критерия порядка  $r$ , то при полученном значении статистики  $S$  и заданном уровне значимости  $\alpha$  нулевая гипотеза  $H_0$  должна быть принята при условии, что

$$K_{\alpha/2} < S < K_{1-\alpha/2}.$$

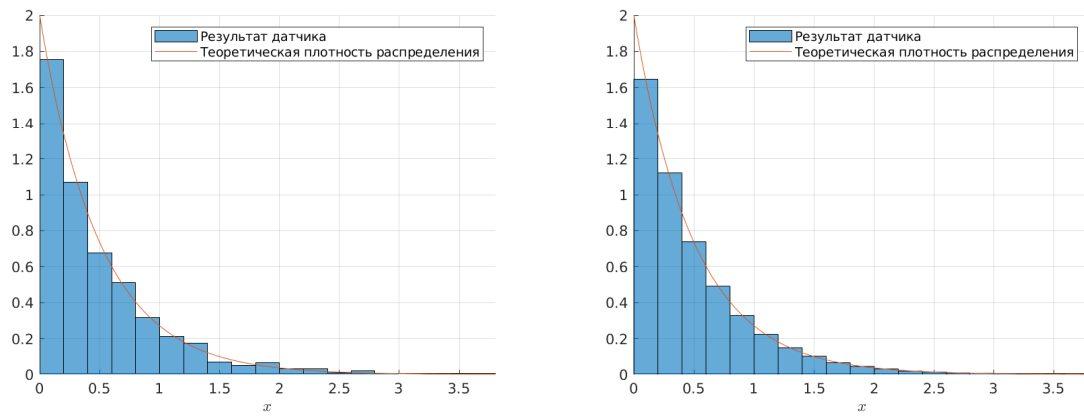


Рис. 3.1: Гистограмма экспоненциального распределения случайной величины с параметром  $\lambda = 2$  при  $10^3$  (слева) и  $10^5$  (справа) испытаний.

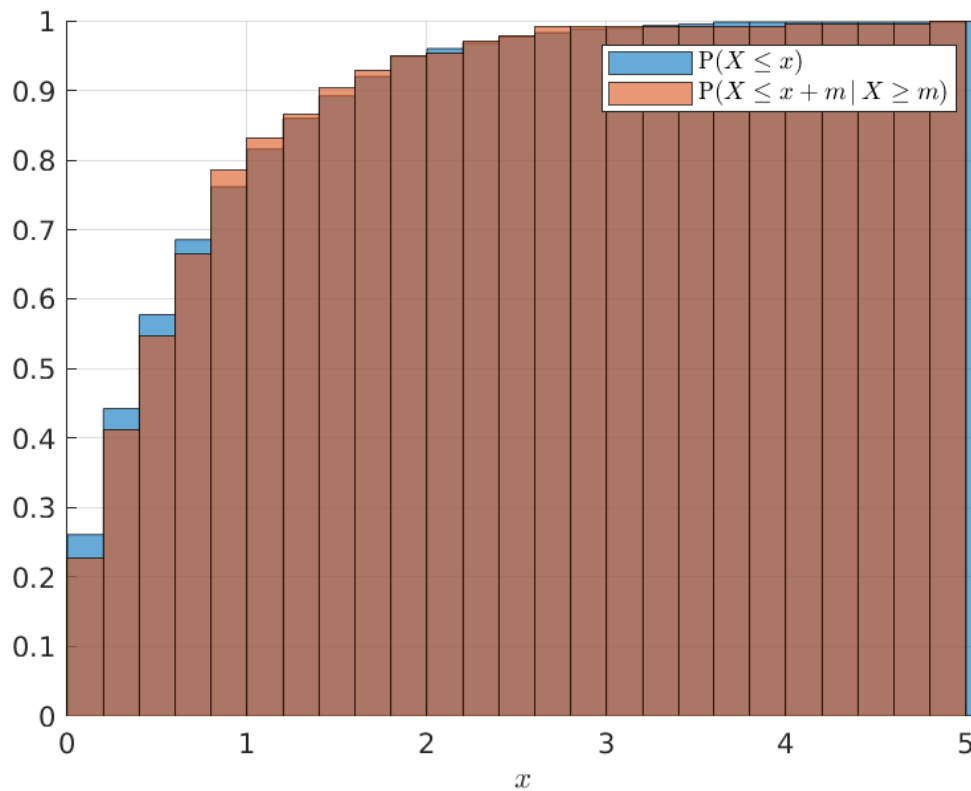


Рис. 3.2: Гистограмма экспоненциального распределения, демонстрирующая его свойство отсутствия памяти. Здесь задан параметр распределения  $\lambda = \frac{3}{2}$ , а также «сдвиг»  $m = 1$ . Проведено  $10^3$  испытаний.



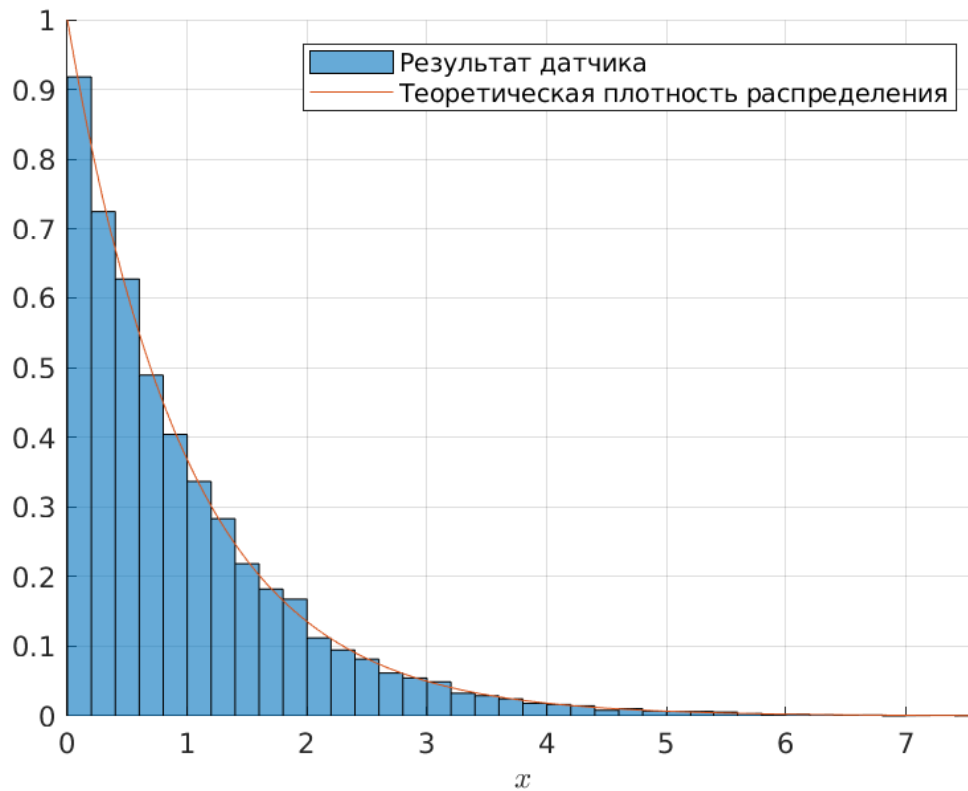


Рис. 3.3: Гистограмма распределения случайной величины  $Y = \min_{i=1,n} X_i$ . Здесь  $\lambda_1 = \lambda_2 = \dots = \lambda_n = \frac{1}{10}$ ,  $n = 10$ . Число испытаний  $10^4$ .

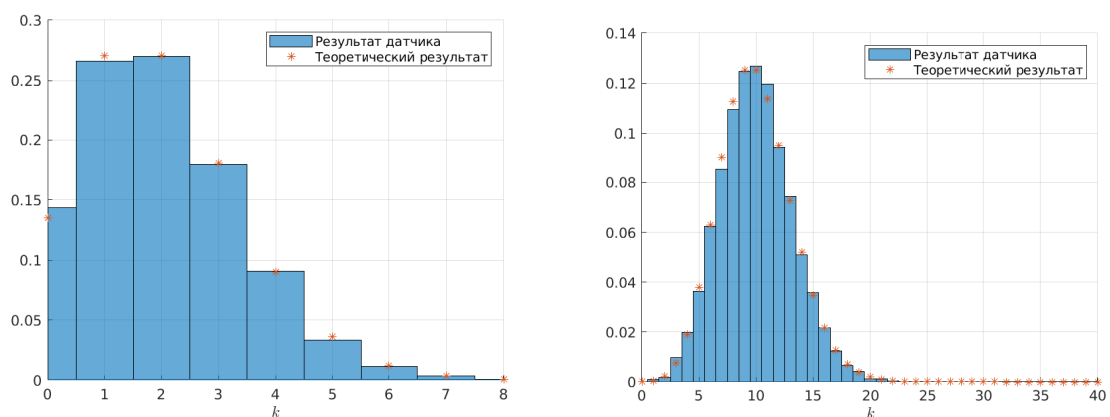


Рис. 3.4: Гистограмма распределения Пуассона случайной величины с параметром  $\lambda = 2$  (слева) и  $\lambda = 10$  (справа) при  $10^4$  испытаний.

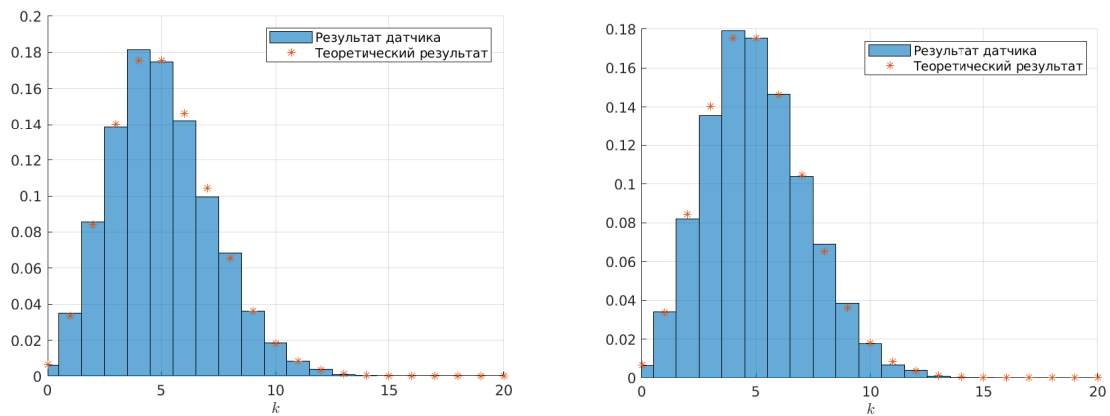


Рис. 3.5: Гистограмма распределения Пуассона случайной величины с параметром  $\lambda = 5$  как приближение биномиальным распределением с параметром  $n = 5 \cdot 10^2$  (слева) и  $n = 5 \cdot 10^4$  (справа) при  $10^4$  испытаний.

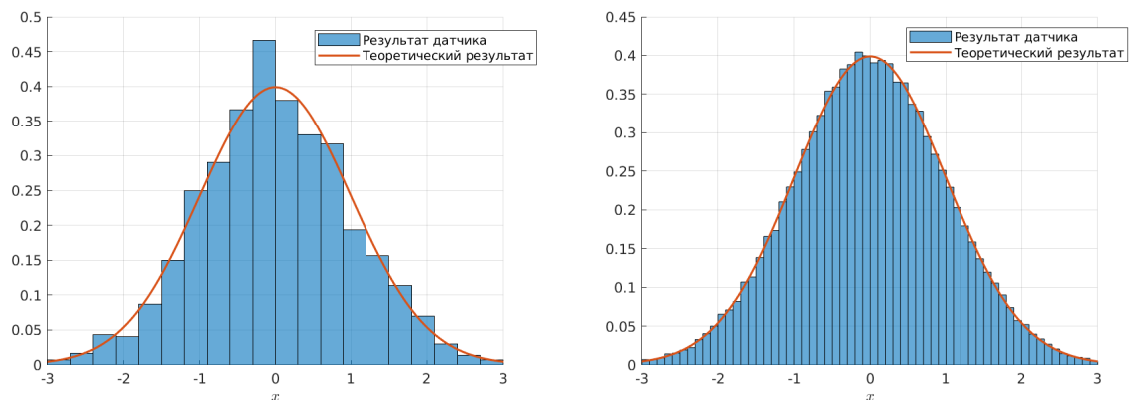


Рис. 3.6: Гистограмма стандартного нормального распределения случайной величины при  $10^3$  (слева) и  $10^5$  (справа) испытаний.

## 4 Задание №4

1. Построить датчик распределения Коши.
2. На основе датчика распределения Коши с помощью метода фон Неймана построить датчик стандартного нормального распределения. При помощи функции `normal probability plot` убедиться в корректности работы построенного датчика и обосновать наблюдаемую линейную зависимость.
3. Сравнить скорость моделирования стандартного нормального распределения в заданиях 3 и 4.

### 4.1 Построение датчика распределения Коши

**Определение 4.1.** Будем говорить, что случайная величина  $X$  имеет *распределение Коши* с параметрами  $a$  и  $b$ , если ее функция распределения имеет вид:

$$F_X(x) = \frac{1}{\pi} \operatorname{arctg} \left( \frac{x - a}{b} \right) + \frac{1}{2}.$$

Будем обозначать такие случайные величины

$$X \sim C(a, b).$$

Функция распределения Коши имеет обратную, поэтому воспользовавшись методом обратной функции распределения (смотри теорему 2.1), получим

$$X = F_X^{-1}(\xi) = a + b \operatorname{tg} [\pi(\xi - 1/2)], \quad \text{где } \xi \sim U[0, 1].$$

### 4.2 Построение датчика стандартного нормального распределения методом фон Неймана

Метод фон Неймана заключается в моделировании нормального распределения путем мажорирования плотностью распределения Коши с параметрами  $a$  и  $b$ . Для достижения наилучшей оценки, будем подбирать значения параметров  $a$  и  $b$ . Выпишем плотности стандартного нормального распределения и распределения Коши:

$$\rho_N(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$
$$\rho_C(x) = \frac{1}{\pi} \cdot \frac{b}{(x - a)^2 + b^2}.$$

При моделировании будем использовать следующий алгоритм:

1. возьмем некоторое число  $k > 0$  такое, что  $\rho_N(x) \leq k\rho_C(x) \forall x \in \mathbb{R}$ ;
2. рассмотрим значение случайной величины  $\hat{\xi} = \xi$ , где  $\xi \sim C(a, b)$ ;
3. сгенерируем случайную величину  $\hat{\eta} = \eta$ , где  $\eta \sim \operatorname{Bern} \left( \frac{\rho_N(\hat{\xi})}{k\rho_C(\hat{\xi})} \right)$ ;
4. если  $\hat{\eta} = 1$ , то  $\hat{\xi}$  является значением случайной величины из распределения с плотностью  $\rho_N(x)$ , иначе, продолжаем моделирование, начиная с пункта 2.

Данный алгоритм работает тем быстрее, чем ближе отношение  $\frac{\rho_N(x)}{k\rho_C(x)}$  к единице, поэтому в качестве  $k$  возьмем  $k^* = \min_{a,b} \max_x \frac{\rho_N(x)}{\rho_C(x)}$ . Рассмотрим это отношение:

$$\frac{\rho_N(x)}{\rho_C(x)} = \frac{\sqrt{\pi}}{b\sqrt{2}} \cdot e^{-\frac{x^2}{2}} \cdot [(x-a)^2 + b^2].$$

Допустим, что оптимальное значение параметра  $a = 0$ . Рассмотрим вспомогательную функцию

$$g(x) = e^{-\frac{x^2}{2}} (x^2 + b^2)$$

и найдем ее точки максимума:

$$g'(x) = xe^{-\frac{x^2}{2}} (2 - b^2 - x^2) \implies \begin{cases} x = 0, & |b| > \sqrt{2}, \\ x = \pm\sqrt{2-b^2}, & 0 < |b| \leq \sqrt{2}, \end{cases}$$

Следовательно,

$$k^* = \min \left\{ \min_{|b| > \sqrt{2}} \sqrt{\frac{\pi}{2}} b, \min_{0 < |b| \leq \sqrt{2}} \frac{\sqrt{2\pi}}{b} e^{\frac{b^2}{2}-1} \right\}.$$

Поскольку  $k > 0$  и  $b > 0$ , то найдем минимум еще одной вспомогательной функции:

$$h(b) = \frac{e^{\frac{b^2}{2}-1}}{b}, \quad h'(b) = \frac{1-b^2}{b^2} e^{\frac{b^2}{2}-1} \implies b = 1.$$

Таким образом, при  $a^* = 0$ ,  $b^* = 1$ :

$$k^* = \min \left\{ \sqrt{\pi}, \sqrt{\frac{2\pi}{e}} \right\} = \sqrt{\frac{2\pi}{e}}.$$

Напоследок покажем, что  $a^* = 0$  действительно является оптимальным значением параметра:

$$\begin{aligned} k^* &= \min_{a,b} \max_x \frac{\rho_N(x)}{\rho_C(x)} = \min_{a,b} \max_x \frac{\sqrt{\pi}}{b\sqrt{2}} \cdot e^{-\frac{x^2}{2}} \cdot [(x-a)^2 + b^2] = \\ &= \min_a \left\{ \min_{b > \sqrt{2}} \frac{\rho_N(x)}{\rho_C(x)} \Big|_{x=0}, \min_{0 < b \leq \sqrt{2}} \frac{\rho_N(x)}{\rho_C(x)} \Big|_{x=\pm\sqrt{2-b^2}} \right\} > \\ &> \min_a \left\{ \min_{b > \sqrt{2}} \frac{\sqrt{\pi}}{b\sqrt{2}} (a^2 + b^2), \min_{0 < b \leq \sqrt{2}} (\sqrt{2-b^2} + |a|) \right\}. \end{aligned}$$

Минимум же последнего выражения достигается при  $a = 0$ .

### 4.3 Дополнительно о методе фон Неймана

Проиллюстрируем результат работы датчика при помощи функции `normal probability plot`. Данная функция строит график распределения в следующих осях: на оси абсцисс откладываются точки выборки, а на оси ординат — квантили стандартного нормального распределения.

Рассмотрим случайную величину  $X \sim N(\mu, \sigma^2)$ . Тогда ее функция распределения

$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Произведем замену переменной  $s = \frac{t-\mu}{\sigma}$ , тогда

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{s^2}{2}} ds = F_N\left(\frac{x-\mu}{\sigma}\right),$$

где  $F_N(x)$  — функция стандартного нормального распределения.

Получается, что квантили различных нормальных распределений связаны между собой линейно. То есть любая случайная величина  $X \sim N(\mu, \sigma^2)$  представима в виде  $X = \sigma Y + \mu$ , где  $Y \sim N(0, 1)$ .

Функция **normal probability plot** для каждой нормально распределенной выборки должна показать прямую со сдвигом  $\mu$  и угловым коэффициентом  $\sigma$ .

В конце хотелось бы сравнить скорости работы моделирования стандартной нормальной случайной величины при методах моделирования парами (смотри предыдущее задание) и фон Неймана. Для этого мы приведем график скорости работы каждого метода в зависимости от объема генерируемой выборки. Стоит отметить, что программное решение задачи написано в системе **Matlab**, где на системном уровне реализованы матричные вычисления, которые работают гораздо быстрее циклов. Метод моделирования парами работает использует только матричные операции, в то время как в методе фон Неймана используется цикл. К тому же матричных операций в методе фон Неймана также используется больше.

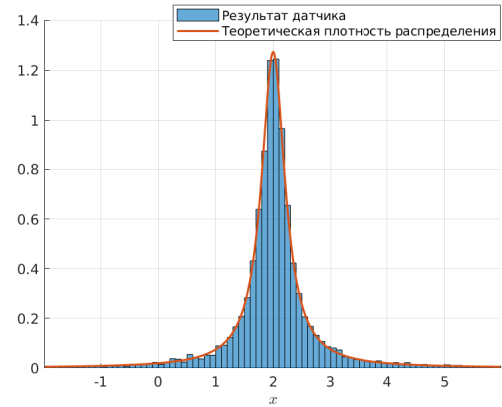
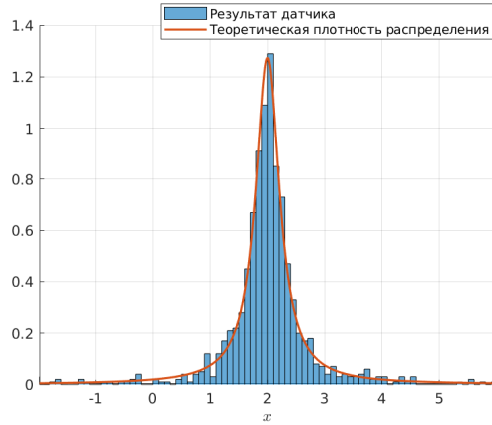


Рис. 4.1: Гистограмма распределения Коши случайной величины с параметрами  $a = 2$  и  $b = 1/4$  при  $10^3$  (слева) и  $10^4$  (справа) испытаний.

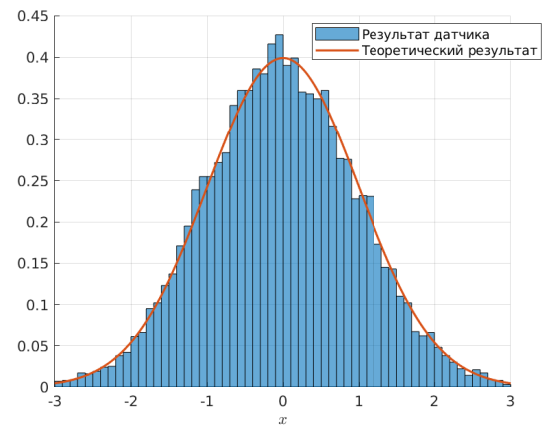
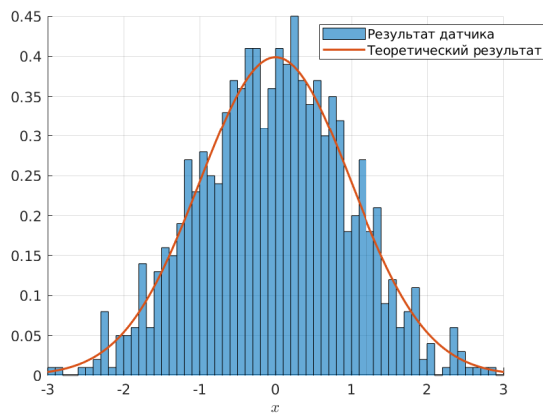


Рис. 4.2: Гистограмма стандартного нормального распределения случайной величины, смоделированной методом фон Неймана при  $10^3$  (слева) и  $10^4$  (справа) испытаний.

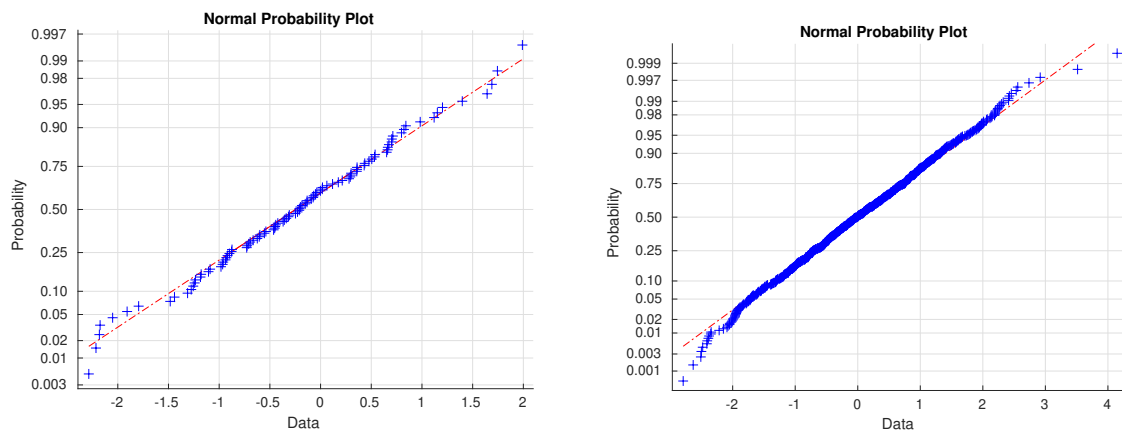


Рис. 4.3: Результат работы функции `normal probability plot` для датчика стандартного нормального распределения методом фон Неймана при  $10^2$  (слева) и  $10^3$  (справа) испытаний.

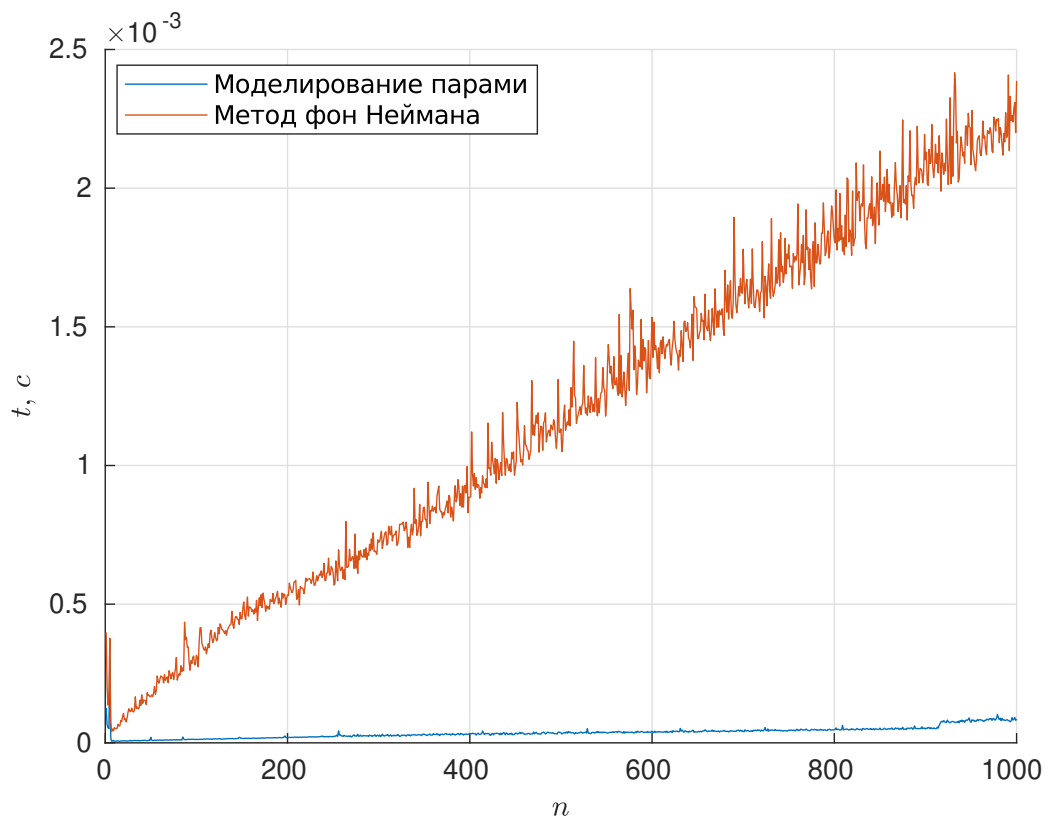


Рис. 4.4: Сравнение времени работы методов моделирования стандартного нормального распределения в зависимости от объема выборки.

## 5 Задание №5

1. Пусть  $X \sim N(\mu, \sigma^2)$ . Убедиться в справедливости ЗБЧ и ЦПТ, то есть исследовать поведение суммы  $S_n$  и эмпирического распределения случайной величины

$$\sqrt{n} \left( \frac{S_n}{n} - \mu \right).$$

2. Считая  $\mu$  и  $\sigma$  неизвестными, для пункта 1 построить доверительные интервалы для среднего и дисперсии.
3. Пусть  $X \sim C(a, b)$  имеет распределение Коши со сдвигом  $a$  и масштабом  $b$ . Проверить эмпирически, как ведут себя суммы  $\frac{S_n}{n}$ . Результат объяснить, а также найти закон распределения данных сумм.

### 5.1 Закон больших чисел. Центральная предельная теорема.

**Теорема 5.1** (Закон больших чисел). Пусть  $X_1, X_2, \dots, X_n, \dots$  — последовательность независимых одинаково распределенных случайных величин, определенных на одном вероятностном пространстве  $(\Omega, \mathcal{F}, \mathbb{P})$ , с конечным первым моментом, равным  $\mathbb{E} X_i = \mu$ . Обозначим за  $S_n = \sum_{i=1}^n X_i$ . Тогда

$$\frac{S_n}{n} \xrightarrow{\mathbb{P}} \mu,$$

то есть

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| \frac{S_n}{n} - \mu \right| < \varepsilon \right) = 1.$$

**Теорема 5.2** (Центральная предельная теорема). Пусть  $X_1, X_2, \dots, X_n, \dots$  — последовательность независимых одинаково распределенных случайных величин, определенных на одном вероятностном пространстве  $(\Omega, \mathcal{F}, \mathbb{P})$ , с конечным первым моментом, равным  $\mathbb{E} X_i = \mu$  и конечной дисперсией  $\text{Var} X_i = \sigma^2 \neq 0$ . Обозначим за  $Y_n = \frac{S_n - \mu n}{\sigma \sqrt{n}}$ , тогда

$$Y_n \xrightarrow{\text{dist.}} N(0, 1),$$

то есть

$$\lim_{n \rightarrow \infty} \mathbb{P}(Y_n < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = F_N(x).$$

### 5.2 Построение доверительных интервалов для среднего и дисперсии нормальной случайной величины.

Пусть теперь  $X_1, X_2, \dots, X_n$  — независимая выборка из некоторого нормального распределения  $N(\mu, \sigma^2)$  с неизвестными параметрами. Построим доверительный интервал для неизвестного среднего.

**Теорема 5.3.** Случайная величина

$$T = \sqrt{n} \cdot \frac{\bar{X} - \mu}{s},$$



где  $\bar{X}$  — выборочное среднее, а  $s$  — несмещенное выборочное стандартное отклонение, имеет распределение Стьюдента с  $(n - 1)$  степенью свободы.

Обозначим за  $t_\alpha$  квантиль распределения Стьюдента с  $(n - 1)$  степенью свободы порядка  $\alpha$ . В силу симметрии данного распределения имеем, что

$$\mathbb{P}(-t_{1-\alpha/2} \leq T \leq t_{1-\alpha/2}) = 1 - \alpha,$$

$\Downarrow$

$$\mathbb{P}\left(\bar{X} - \frac{s}{\sqrt{n}}t_{1-\alpha/2} \leq \mu \leq \bar{X} + \frac{s}{\sqrt{n}}t_{1-\alpha/2}\right) = 1 - \alpha.$$

Таким образом мы получили доверительный интервал для математического ожидания  $\mu$ :

$$\bar{X} - \frac{s}{\sqrt{n}}t_{1-\alpha/2} \leq \mu \leq \bar{X} + \frac{s}{\sqrt{n}}t_{1-\alpha/2},$$

длина которого равна

$$L(n) = 2 \frac{s}{\sqrt{n}} t_{1-\alpha/2} = O(n^{-1/2}) \xrightarrow{n \rightarrow \infty} 0.$$

Теперь будем строить доверительный интервал для дисперсии. В этом нам поможет теорема Фишера для нормальных выборок.

**Теорема 5.4** (Фишер). *Случайная величина*

$$H = \frac{s^2}{\sigma^2}(n - 1)$$

*имеет распределение хи-квадрат с  $(n - 1)$  степенью свободы.*

Обозначим за  $\chi_\alpha^2$  квантиль распределения хи-квадрат с  $(n - 1)$  степенью свободы порядка  $\alpha$ . Аналогично предыдущим рассуждениям имеем:

$$\mathbb{P}\left(\chi_{\frac{1+\alpha}{2}}^2 \leq H \leq \chi_{\frac{1-\alpha}{2}}^2\right) = \alpha,$$

$\Downarrow$

$$\mathbb{P}\left(\frac{s^2}{\chi_{\frac{1+\alpha}{2}}^2}(n - 1) \leq \sigma^2 \leq \frac{s^2}{\chi_{\frac{1-\alpha}{2}}^2}(n - 1)\right) = \alpha.$$

Таким образом мы получили доверительный интервал для дисперсии

$$\frac{s^2}{\chi_{\frac{1+\alpha}{2}}^2}(n - 1) \leq \sigma^2 \leq \frac{s^2}{\chi_{\frac{1-\alpha}{2}}^2}(n - 1),$$

длина которого равна

$$L(n) = s^2(n - 1) \left( \frac{s^2}{\chi_{\frac{1-\alpha}{2}}^2} - \frac{s^2}{\chi_{\frac{1+\alpha}{2}}^2} \right) \xrightarrow{n \rightarrow \infty} 0.$$

### 5.3 Поведение частичных сумм распределения Коши

Рассмотрим случайную величину  $X \sim C(a, b)$ . Эмпирически можно увидеть, что выборка таких случайных величин не удовлетворяет закону больших чисел. Разберемся почему это так.

**Определение 5.1.** Пусть  $\xi$  — произвольная случайная величина, определенная на вероятностном пространстве  $(\Omega, \mathcal{F}, \mathbb{P})$ . Тогда ее *математическим ожиданием* называется

$$\mathbb{E} \xi = \int_{\Omega} \xi(w) \mathbb{P}(dw).$$

Получается, что, так как значение

$$\mathbb{E} X = \frac{b}{\pi} \int_{-\infty}^{+\infty} \frac{1}{(x-a)^2 + b^2} dx$$

не определено, то не выполняется одно из условий закона больших чисел. Как же ведут себя частичные суммы распределения Коши? На этот вопрос можно ответить при помощи следующего утверждения.

**Утверждение 5.1** (Свойство устойчивости). Пусть  $X_1, X_2, \dots, X_n, \dots$  — независимые одинаково распределенные случайные величины, имеющие распределение Коши с параметрами  $a$  и  $b$ . Тогда случайная величина

$$S_n = \frac{1}{n} \sum_{k=1}^n X_k$$

также имеет распределение Коши с параметрами  $a$  и  $b$ , причем для любого  $n$ .

**Доказательство.** Воспользуемся аппаратом характеристических функций. Известно, что характеристическая функция случайной величины  $X \sim C(a, b)$  имеет вид

$$\varphi_X(x) = e^{aix - b|x|}.$$

Вспомним следующие свойства характеристических функций:

$$\varphi_{aX}(x) = \varphi_X(ax),$$

$$\varphi_{\sum_{k=1}^n X_k}(x) = \prod_{k=1}^n \varphi_{X_k}(x).$$

Теперь мы можем легко найти характеристическую функцию случайной величины  $S_n$ :

$$\varphi_{S_n}(x) = \varphi_{\sum_{k=1}^n X_k}(x/n) = \prod_{i=1}^n \varphi_{X_k}(x/n) = (\varphi_{X_1}(x/n))^n = \left( e^{\frac{aix}{n} - \left| \frac{x}{n} \right| b} \right)^n = e^{aix - b|x|} = \varphi_{X_1}(x).$$

Так как характеристические функции  $S_n$  и  $X_k$  совпадают, то их распределения тоже совпадают. Утверждение доказано. ■

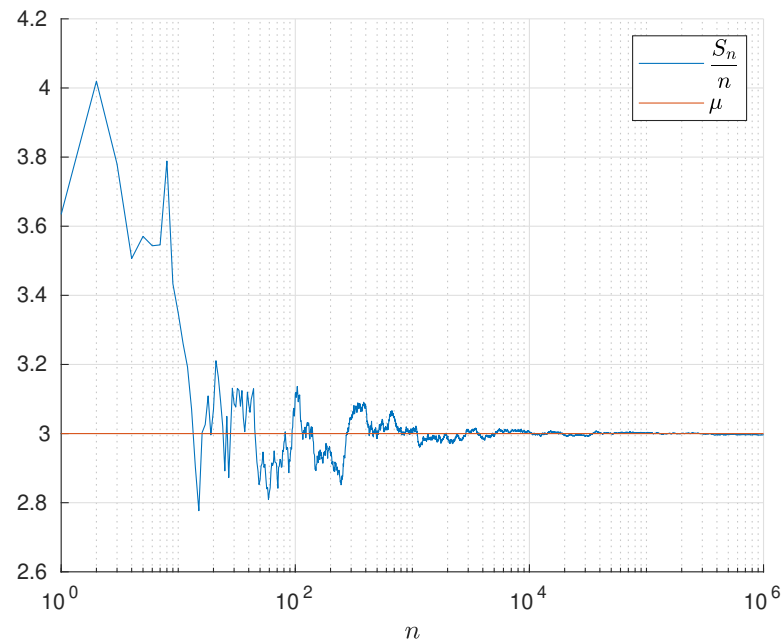


Рис. 5.1: Иллюстрация выполнения закона больших чисел для нормально распределенной случайной величины с параметрами  $\mu = 3$ ,  $\sigma^2 = 4$ .

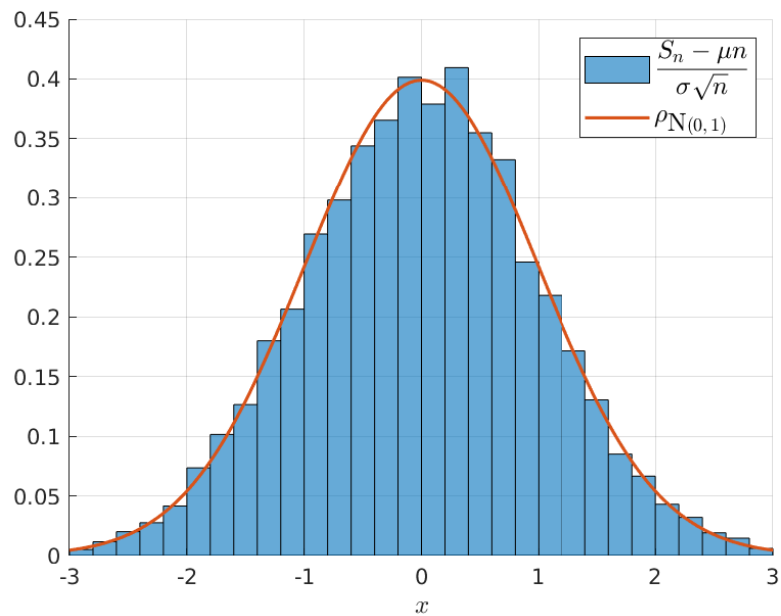


Рис. 5.2: Иллюстрация выполнения центральной предельной теоремы для нормально распределенной случайной величины с параметрами  $\mu = 3$ ,  $\sigma^2 = 4$ . Было проведено  $10^4$  испытаний по генерации  $S_n$ , где  $n = 10^4$ .

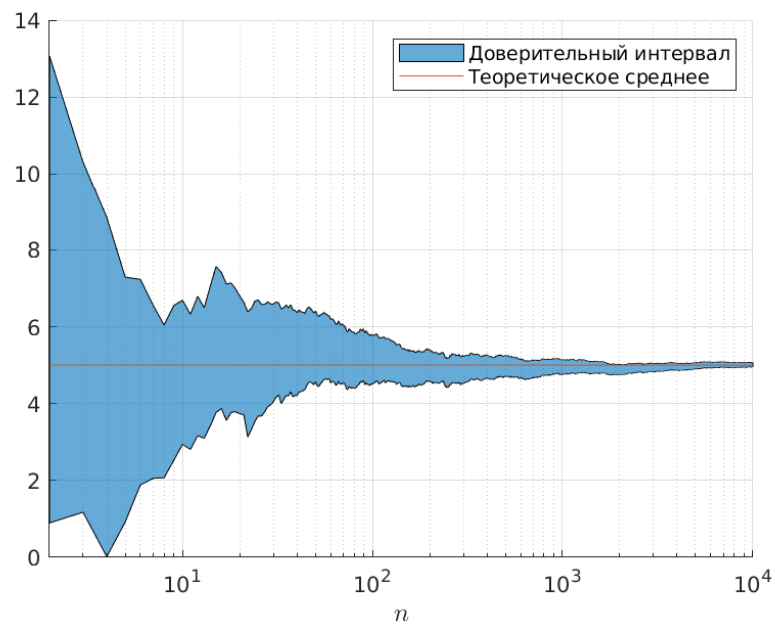


Рис. 5.3: Доверительный интервал для математического ожидания случайной величины нормального распределения с параметрами  $\mu = 5$ ,  $\sigma^2 = 9$ .

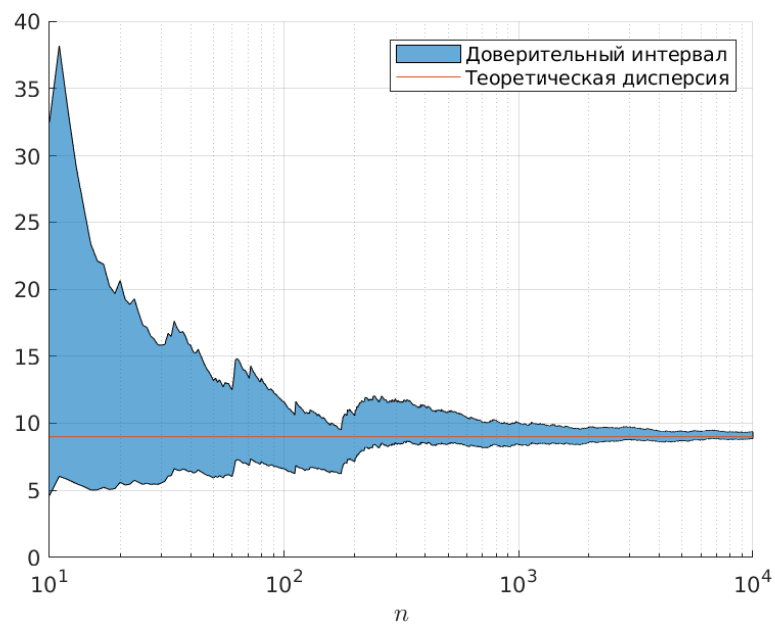


Рис. 5.4: Доверительный интервал для дисперсии случайной величины нормального распределения с параметрами  $\mu = 5$ ,  $\sigma^2 = 9$ .

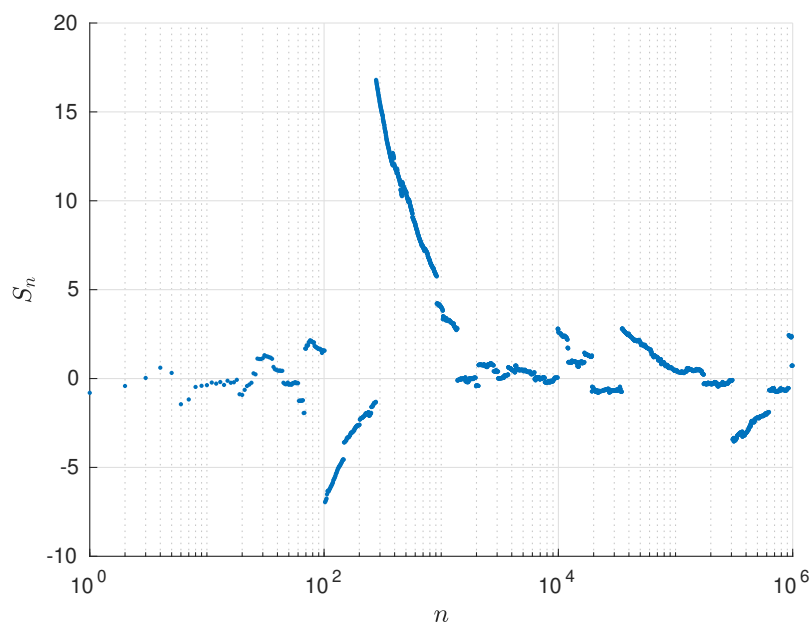


Рис. 5.5: Иллюстрация неприменимости закона больших чисел к распределению Коши. Поведение частичных сумм случайной величины Коши с параметрами  $a = 0$ ,  $b = 1, 5$ .

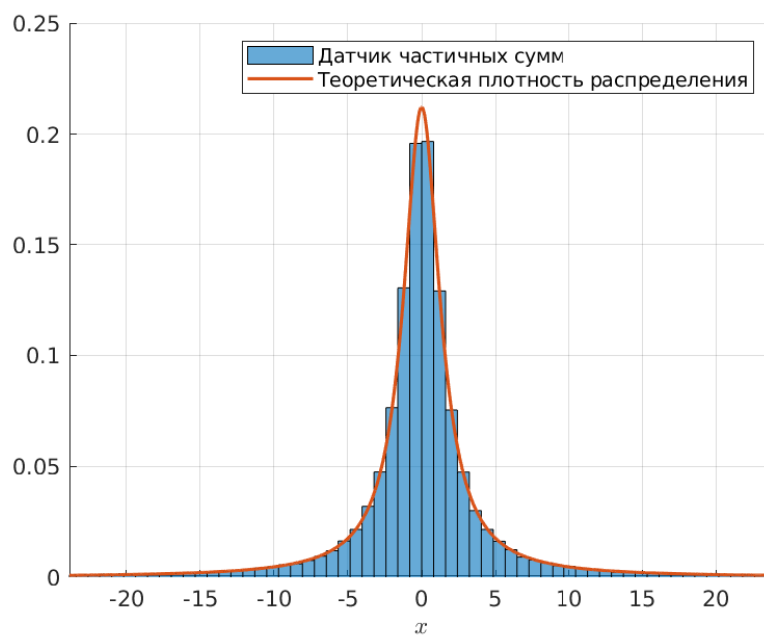


Рис. 5.6: Иллюстрация свойства устойчивости распределения Коши. Представлена гистограмма распределения частичных сумм  $S_{100}$  случайной величины Коши с параметрами  $a = 0$ ,  $b = 1, 5$ . Объем выборки —  $10^5$ .

## 6 Задание №6

1. Посчитать интеграл

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \frac{e^{-\left(x_1^2 + \dots + x_{10}^2 + \frac{1}{2^7 \cdot x_1^2 \cdot \dots \cdot x_{10}^2}\right)}}{x_1^2 \cdot \dots \cdot x_{10}^2} dx_1 dx_2 \dots dx_{10}$$

- методом Монте–Карло;
- методом квадратур, сводя задачу к вычислению собственного интеграла Римана.

2. Для каждого случая оценить точность вычислений.

### 6.1 Метод Монте–Карло

В данном разделе найдем методом Монте–Карло искомый интеграл и оценим точность вычислений. Для начала обозначим данный интеграл и подынтегральную функцию

$$I = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \underbrace{\frac{e^{-\left(x_1^2 + \dots + x_{10}^2 + \frac{1}{2^7 \cdot x_1^2 \cdot \dots \cdot x_{10}^2}\right)}}{x_1^2 \cdot \dots \cdot x_{10}^2}}_{f(x)} dx_1 dx_2 \dots dx_{10}.$$

Пусть  $\rho(x)$  — плотность некоторой случайной величины  $X$ , имеющей абсолютно непрерывное распределение. Обозначим за  $\hat{f}(x) = \frac{f(x)}{\rho(x)}$ . Тогда

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n} \frac{f(x)}{\rho(x)} \rho(x) dx = \int_{\mathbb{R}^n} \hat{f}(x) \rho(x) dx = \mathbb{E} \hat{f}(X).$$

Получается, что задача вычисления интеграла сводится к вычислению математического ожидания случайной величины  $Y = \hat{f}(X)$ . Согласно закону больших чисел,

$$Y_n = \frac{\sum_{i=1}^n \hat{f}(x^i)}{n} \xrightarrow{n \rightarrow \infty} \mathbb{E} \hat{f}(X), \quad \text{где } \{x^i\}_{i=1}^n \text{ — выборка случайной величины } X.$$

Поэтому в дальнейшем будем считать математическое ожидание, полагая

$$\mathbb{E} \hat{f}(X) \approx Y_n.$$

Для подсчета интеграла  $I$  возьмем случайную величину  $X$ , имеющую десятимерное стандартное нормальное распределение, то есть  $X \sim N(O, I)$ , где  $O, I \in \mathbb{R}^{10 \times 10}$ ,  $O$  — нулевая матрица,  $I$  — единичная матрица. Плотность такой случайной величины выражается формулой

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

Тогда

$$\hat{f}(x) = \frac{f(x)}{\rho(x)} = (2\pi)^5 \frac{e^{-\left(\frac{x_1^2 + \dots + x_{10}^2}{2} + \frac{1}{2^7 \cdot x_1^2 \cdot \dots \cdot x_{10}^2}\right)}}{x_1^2 \cdot \dots \cdot x_{10}^2}.$$

Таким образом, мы можем посчитать значение интеграла  $I$  по формуле

$$I \approx \frac{\sum_{i=1}^n \hat{f}(x^i)}{n}, \quad \text{где } x^i = (x_1^i, \dots, x_{10}^i), \quad x_k^i \sim N(0, 1), \quad k = \overline{0, 10}.$$

Теперь оценим точность вычисления интеграла методом Монте-Карло. Для этого воспользуемся центральной предельной теоремой:

$$\begin{aligned} \mathbb{P}(|Y_n - I| < \varepsilon) &= \mathbb{P}\left(\left|\frac{S_n - nI}{n}\right| < \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - nI}{\sqrt{n}\sigma_n}\right| < \frac{\varepsilon\sqrt{n}}{\sigma_n}\right) = \\ &= \mathbb{P}\left(\frac{S_n - nI}{\sqrt{n}\sigma_n} < \frac{\varepsilon\sqrt{n}}{\sigma_n}\right) - \mathbb{P}\left(\frac{S_n - nI}{\sqrt{n}\sigma_n} < -\frac{\varepsilon\sqrt{n}}{\sigma_n}\right) \approx F_N\left(\frac{\varepsilon\sqrt{n}}{\sigma_n}\right) - F_N\left(-\frac{\varepsilon\sqrt{n}}{\sigma_n}\right) = \\ &= 2F_N\left(\frac{\varepsilon\sqrt{n}}{\sigma_n}\right) - 1 = 1 - \alpha. \end{aligned}$$

$\Downarrow$

$$\varepsilon = K_{1-\alpha/2} \cdot \frac{\sigma_n}{\sqrt{n}},$$

где  $\sigma_n$  — несмещенная выборочная дисперсия,  $F_N$  — функция стандартного нормального распределения, а  $K_\alpha$  — его квантиль порядка  $\alpha$ . Таким образом для уровня значимости  $\alpha = 0,05$  погрешность будет равна  $\varepsilon = 1,96 \frac{\sigma_n}{\sqrt{n}}$ .

Число испытаний	Результат	Погрешность	Время работы, с
$10^5$	127,814	23,0369	0,016521
$10^6$	124,315	2,27042	0,997902
$10^7$	124,791	0,720674	9,60839
$10^8$	124,735	0,227662	104,135

Таблица 7: Результат подсчета интеграла  $I$  методом Монте-Карло.

## 6.2 Метод квадратур

В методе квадратур мы должны свести задачу к вычислению собственного интеграла Римана. Для этого сделаем замену переменных

$$x_i = \operatorname{tg}\left(\frac{\pi}{2}t_i\right), \quad 0 \leq t_i \leq 1, \quad i = \overline{1, 10}.$$

Таким образом, исходный интеграл равен интегралу с ограниченной областью интегрирования:

$$I = \left(\frac{\pi}{2}\right)^{10} \int_0^1 \dots \int_0^1 \frac{\exp\left\{-\left[\sum_{k=1}^{10} \operatorname{tg}\left(\frac{\pi}{2}t_k\right)^2 + \frac{1}{2^7 \cdot \prod_{k=1}^{10} \operatorname{tg}\left(\frac{\pi}{2}t_k\right)^2}\right]\right\}}{\prod_{k=1}^{10} \operatorname{tg}\left(\frac{\pi}{2}t_k\right)^2 \cdot \prod_{k=1}^{10} \cos\left(\frac{\pi}{2}t_k\right)^2} dt_1 \dots dt_{10}.$$

Теперь обозначим подынтегральную функцию последнего выражения за  $F(t) = F(t_1, \dots, t_{10})$  и рассмотрим равномерное разбиение отрезка  $[0, 1]$  на  $N$  частей:

$$0 = t^0 < t^1 < \dots < t^{10} = 1, \quad t^k = \frac{k}{N}, \quad k = \overline{0, N}$$

если взять за  $\xi_k = \frac{t^{k-1} + t^k}{2}$  середины отрезков  $[t^{k-1}, t^k]$ ,  $k = \overline{1, N}$ , то приближенно значение интеграла можно посчитать по формуле

$$I \approx \sum_{k_1=1}^N \dots \sum_{k_{10}=1}^N F(\xi_{k_1}, \dots, \xi_{k_{10}}) \Delta, \quad \text{где } \Delta = \frac{1}{N^{10}}.$$

Погрешность метода прямоугольников имеет следующий вид:

$$\varepsilon = \frac{h^2}{24}(b-a) \sum_{i,j=1}^n \max |f''_{x_i x_j}| = \frac{1}{6N^2} \sum_{i,j=1}^n \max |f''_{x_i x_j}|.$$

Число разбиений	Результат	Время работы, с
3	0,086797	0,007513
4	272,6029	0,073578
5	183,4886	0,889859
6	116,3903	8,377370

Таблица 8: Результат подсчета интеграла  $I$  методом квадратур. Видно, что метод Монте–Карло гораздо более оптимальный по времени.



## 7 Задание №7

1. Методом случайного поиска найти минимальное значение функции  $f$  на множестве  $A = \{x_1, x_2 : x_1^2 + x_2^2 \leq 1\}$ , т.е.  $y = \min_{x \in A} f(x)$ , где

$$f(x) = x_1^3 \sin\left(\frac{1}{x_1}\right) + 10x_1x_2^4 \cos\left(\frac{1}{x_2}\right)$$

при  $x_1 \neq 0$  и  $x_2 \neq 0$ , функция доопределяется по непрерывности при  $x_1 = 0$  или  $x_2 = 0$ .

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

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

3. Оценить точность. Сравнить результаты со стандартными методами оптимизации.

### 7.1 Метод случайного поиска

Алгоритм поиска заключается в следующем: мы будем генерировать случайные величины, равномерно распределенные на единичной окружности, а затем сравним значения функции в этих точках и найдем среди них наименьшее.

Для начала предложим способ генерации таких случайных величин. Воспользуемся методом перехода в полярные координаты. Совместная плотность распределения случайных величин  $x_1, x_2$  равна

$$\rho_{x_1, x_2}(x_1, x_2) = \begin{cases} \frac{1}{\pi}, & \text{при } x_1^2 + x_2^2 \leq 1, \\ 0, & \text{иначе.} \end{cases}$$

Тогда

$$\mathbb{P}((x_1, x_2)^T \in A) = \iint_{x_1^2 + x_2^2 \leq 1} \frac{1}{\pi} dx dy = \left\{ \begin{array}{l} x_1 = r \cos \varphi, \\ x_2 = r \sin \varphi, \\ 0 \leq r \leq 1, \\ 0 \leq \varphi \leq 2\pi \end{array} \right\} = \frac{1}{\pi} \int_0^1 r dr \int_0^{2\pi} d\varphi = \int_0^1 dr^2 \int_0^{2\pi} \frac{1}{2\pi} d\varphi.$$

В правой части последнего равенства получили произведение двух интегралов. Первый показывает, что функция распределения случайной величины  $r$  на отрезке  $[0, 1]$  равна  $F_r(x) = x^2$ ; второй — что плотность распределения случайной величины  $\varphi$  на отрезке  $[0, 2\pi]$  равна  $\rho_\varphi(x) = 1/2\pi$ . То есть

$$r = \sqrt{q}, \text{ где } q \sim U[0, 1], \quad \text{а } \varphi \sim U[0, 2\pi].$$

Теперь оценим точность предложенного алгоритма. Пусть  $(x_1, x_2)$  — точка минимума, полученная описанным выше методом, а  $(x_1^*, x_2^*)$  — реальная точка минимума. Тогда будем оценивать разность  $|f(x_1, x_2) - f(x_1^*, x_2^*)|$ . В этом нам поможет следующая теорема.

**Теорема 7.1.** Пусть функция  $f : \Omega \rightarrow \mathbb{R}$  непрерывно дифференцируема в выпуклой компактной области  $\Omega$  пространства  $\mathbb{R}^n$ . Тогда  $|f(x) - f(y)| \leq \sup_{\xi \in \Omega} |\text{grad } f(\xi)| \cdot |x - y|$ .

Получается, что мы можем сделать следующее:

$$|f(x_1, x_2) - f(x_1^*, x_2^*)| \leq \max_{\xi_1^2 + \xi_2^2 \leq 1} |\text{grad } f(\xi_1, \xi_2)| \cdot \underbrace{\text{dist}((x_1, x_2)^T, (x_1^*, x_2^*)^T)}_{\delta}.$$

$$\begin{aligned} \left| \frac{\partial f}{\partial x_1} \right| &= \left| 3x_1^2 \sin \frac{1}{x_1} - x_1 \cos \frac{1}{x_1} + 10x_2^4 \cos \frac{1}{x_2} \right| \leq \\ &\leq 3x_1^2 \left| \sin \frac{1}{x_1} \right| + |x_1| \left| \cos \frac{1}{x_1} \right| + 10x_2^4 \left| \cos \frac{1}{x_2} \right| \leq 3x_1^2 + |x_1| + 10x_2^4 = \\ &= 3x_1^2 + |x_1| + 10(1 - x_1^2)^2 \leq 10x_1^4 - 17x_1^2 + 11 \leq 11, \end{aligned}$$

$$\begin{aligned} \left| \frac{\partial f}{\partial x_2} \right| &= \left| 40x_1x_2^3 \cos \frac{1}{x_2} + 10x_1x_2^2 \sin \frac{1}{x_2} \right| \leq 40|x_1||x_2| \left| \cos \frac{1}{x_2} \right| + 10|x_1|x_2^2 \left| \sin \frac{1}{x_2} \right| = \\ &= 10|x_1|(1 - x_1^2) \left( 4\sqrt{1 - x_1^2} + 1 \right) \leq 50|x_1|(1 - x_1^2) \leq 19,245. \end{aligned}$$

↓

$$|f(x_1, x_2) - f(x_1^*, x_2^*)| \leq \sqrt{11^2 + 19,245^2} \cdot \delta \leq \underbrace{22,17}_C \delta.$$

Оценка получена, но она зависит от того, насколько близко построенная нами случайная величина окажется к реальной точке минимума. Оценим вероятность того, что хотя бы одна случайная величина из  $n$  окажется в  $\delta$ -окрестности искомой точки минимума. В худшем случае точка минимума может оказаться на границе множества  $A$  (в таком случае  $\delta$ -окрестность будет иметь минимальную площадь, а, значит, вероятность попадания туда равномерно распределенной случайной величины будет минимальной), поэтому оценивать будем вероятность попадания именно в граничную  $\delta$ -окрестность. Также учтем, что заданная функция  $f$  является четной по переменной  $x_2$ , что говорит о наличии как минимум двух точек минимума, поэтому вероятность попадания в окрестности минимумов увеличивается в 2 раза.

$$p \geq 1 - \left( 1 - \frac{\arccos(2 - \frac{\delta^2}{2}) - 2 \sin \left( \arccos(1 - \frac{\delta^2}{2}) \right) + \pi \delta^2}{\pi} \right)^n \approx 1 - (1 - \delta^2)^n.$$

Теперь, мы можем определить сколько потребуется генераций случайной величины, чтобы найти минимум функции с заданной погрешностью и уровнем доверия. В рамках же задания мы пойдем обратным путем: мы будем вычислять погрешность, исходя из уровня доверия и количества генераций:

$$\varepsilon = C \left( 1 - (1 - p)^{\frac{1}{n}} \right)^{\frac{1}{2}} \approx C \sqrt{\frac{p}{n}} = 22,17 \sqrt{\frac{p}{n}}.$$

Число генераций	Результат	Погрешность
$10^5$	-1,286	0,0698
$10^6$	-1,2879	0,0221
$10^7$	-1,2882	0,007
$10^8$	-1,2884	0,0022

Таблица 9: Результат работы метода случайного поиска при различном количестве генераций при уровне доверия  $p = 0,99$ .

## 7.2 Метод имитации отжига

Алгоритм основывается на имитации физического процесса, который происходит при кристаллизации вещества, в том числе при отжиге металлов. Предполагается, что атомы уже выстроились в кристаллическую решётку, но ещё допустимы переходы отдельных атомов из одной ячейки в другую. Предполагается, что процесс протекает при постепенно понижающейся температуре. Переход атома из одной ячейки в другую происходит с некоторой вероятностью, причём вероятность понижается с понижением температуры. Устойчивая кристаллическая решётка соответствует минимуму энергии атомов, поэтому атом либо переходит в состояние с меньшим уровнем энергии, либо остаётся на месте.

При помощи моделирования такого процесса ищется такая точка или множество точек, на котором достигается минимум числовой функции  $g(x)$ , где  $x = (x_1, \dots, x_m) \in X$ . Решение ищется последовательным вычислением точек  $x^0, x^1, \dots$  пространства  $X$ ; каждая точка, начиная с  $x^1$ , «претендует» на то, чтобы лучше предыдущих приближать решение. Алгоритм принимает точку  $x^0$  как исходные данные. На каждом шаге алгоритм (который описан ниже) вычисляет новую точку и понижает значение величины (изначально положительной), понимаемой как «температура». Алгоритм останавливается по достижении точки, которая оказывается при температуре ноль.

Точка  $x^{i+1}$  по алгоритму получается на основе текущей точки  $x^i$  следующим образом. К точке  $x_i$  применяется оператор  $A$ , который случайным образом модифицирует соответствующую точку, в результате чего получается новая точка  $x^*$ . Точка  $x^*$  становится точкой  $x^{i+1}$  с вероятностью  $p(x^*, x^{i+1})$ , которая вычисляется в соответствии с распределением Гиббса:

$$p(x^* \rightarrow x^{i+1} | x^i) = \begin{cases} 1, & \text{при } g(x^*) - g(x^i) < 0, \\ \exp\left(-\frac{g(x^*) - g(x^i)}{t_i}\right), & \text{при } g(x^*) - g(x^i) \geq 0. \end{cases}$$

Здесь  $t_i > 0$  — элементы произвольной убывающей, сходящейся к нулю положительной последовательности, которая задаёт аналог падающей температуры в кристалле. Скорость убывания и закон убывания могут быть заданы по желанию создателя алгоритма.

Алгоритм имитации отжига похож на градиентный спуск, но за счёт случайности выбора промежуточной точки должен попадать в локальные минимумы реже, чем градиентный спуск. Алгоритм имитации отжига не гарантирует нахождения минимума функции, однако при правильной политике генерации случайной точки в пространстве  $X$ , как правило, происходит улучшение начального приближения.

Число генераций	Результат
$10^2$	$6,1673 \cdot 10^{-6}$
$10^3$	$1,2550 \cdot 10^{-7}$
$10^4$	$2,8228 \cdot 10^{-8}$

Таблица 10: Результат работы метода имитации отжига.

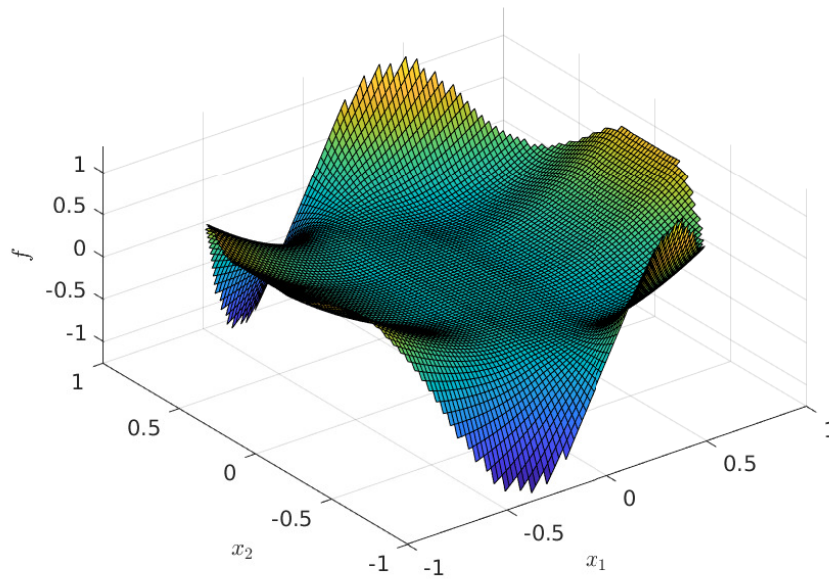


Рис. 7.1: Внешний вид функции  $f(x)$  на множестве  $x_1^2 + x_2^2 \leq 1$ . Видно, где достигается минимум и его примерное значение, поэтому предложенное решение не единственное. Действительно, из-за того, что минимум достигается на границе, мы могли сузить область случайного поиска, например, на граничную окружность.

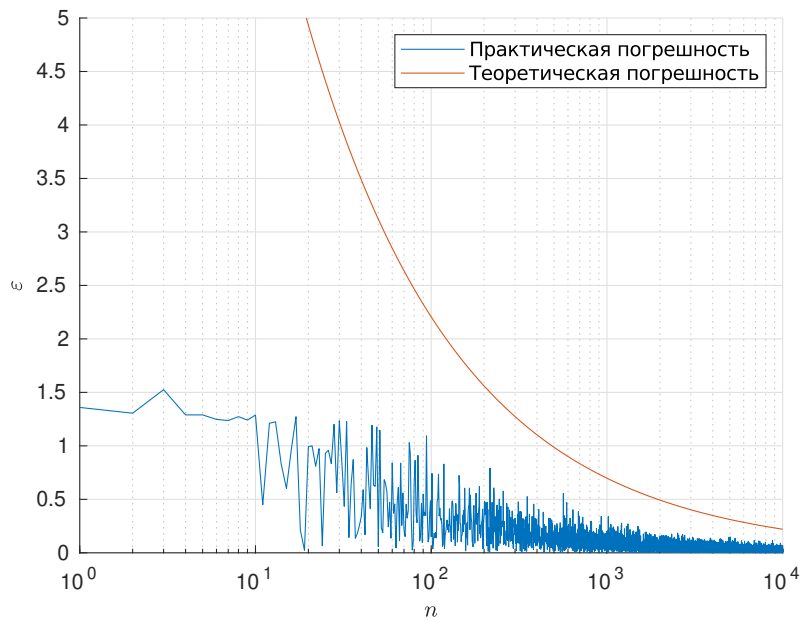


Рис. 7.2: Скорость сходимости метода случайного поиска и его верхняя оценка с уровнем доверия  $p = 0,99$  для заданной функции  $f(x)$ . Для поиска практической погрешности было найдено реальное значение минимума, равное  $-1,2885$ , при помощи метода градиентного спуска.

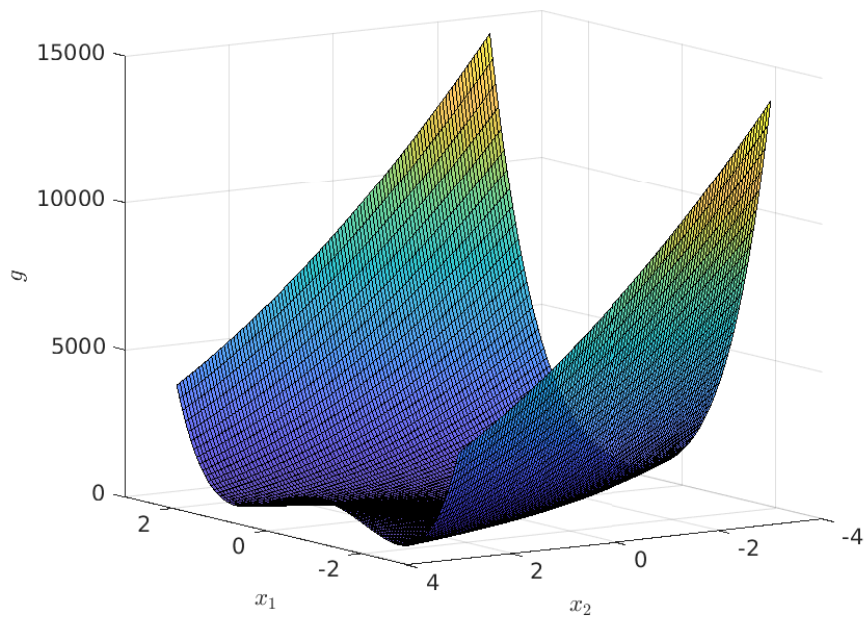


Рис. 7.3: Внешний вид функции  $g(x_1, x_2)$  на квадрате со стороной 3.

## 8 Задание №8

Применить метод Монте–Карло к решению первой краевой задачи для двумерного уравнения Лапласа в единичном круге:

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

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

### 8.1 Алгоритм решения задачи

Построим разностную схему для данной задачи Дирихле. Для этого выберем достаточно мелкую квадратную сетку с шагом  $h$ . Координаты узлов пусть будут  $x_i = ih$ ,  $y_j = jh$ , а значения  $u(x_i, y_j)$  и  $f(x_i, y_j)$  для краткости обозначим за  $u_{i,j}$  и  $f_{i,j}$ .

**Определение 8.1.** Будем называть узел сетки  $(i, j)$  *внутренним*, если он и все четыре соседних с ним узла:  $(i-1, j)$ ,  $(i+1, j)$ ,  $(i, j-1)$ ,  $(i, j+1)$  принадлежат области  $D + \delta D$ ; в противном случае узел  $(i, j)$ , принадлежащий  $D + \delta D$ , будем называть *граничным*.

Обозначим за  $D_h$  множество всех внутренних точек, а за  $\delta D_h$  — множество всех граничных точек.

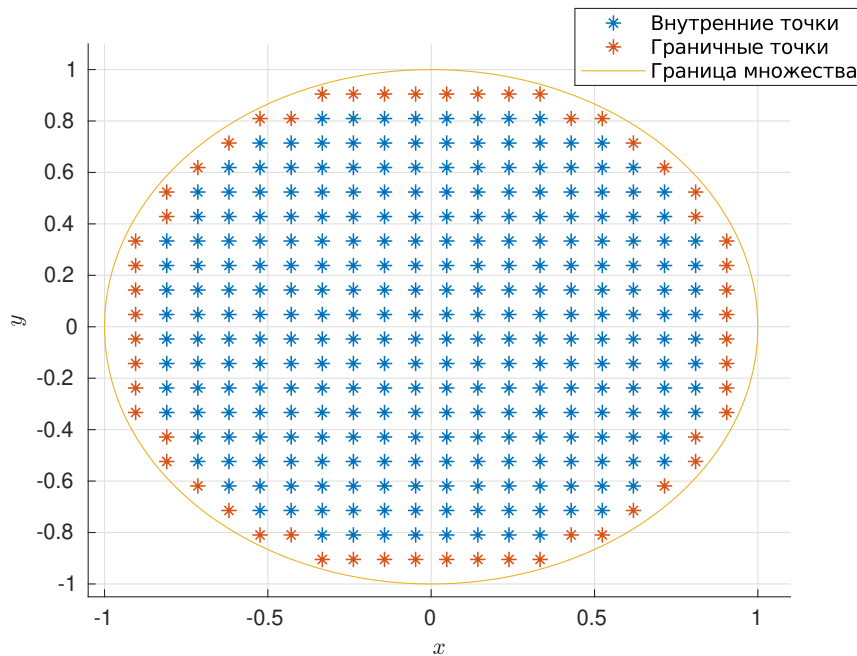


Рис. 8.1: Иллюстрация определения внутренних и граничных точек.

Во внутреннем узле  $(x_i, y_j)$  уравнение Лапласа  $u_{xx} + u_{yy} = 0$  заменим разностным уравнением

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h^2} = 0,$$

которое можно переписать в виде

$$u_{i,j} = \frac{1}{4}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}).$$

В граничном же узле положим

$$u_{i,j} = f_{i,j}.$$

Представим себе частицу, которая совершает равномерное случайное блуждание по узлам сетки. А именно, находясь во внутреннем узле  $(x_i, y_j)$  сетки, эта частица за один переход с одинаковой вероятностью  $1/4$  может переместиться в один из четырёх соседних узлов, причём каждый такой единичный переход случаен и не зависит от положения частицы и истории её передвижений. Будем считать, что блуждание заканчивается, как только частица попадает в граничный узел.

Пусть  $P(i, j, p, q)$  — вероятность того, что траектория частицы, вышедшей из узла  $(x_i, y_j)$ , закончится в граничном узле  $(x_p, y_q)$ . Так как блуждание точки неизбежно заканчивается на границе в первой же точке выхода её на границу, то

$$\sum_{(x_p, y_q) \in \delta D_h} P(i, j, p, q) = 1,$$

причём если  $(p', q'), (p, q) \in \delta D_h$ , то

$$P(p', q', p, q) = \begin{cases} 1, & \text{при } (p' - p)^2 + (q' - q)^2 = 0, \\ 0, & \text{при } (p' - p)^2 + (q' - q)^2 \neq 0. \end{cases}$$

Теперь составим сумму

$$v_{i,j} = \sum_{(x_p, y_q) \in \delta D_h} P(i, j, p, q) f_{pq}.$$

Если рассматривать функцию  $f(x, y)$  как случайную величину, принимающую значения  $f_{pq}$  на границе  $\delta D_h$ , то написанная выше сумма представляет собой математическое ожидание функции  $f(x, y)$  на границе  $\delta D_h$  для траекторий, начинающихся в узле  $(x_i, y_j)$ . Тогда в силу закона больших чисел можно аппроксимировать математическое ожидание выборочным средним:

$$v_{i,j} \approx \frac{1}{N} \sum_{k=1}^N f(x_p^{(k)}, y_q^{(k)}).$$

Частица, начавшая своё случайное блуждание из внутреннего узла  $(x_i, y_j)$ , после первого шага с вероятностью, равной  $1/4$ , попадает в один из соседних четырёх узлов. Откуда по формуле полной вероятности

$$\begin{aligned} v_{i,j} &= \frac{1}{4} \sum_{(x_p, y_q) \in \delta D_h} (P(i-1, j, p, q) + P(i+1, j, p, q) + P(i, j-1, p, q) + P(i, j+1, p, q)) f_{pq} = \\ &= \frac{1}{4} (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}). \end{aligned}$$

То есть во внутреннем узле  $(x_i, y_j)$

$$v_{i,j} = \frac{1}{4}(v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}),$$

а в граничном узле

$$v_{i,j} = f_{i,j}.$$

Теперь мы можем описать алгоритм построения численного решения задачи:

1. Построим квадратную сетку с шагом  $h$  на заданном множестве  $D$ , как это сделано, например, на рисунке 8.1.
2. В каждом граничном узле этой сетки положим

$$u(x, y) = f(x, y).$$

3. В каждом же внутреннем узле  $(x_i, y_j)$  сетки проведем  $n$  случайных блужданий. Тогда значение искомой функции в этом узле можно положить

$$u(x_i, y_j) = \frac{1}{n} \sum_{k=1}^n f(x_{k_i}, y_{k_j}),$$

где  $(x_{k_i}, y_{k_j})$  — граничный узел, в котором завершилось  $k$ -ое случайное блуждание, “выпущенное” из узла  $(x_i, y_j)$ .

Проверим правильность работы приведенного алгоритма для конкретной функции  $f(x, y) = x^2 - y^2$ . Для того, чтобы нам было с чем сравнивать, найдем аналитическое решение этой задачи. Сразу оговоримся, что по теореме о существовании решения внутренней задачи Дирихле, оно точно существует. Будем искать решение в виде  $u(x, y) = Ax^2 + By^2 + C$ . Подставив его в формулировку задачи, получим следующие условия на коэффициенты:

$$\begin{cases} A + B = 0, \\ A - B = 2, \\ B + C = -1; \end{cases} \iff \begin{cases} A = 1, \\ B = -1, \\ C = 0; \end{cases}$$

То есть мы получили, что функция  $u(x, y) = x^2 - y^2$  является решением задачи, и причём это решение единственно.



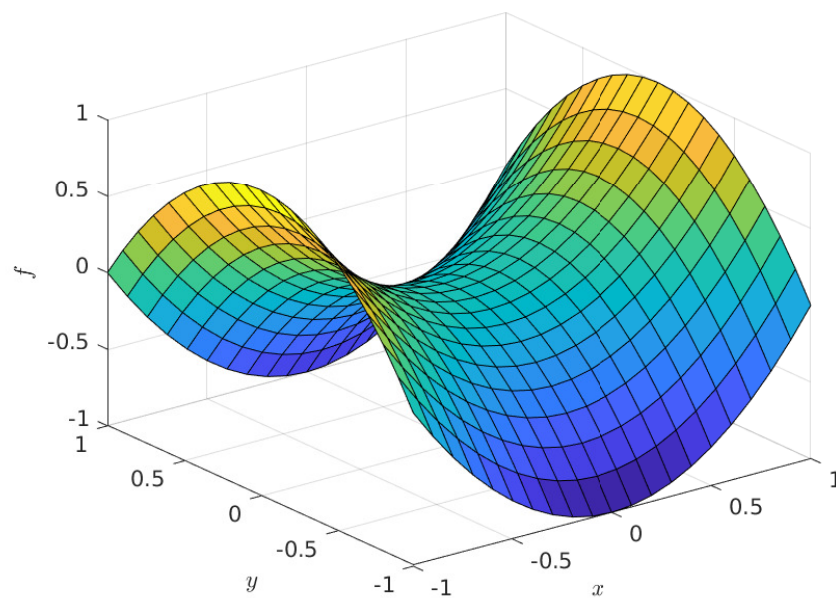


Рис. 8.2: Аналитическое решение задачи, построенное на множестве  $[-1, 1] \times [-1, 1]$ .

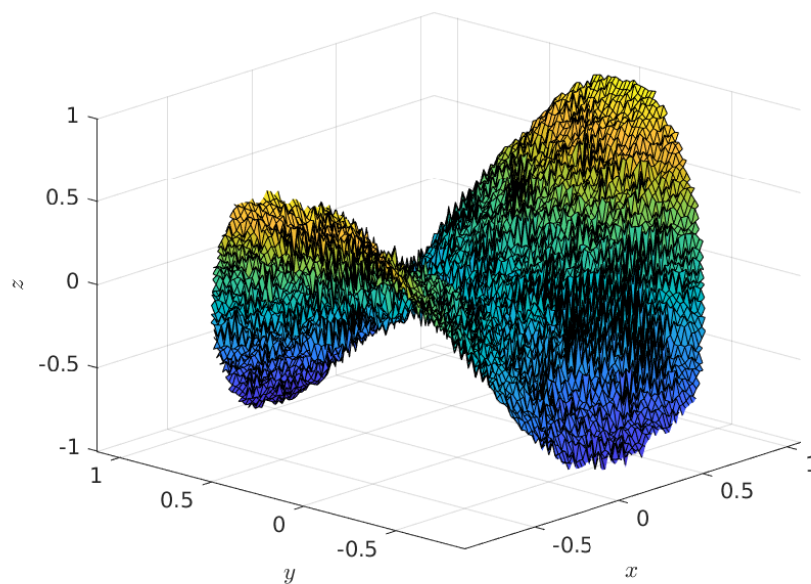


Рис. 8.3: Решение задачи методом Монте–Карло на сетке с радиусом 0,02.

## 9 Задание №9

Рассмотреть два вида процессов:

- Винеровский процесс  $W(t)$ ,  $t \in [0, 1]$ ,  $W(0) = 0$ .
- Процесс Орнштейна–Уленбека  $X(t)$ ,  $t \in [0, 1]$ ,  $X(0) = X_0$ , то есть стационарный гауссовский процесс. Начальные значения  $X_0$  генерируются случайным образом так, чтобы полученный процесс был стационарным.

Для данных гауссовских процессов:

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

### 9.1 Винеровский процесс

**Определение 9.1.** Рассмотрим вероятностное пространство  $(\Omega, \mathcal{F}, \mathbb{P})$ . Тогда назовём *случайным процессом* параметризованное семейство  $\{W_t\}_{t \in T}$  случайных величин

$$W_t : \Omega \rightarrow \mathbb{R}, t \in T, T \subset [0, +\infty).$$

В данном случае множество параметров интерпретируется как некоторый временной интервал.

**Определение 9.2.** Будем называть случайный процесс  $\{W_t\}_{t \in T}$  *гауссовским*, если для любых  $t_0, t_1, \dots, t_n \in T$  соответствующий случайный вектор  $w = (W_{t_1}, W_{t_2}, \dots, W_{t_n})$  имеет многомерное нормальное распределение, то есть имеет плотность

$$\rho(W_{t_1}, \dots, W_{t_n}) = \frac{1}{(2\pi)^{n/2} |R|^{1/2}} \cdot \exp \left\{ -\frac{1}{2} \cdot \langle R^{-1}(w - m), w - m \rangle \right\},$$

где  $m = (m_1, m_2, \dots, m_n)^T$  — вектор средних, а  $R \in \mathbb{R}^{n \times n}$  — ковариационная матрица:  $R = \|\text{cov}(t_i, t_j)\|_{i,j}$ ,  $R = R^T > 0$ .

**Определение 9.3.** Определим *винеровский процесс* как гауссовский процесс на отрезке  $[0, 1]$  с нулевым средним и ковариационной функцией  $\text{cov}(W(t_i), W(t_j)) = \min\{t_i, t_j\}$ .

Выпишем основные свойства этого процесса. Доказательство этих свойств можно найти у Ширяева. Итак:

1.  $W(0) = 0$  почти наверное;
2.  $W(t)$  является непрерывной функцией (по переменной  $t$ );
3. Приращения функции  $W(t)$  независимы и имеют стандартное нормальное распределение, то есть  $W(t_2) - W(t_1) \sim N(0, 1)$ , для любых  $t_1 < t_2$ .

### 9.1.1 Переходные вероятности

Рассмотрим отрезок  $[t_1, t_2] \subset T$  и его внутреннюю точку  $t = t_1 + \alpha(t_2 - t_1)$ ,  $0 < \alpha < 1$ . И найдем условную плотность

$$\rho_{W(t)}(x \mid W(t_1) = x_1, W(t_2) = x_2) = \frac{\rho_{W(t_1), W(t), W(t_2)}(x_1, x, x_2)}{\rho_{W(t_1), W(t_2)}(x_1, x_2)}.$$

Обозначим векторы  $\hat{x} = [x_1, x_2]^T$  и  $\hat{\hat{x}} = [x_1, x, x_2]^T$ . Тогда плотности вероятностей равны

$$\rho_{W(t_1), W(t_2)}(\hat{x}) = \frac{1}{2\pi\sqrt{|R_2|}} \cdot \exp \left\{ -\frac{1}{2} \hat{x}^T R_2^{-1} \hat{x} \right\},$$

$$\rho_{W(t_1), W(t), W(t_2)}(\hat{\hat{x}}) = \frac{1}{(2\pi)^{3/2} \sqrt{|R_3|}} \cdot \exp \left\{ -\frac{1}{2} \hat{\hat{x}}^T R_3^{-1} \hat{\hat{x}} \right\}.$$

Здесь за  $R_2$  и  $R_3$  обозначены соответствующие ковариационные матрицы:

$$R_2 = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix}, \quad R_3 = \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t & t \\ t_1 & t & t_2 \end{pmatrix}.$$

Теперь посчитаем определители и обратные ковариационных матриц:

$$|R_2| = t_1(t_2 - t_1), \quad |R_3| = t_1(t - t_1)(t_2 - t).$$

$$R_2^{-1} = \begin{pmatrix} \frac{t_2}{t_1(t_2 - t_1)} & -\frac{1}{t_2 - t_1} \\ -\frac{1}{t_2 - t_1} & \frac{1}{t_2 - t_1} \end{pmatrix}$$

$$R_3^{-1} = \begin{pmatrix} \frac{t}{t_1(t - t_1)} & -\frac{1}{t - t_1} & 0 \\ -\frac{1}{t - t_1} & \frac{t_2 - t_1}{(t_2 - t)(t - t_1)} & -\frac{1}{t_2 - t} \\ 0 & -\frac{1}{t_2 - t} & \frac{1}{t_2 - t} \end{pmatrix}$$

Соединим все вместе и получим

$$\rho_{W(t)}(x \mid W(t_1) = x_1, W(t_2) = x_2) = \frac{1}{\sqrt{2\pi(1 - \alpha)(t_2 - t_1)}} \exp \left\{ -\frac{(x - (1 - \alpha)x_1 + \alpha x_2)^2}{2\alpha(1 - \alpha)(t_2 - t_1)} \right\}.$$

### 9.1.2 Алгоритм

1. Сначала определим значения винеровского процесса на концах отрезка  $[0, 1]$ . По условию  $t_0 = 0$ ,  $W(t_0) = 0$ ,  $t_1 = 1$ . Величину  $W(t_1)$  разыграем как  $N(0, 1)$ .
2. Будем рекурсивно делить отрезки  $[t_0, t_1]$  в некотором отношении  $\alpha$ . Затем считать значение случайной величины  $W(t)$  по выведенной формуле условной плотности. Это просто, если заметить тот факт, что:

$$W(t) \sim N((1 - \alpha)x_1 + \alpha x_2, \alpha(1 - \alpha)(t_2 - t_1)).$$

3. Остановим алгоритм по достижении заданной точности  $t_{k+1} - t_k < \varepsilon$ .

*Замечание 9.1.* Проверить правильность вычислений можно дополнительно построив доверительный интервал. В нашем случае можно воспользоваться правилом трех сигм.

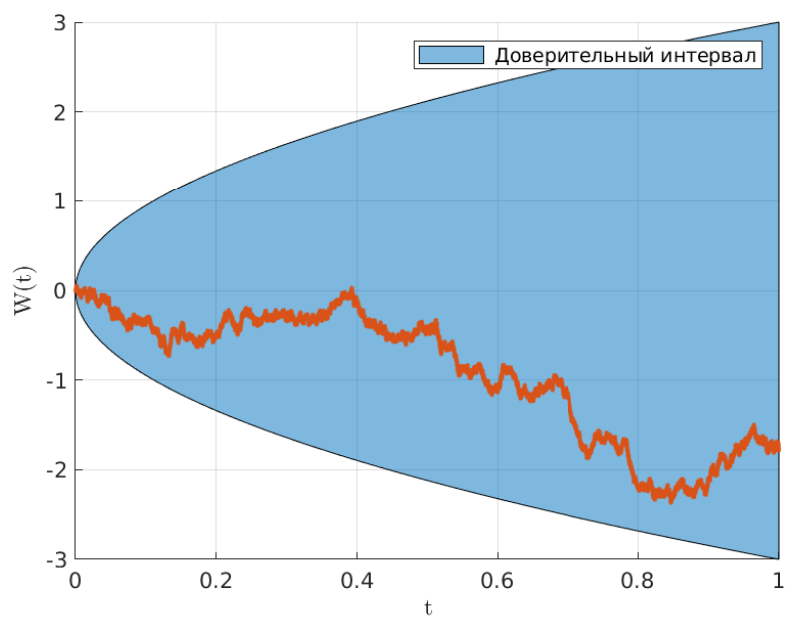


Рис. 9.1: Поточечный график винеровского процесса демонстрирует его непрерывность. Параметрами алгоритма были взяты  $\alpha = 0,3$ ,  $\varepsilon = 10^{-4}$ .

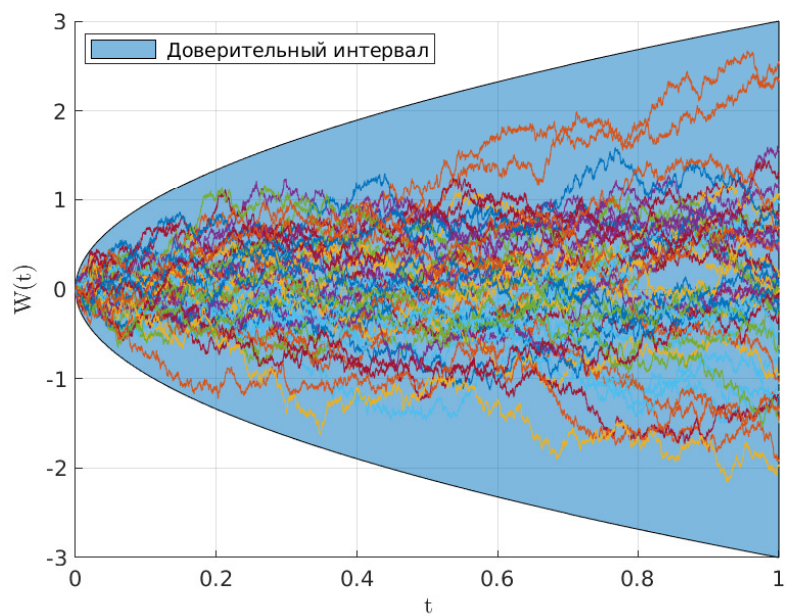


Рис. 9.2: Представлены 50 винеровских процессов.

## 9.2 Процесс Орнштейна–Уленбека

**Определение 9.4.** Случайный процесс  $\{W_t\}_{t \in T}$  называется *стационарным*, если конечномерные распределения инвариантны относительно сдвига времени.

**Определение 9.5.** Гауссовский процесс  $\{W_t\}_{t \in T}$  называется *процессом Орнштейна–Уленбека*, если он является стационарным и марковским.

Из-за того, что процесс Орнштейна–Уленбека является стационарным, он обладает следующими свойствами:

$$\mathbb{E} W(t) = \mu = \text{const}, \quad R(t, s) = R(|s - t|).$$

Не ограничивая общностей, будем рассматривать процесс при  $\mu = 0$ . Введем обозначение  $\text{Var } W(t) = \sigma^2$ . Тогда

$$R(t, s) = \sigma^2 \rho(s, t), \quad \text{где } \rho(s, t) \text{ коэффициент корреляции.}$$

**Теорема 9.1.** Для того, чтобы последовательность  $W_1, \dots, W_n$  нормально распределенных случайных величин была марковской, необходимо и достаточно, чтобы

$$\rho(j, k) = \rho(j, i) \cdot \rho(i, k), \quad \forall i, j, k : j \leq i < k \leq n,$$

где  $\rho(i, j)$  — коэффициент корреляции случайных величин  $W_i$  и  $W_j$ .

В силу того, что процесс  $W(t)$  является марковским, то  $\rho(s, t) = \rho(s, \tau) \cdot \rho(\tau, t)$ . В силу же того, что  $R(s, t) = R(|s - t|)$ , то  $\rho(s, t) = \rho(s - t)$ . Тогда введем замену переменных:  $x = s - \tau$ ,  $y = \tau - t$ . Получим, что

$$\rho(x + y) = \rho(x) \cdot \rho(y).$$

**Теорема 9.2.** Пусть функция  $f(t)$  определена при  $t > 0$  и ограничена на каждом конечном интервале. Если  $f(t)$  удовлетворяет соотношению  $f(t + s) = f(t)f(s)$ , то или  $f(t) \equiv 0$ , или  $f(t) = e^{-\lambda t}$ , где  $\lambda$  — некоторая положительная константа.

Получается, что задача распадается на два варианта: когда  $\rho(t) \equiv 0$  и иначе. Первый случай равносильен тому, что ковариационная функция  $\text{cov}(W(t), W(s))$  также равна нулю. Это значит, что случайные величины  $W(t)$  независимы в совокупности. Поэтому будем моделировать каждую случайную величину как  $N(\mu, \sigma^2)$ .

Теперь рассмотрим случай  $\rho(s, t) = e^{-\lambda|s-t|}$ ,  $\lambda > 0$ . Тогда ковариационная функция рассматриваемого процесса имеет вид  $R(s, t) = \sigma^2 e^{-\lambda|s-t|}$ . Найдем переходную плотность:

$$\rho_{W(t)}(x_1 | W(s) = x_2) = \frac{\rho_{W(t), W(s)}(x_1, x_2)}{\rho_{W(s)}(x_2)}.$$

Мы можем выписать формулу для каждой из плотностей, участвующих в выражении, так как рассматривается гауссовский процесс:

$$\rho_{W(t), W(s)}(x_1, x_2) = \frac{1}{2\pi|C|^{1/2}} \exp \left\{ -\frac{1}{2} \langle x, C^{-1}x \rangle \right\}, \quad \text{где } x = (x_1, x_2)^T.$$

$$\rho_{W(s)}(x_2) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{x_2^2}{2\sigma^2} \right\}.$$

Здесь  $C$  — ковариационная матрица. Запишем ее, ее определитель и обратную:

$$C = \begin{pmatrix} \sigma^2 & R(t, s) \\ R(t, s) & \sigma^2 \end{pmatrix}, \quad |C| = \sigma^4 - R^2(t, s), \quad C^{-1} = \frac{1}{|C|} \begin{pmatrix} \sigma^2 & -R(t, s) \\ -R(t, s) & \sigma^2 \end{pmatrix}.$$

Соберем все вместе:

$$\rho_{W(t)}(x_1 | W(s) = x_2) = \frac{1}{\left(2\pi \left(\sigma^2 - \frac{R^2(t, s)}{\sigma^2}\right)\right)^{1/2}} \exp \left\{ -\frac{\left(x_1 - \frac{R(t, s)}{\sigma^2} x_2\right)^2}{2 \left(\sigma^2 - \frac{R^2(t, s)}{\sigma^2}\right)} \right\}.$$

### 9.2.1 Алгоритм

Процесс Орнштейна–Уленбека марковский. Это говорит нам о том, что имея значения случайных величин  $W(t_1)$  и  $W(t_2)$ , мы всегда можем сгенерировать случайную величину  $W(t)$ , где  $t_1 \leq t \leq t_2$ . Дальнейшие рассуждения проведем аналогично рассмотренному выше винеровскому процессу. То есть зафиксируем точку  $t = \frac{t_1+t_2}{2}$  и найдем условную плотность

$$\rho_{W(t)}(W(t_1) = x_1, W(t_2) = x_2) = \frac{\rho_{W(t_1), W(t), W(t_2)}(x_1, x, x_2)}{\rho_{W(t_1), W(t_2)}(x_1, x_2)}.$$

Так как процесс  $W(t)$  гауссовский, то мы знаем эти плотности:

$$\rho_{W(t_1), W(t_2)} = \frac{1}{2\pi \sqrt{|R_2|}} \exp \left\{ -\frac{1}{2} \hat{x}^T R_2^{-1} \hat{x} \right\},$$

$$\rho_{W(t_1), W(t), W(t_2)}(x_1, x, x_2) = \frac{1}{(2\pi)^{3/2} \sqrt{|R_3|}} \exp \left\{ -\frac{1}{2} \hat{\hat{x}}^T R_1^{-1} \hat{\hat{x}} \right\}.$$

Здесь  $\hat{x} = [x_1, x_2]$ ,  $\hat{\hat{x}} = [x_1, x, x_2]$ ,

$$R_2 = \sigma^2 \begin{pmatrix} 1 & e^{-\lambda(t_2-t_1)} \\ e^{-\lambda(t_2-t_1)} & 1 \end{pmatrix}, \quad R_3 = \sigma^2 \begin{pmatrix} 1 & e^{-\lambda(t-t_1)} & e^{-\lambda(t_2-t_1)} \\ e^{-\lambda(t-t_1)} & 1 & e^{-\lambda(-\lambda)(t_2-t)} \\ e^{-\lambda(t_2-t_1)} & e^{-\lambda(t_2-t_1)} & 1 \end{pmatrix}.$$

Если подставить результаты в формулу условной плотности, то получим следующий результат:

$$W(t) \sim N \left( (x_1 + x_2) \frac{e^{-\frac{\lambda(t_2-t_1)}{2}}}{1 + e^{-\lambda(t_2-t_1)}}, \sigma^2 \frac{1 - e^{-\lambda(t_2-t_1)}}{1 + e^{-\lambda(t_2-t_1)}} \right).$$

В качестве граничных значений возьмем:

$$W(0) \sim N(0, \sigma^2), \quad W(1) \sim N(x_0 e^{-\lambda T}, \sigma^2 (1 - e^{-2\lambda T})).$$

*Замечание 9.2.* Также построим доверительный интервал правилом трех сигм.

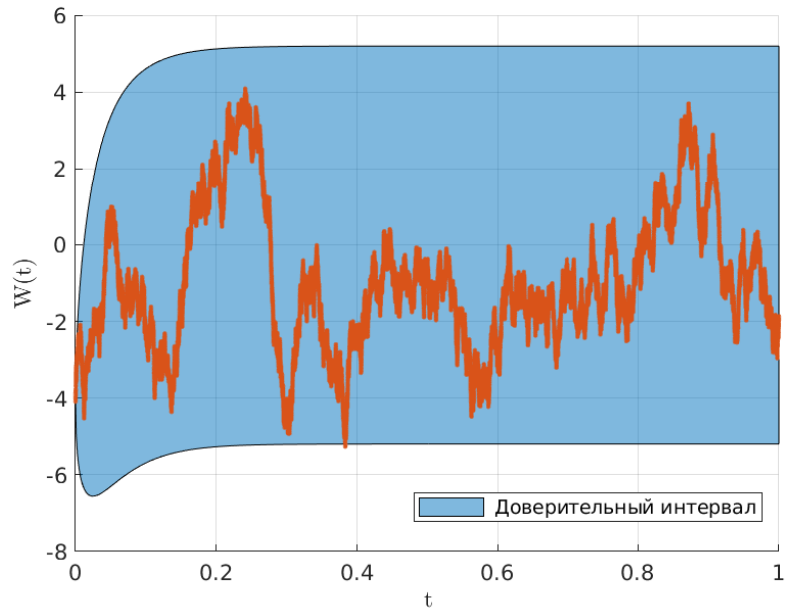


Рис. 9.3: Поточечный график процесса Орштейна–Уленбека с параметрами  $\lambda = 20, \sigma^2 = 3$  демонстрирует его непрерывность. Погрешность алгоритма была взята  $\varepsilon = 10^{-5}$ .

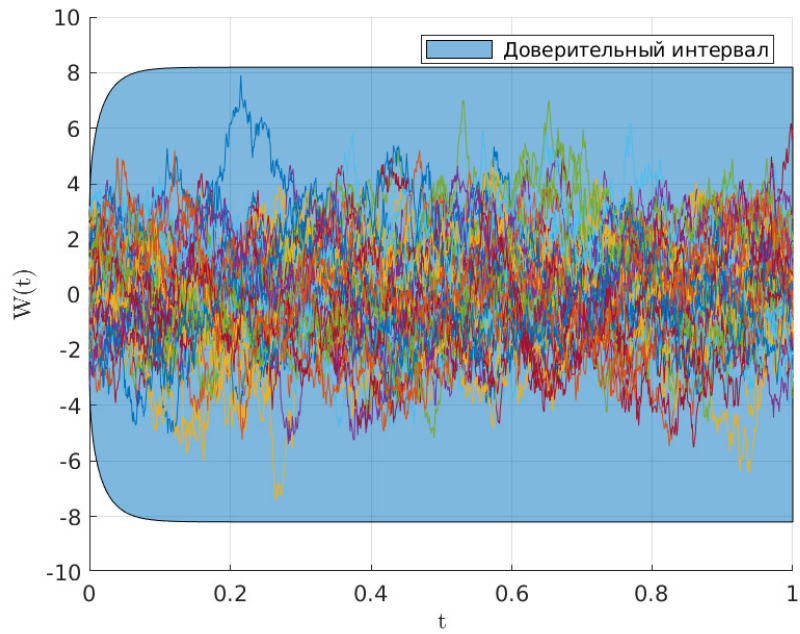


Рис. 9.4: Представлены 50 процессов Орштейна–Уленбека с параметрами  $\lambda = 20, \sigma^2 = 3$ .

## 10 Задание №10

Провести фильтрацию одномерного процесса Орнштейна–Уленбека:

1. Используя генератор белого шума, добавить случайную ошибку с известной дисперсией к реализации процесса Орнштейна–Уленбека.
2. При помощи фильтра Калмана оценить траекторию процесса по зашумленному сигналу. Параметры процесса и белого шума считать известными.
3. Рассмотреть случай, когда шум
  - является гауссовским,
  - имеет распределение Коши.

### 10.1 Фильтр Калмана для гауссовского шума

Рассмотрим следующую систему:

$$\begin{cases} x_{k+1} = A_k x_k + w_k, \\ y_{k+1} = C_{k+1} x_{k+1} + v_{k+1}. \end{cases}$$

Здесь  $x_0, w_0, \dots, w_{N-1}, v_0, \dots, v_{N-1}$  независимые в совокупности случайные величины.  $Y_{N-1} = (y_0, \dots, y_{N-1})^T$  — все наблюдения, а  $X_{N-1} = (x_0, \dots, x_{N-1})$  — исходный процесс, который надо найти. Для этого воспользуемся фильтром Калмана, а точнее его схемой „шагаем мерим“, общий вид которой совпадает с вышеприведенной системой.

Фильтр Калмана для схемы „шагаем мерим“ имеет вид:

$$\begin{cases} x_{k+1|k} = A_k x_{k|k}, \\ x_{k+1|k+1} = x_{k+1|k} + R_{k+1|k} C_{k+1}^T (C_{k+1} R_{k+1|k} C_{k+1}^T + N_{k+1})^{-1} (y_{k+1} - C_{k+1} x_{k+1|k}), \\ R_{k+1|k} = A_k R_{k|k} A_k^T + M_k, \\ R_{k+1|k+1} = R_{k+1|k} - R_{k+1|k} C_{k+1}^T (C_{k+1} R_{k+1|k} C_{k+1}^T + N_{k+1})^{-1} C_{k+1} R_{k+1|k}, \\ x_{0|0} = \mathbb{E} x_0, \\ R_{0|0} = \text{Var } x_0. \end{cases}$$

*Замечание 10.1.* Нам действительно достаточно рассматривать только первые и вторые моменты, потому что исследуемый случайный процесс имеет гауссовское распределение, а значит полностью определяется этими величинами.

В нашей задаче  $x_k$  — процесс Орнштейна–Уленбека с параметрами  $\sigma$  и  $\lambda$ ,  $y_{k+1} = x_{k+1} + v_{k+1}$ . То есть сразу имеем, что  $C = 1$ , а  $N_k$  — дисперсия белого шума (обозначим эту дисперсию за  $\sigma_v^2$ ).

В системе нам пока неизвестны величины  $A_k$  и  $M_k$ . Найдем их. Для начала условимся, что  $t_{k+1} - t_k = \Delta t$  постоянная величина для каждого испытания. Тогда с одной стороны мы имеем

$$\text{Var } x_{k+1} = A_k^2 \text{Var } x_k + \text{Var } w_k = A_k^2 \text{Var } x_k + M_k,$$

$$\begin{aligned} \text{cov}(x_{k+1}, x_k) &= \mathbb{E}(x_{k+1} x_k) - \mathbb{E} x_{k+1} \mathbb{E} x_k = \mathbb{E}(X_k x_k^2 + w_{k+1} x_k) - A_k (\mathbb{E} x_k)^2 = \\ &= \{ \mathbb{E} w_{k+1} = 0, w_{k+1} \text{ и } x_k \text{ независимы} \} = A_k (\mathbb{E} x_k^2 - (\mathbb{E} x_k)^2) = A_k \text{Var } x_k. \end{aligned}$$



*Замечание 10.2.* Здесь и далее считаем распределение помехи  $v$  гауссовским. В случае распределения Коши матожидание помехи не определено и фильтрацию получить не получится.

С другой стороны мы знаем ковариационную функцию процесса Орнштейна–Уленбека  $R(t, s) = \sigma^2 e^{-\lambda|t-s|}$ . Это дает нам следующую систему:

$$\begin{cases} A_k^2 \text{Var } x_k + M_k = \sigma^2, \\ A_k \text{Var } x_k = \sigma^2 e^{-\lambda \Delta t}, \\ \text{Var } x_k = \sigma^2. \end{cases}$$

Как итог получаем все необходимые значения:

$$A_k = e^{-\lambda \Delta t} \quad \text{и} \quad M_k = \sigma^2 (1 - e^{-2\lambda \Delta t}).$$

Собирая все вместе, получаем фильтр Калмана для нашей задачи:

$$\begin{cases} x_{k+1|k} = e^{-\lambda \Delta t} x_{k|k}, \\ x_{k+1|k+1} = x_{k+1|k} + R_{k+1|k} (R_{k+1|k} + \sigma_v^2)^{-1} (y_{k+1} - x_{k+1|k}), \\ R_{k+1|k} = e^{-2\lambda \Delta t} R_{k|k} + \sigma^2 (1 - e^{-2\lambda \Delta t}), \\ R_{k+1|k+1} = R_{k+1|k} - R_{k+1|k} (R_{k+1|k} + \sigma_v^2)^{-1} R_{k+1|k}, \\ x_{0|0} = 0, \\ R_{0|0} = \sigma^2. \end{cases}$$

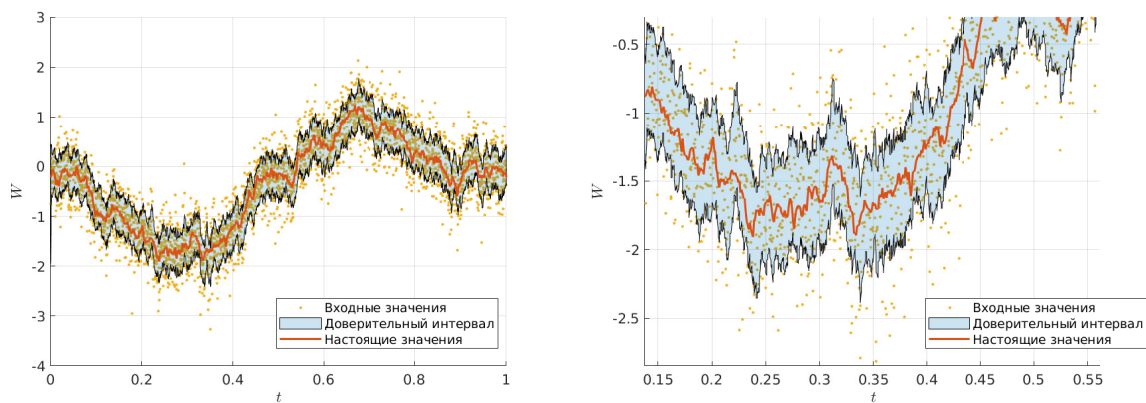


Рис. 10.1: Демонстрация работы фильтра Калмана для гауссовского шума с параметрами распределения  $\mu = 0$ ,  $\sigma^2 = 0.4$  на всем отрезке (слева) и приближенно (справа).

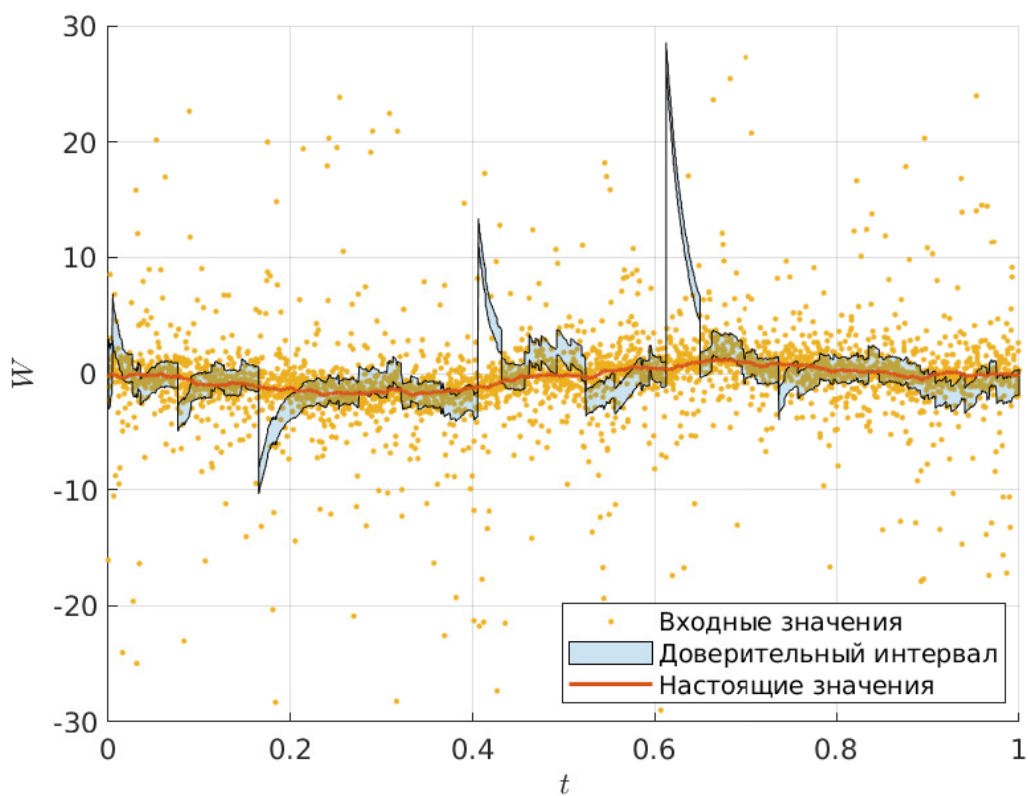


Рис. 10.2: Демонстрация работы фильтра Калмана для шума с распределением Коши с параметрами  $a = 0$ ,  $b = 1$ . Видно, что фильтр Калмана не подходит для избавления от шумов, у которых не определено математическое ожидание.

## 11 Задание №11

Построить двумерное пуассоновское поле, отвечающее сложному пуассоновскому процессу:

1. Первая интерпретация: система массового обслуживания. При этом первая координата поля — время поступления заявки в СМО (равномерное распределение), вторая — время её обслуживания (распределение  $\chi^2$  с 10-ю степенями свободы).
2. Вторая интерпретация: система массового обслуживания с циклической интенсивностью  $\lambda(t) = \lambda_0(1 + \cos(t))$  и единичными скачками. Свести данную задачу моделирования неоднородного пуассоновского процесса при помощи метода Льюиса и Шедлера к моделированию двумерного пуассоновского поля, где первая координата имеет равномерное распределение, а вторая — распределение Бернулли.
3. Третья интерпретация: работа страховой компании. Первая координата — момент наступления страхового случая (равномерное распределение), вторая координата — величина ущерба (распределение Парето). Поступление капитала по времени линейно со скоростью  $c > 0$ , начальный капитал  $W > 0$ .
4. Для каждой системы рассмотреть всевозможные случаи поведения системы в зависимости от значения параметров.

### 11.1 Система массового обслуживания

Пусть  $\lambda$  — интенсивность пуассоновского поля. Время поступления новых заявок генерируются таким образом, что

$$\Delta t_i = t_i - t_{i-1} \sim \text{Exp}(\lambda).$$

**Определение 11.1.** *Распределением  $\chi^2$  с  $k$  степенями свободы* называется распределение суммы квадратов  $k$  независимых стандартных случайных величин.

Время обслуживания  $i$ -ой заявки обозначим за  $s_i$ . Эти времена независимы и генерируются как случайные величины с распределением  $\chi^2(10)$ . Все заявки обрабатываются последовательно. Таким образом, мы можем найти время окончания обработки  $i$ -ой заявки  $Q_i$  так:

$$\begin{cases} Q_i = t_i + s_i, & \text{если } (i-1)\text{-ая заявка обработана,} \\ Q_i = Q_{i-1} + s_i, & \text{иначе.} \end{cases}$$

Или, обобщая, можно написать так:

$$Q_i = t_i + \max\{0, Q_{i-1} - t_i\} + s_i.$$

Для каждой заявки, мы будем считать количество заявок, ожидающих обработки, на момент прихода данной заявки. То есть количество  $n_i$  таких заявок  $j$ , что

$$j < i \quad \text{и} \quad Q_j > t_i.$$

Поскольку среднее время обработки одной заявки равен 10, а средний интервал между поступлениями заявок равен  $\mathbb{E} \Delta_i = 1/\lambda$ , то ожидаем отсутствия очереди при  $\lambda < 0,1$  и неограниченный рост очереди при  $\lambda > 0,1$ .

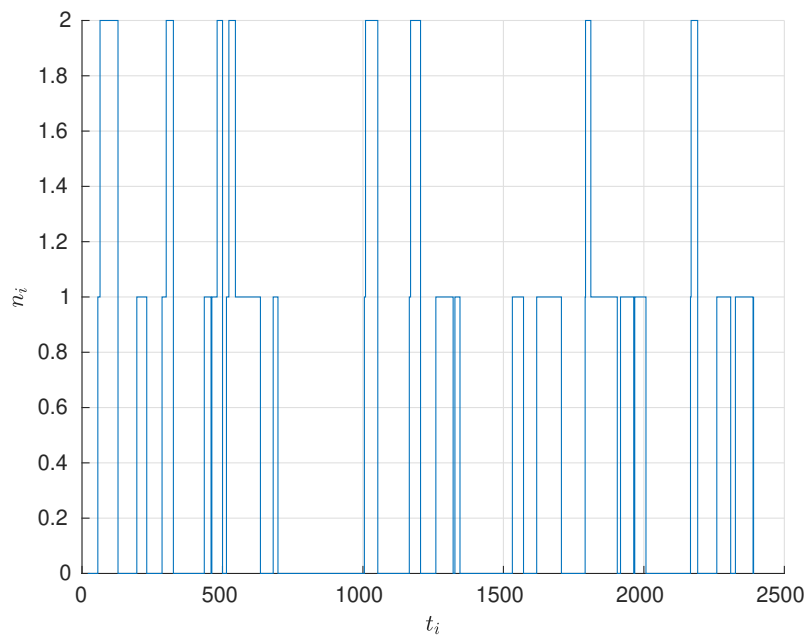


Рис. 11.1: При параметре  $\lambda = 0,05$  очередь почти не образуется.

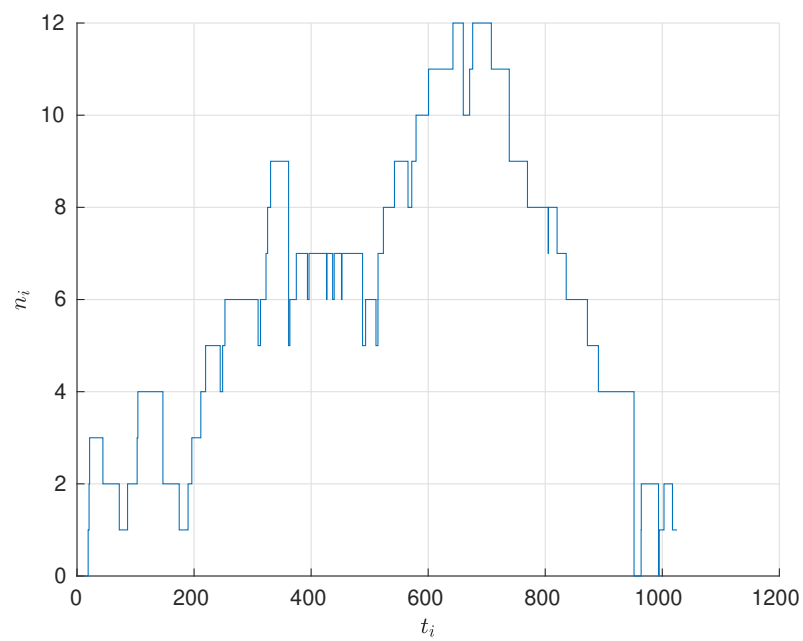


Рис. 11.2: При параметре  $\lambda = 0,1$  наблюдаем относительное равновесие. Очередь то растёт, то уменьшается.

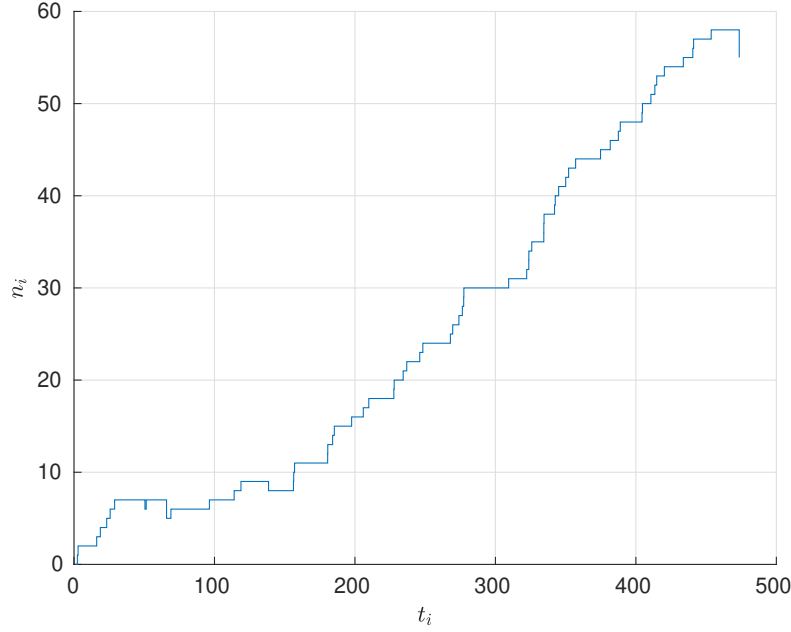


Рис. 11.3: При параметре  $\lambda = 0,02$  очередь неограниченно растёт.

## 11.2 Система массового обслуживания с циклической интенсивностью и единичными скачками

Пусть теперь  $t_1, t_2, \dots, t_n, \dots$  — времена наступления некоторых событий,  $N(t_1, t_2)$  — число событий, произошедших в промежуток времени  $[t_1, t_2]$ . Заметим, что  $t_{i+1} - t_i$  имеет функцию распределения

$$F(x) = 1 - e^{-(\Lambda(t+x) - \Lambda(t))}, \quad \text{где } \Lambda(t) = \int_0^t \lambda(u) du = \lambda(t + \sin t).$$

Здесь  $x \leq 0$ , а сам интеграл  $\Lambda(t)$  неограниченно возрастает при росте  $t$ .

Время  $t_{n+1}$  распределено как  $t_n + F^{-1}(U)$ , где  $U$  равномерно распределённая случайная величина на отрезке  $[0, 1]$ . Стоит отметить, что можно записать случайную величину  $U$  в следующем виде:

$$U = 1 - e^{-E}, \quad \text{где } E \sim \text{Exp}(1).$$

Это соотношение позволяет сделать вывод, что  $t_{n+1}$  будет распределена как  $\Lambda^{-1}(E + \Lambda(t_n))$ .

Мы будем искать обратную функцию к  $\Lambda(t)$  численно, потому что аналитически это сложно. При этом её производная  $\Lambda'(t) = \lambda_0 + \lambda_0 \cos(t)$  почти всюду возрастает. Такой метод моделирования неоднородного пуассоновского процесса называется *методом Льюиса-Шедлера*.

Для того, чтобы не искать обратную функцию, мы можем воспользоваться следующей модификацией этого метода. Обозначим за  $t$  некоторое время — причем нам не важно, произошло ли в это время какое-либо событие или нет. Теперь опишем алгоритм:

1. Будем на каждом шаге алгоритма генерировать случайную величину  $\xi$  с распределением  $\text{Exp}(2\lambda_0)$ .
2. Затем прибавим к переменной  $t$  значение величины  $\xi$  и сгенерируем новую случайную величину  $\eta = \text{Bern}\left(\frac{1+\cos t}{2}\right)$ .
3. Теперь если  $\eta$  приняла значение 1, то положим  $t_{i+1} = t$  и начнем слежующую итерацию ( $i = i + 1$ ). В противном случае, повторим процесс заново.

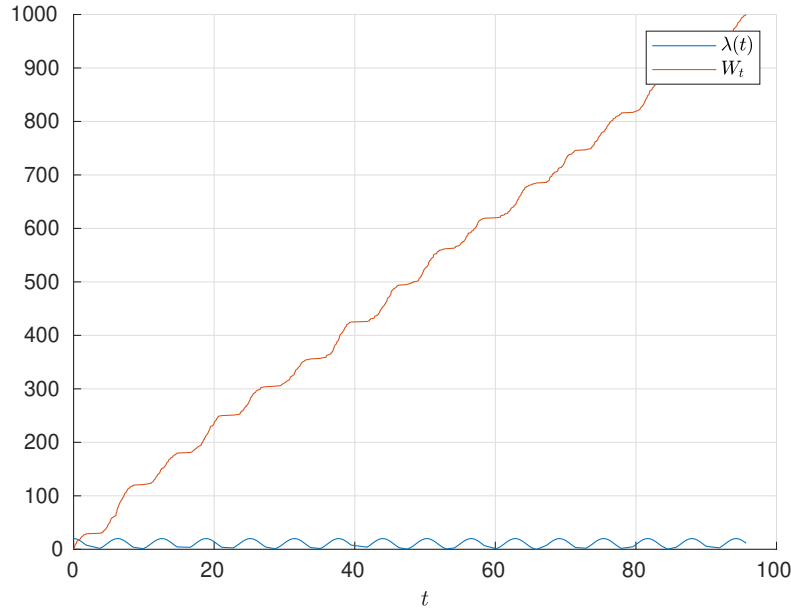


Рис. 11.4: Система массового обслуживания с циклической интенсивностью  $\lambda(t) = \lambda_0(1 + \cos(t))$  и единичными скачками.

### 11.3 Работа страховой компании

**Определение 11.2.** Случайная величина  $\xi$  имеет *распределение Парето* с параметрами  $x_m$  и  $k$ , если ее функция распределения имеет вид

$$F_\xi(x) = 1 - \left(\frac{x_m}{x}\right)^k.$$

Для моделирования паретовской случайной величины, снова воспользуемся методом обратной функции распределения. Обратная функция имеет вид

$$F_\xi^{-1}(x) = \frac{x_m}{1-x}^{1/k}.$$

Будем генерировать времена наступления страховых случаев  $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T$  следующим образом:

$$t_i - t_{i-1} \sim \text{Exp}(\lambda), \text{ где } \lambda > 0.$$

Величину ущерба  $s_i$  страхового случая в момент времени  $t_i$  будем генерировать распределением Парето с параметрами  $x_m$  и  $k$ . То есть по доказанному ранее будет равна  $x_m U^{1/k}$ , где  $U$  — равномерно распределенная на отрезке  $[0, 1]$  случайная величина.

Величина капитала компании в момент времени  $t$  выразится как

$$W(t) = W(0) + ct - s(t).$$

Здесь  $s(t)$  — сумма страховых выплат до момента времени  $t$ . В таком случае время разорения зададим условием:

$$T = \min\{t > 0 \mid W(t) < 0\}.$$

Теперь попробуем найти зависимость между капиталом компании и параметрами  $\lambda, x_m, k, W(0), c$ . Если считать  $k > 0$ , то

$$\mathbb{E} W'(t) = c - \mathbb{E}' s(t) = c - \left( \mathbb{E} \sum_{t_i < t} s_i \right)' = c - (\lambda t \mathbb{E} s_i)' = c - \frac{\lambda k x_m}{k-1}.$$

В таком случае наблюдаем зависимость: если  $c(k-1) > \lambda k x_m$  — капитал будет увеличиваться, если  $c(k-1) = \lambda k x_m$  — относительное равновесие системы,  $c(k-1) < \lambda k x_m$  — капитал будет уменьшаться.

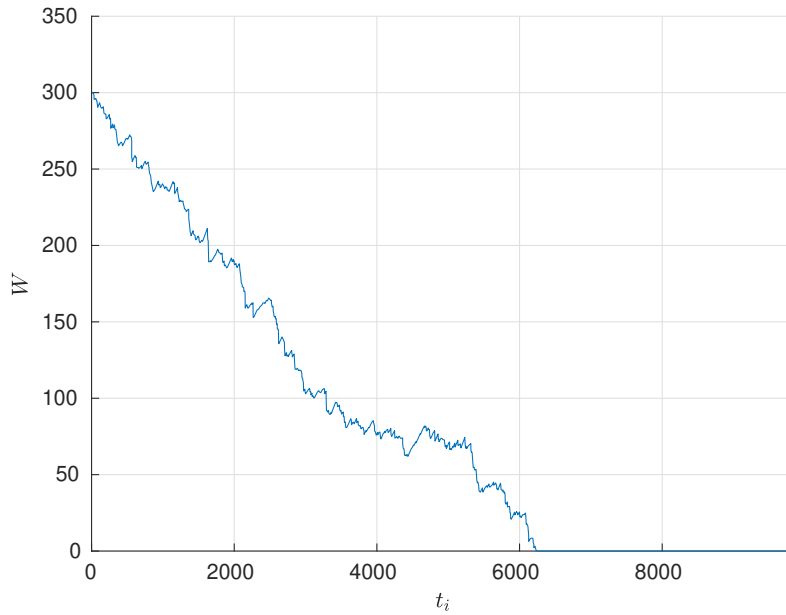


Рис. 11.5: Уменьшение капитала при  $\lambda = 0, x_m = 1, k = 2, W(0) = 300, c = 0,15$ .

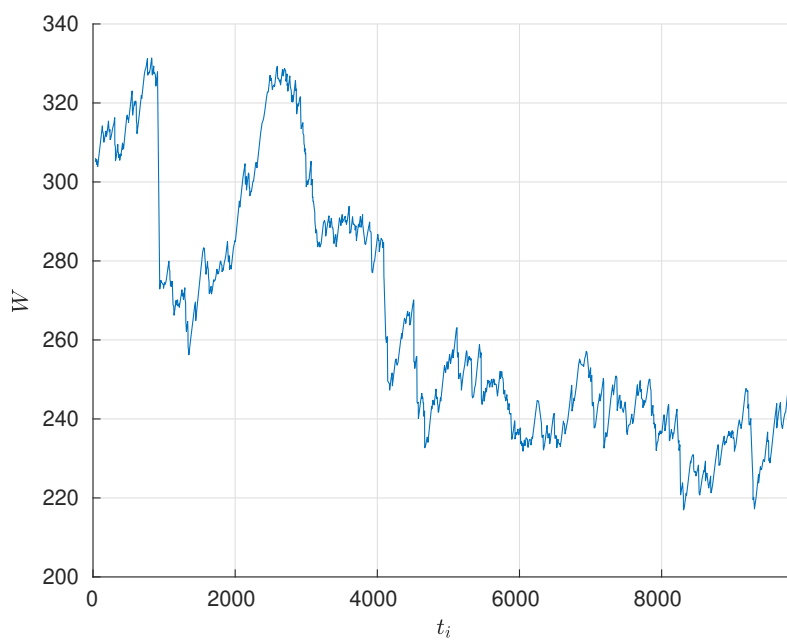


Рис. 11.6: Относительное равновесие при тех же параметрах, кроме  $c = 0,2$ .

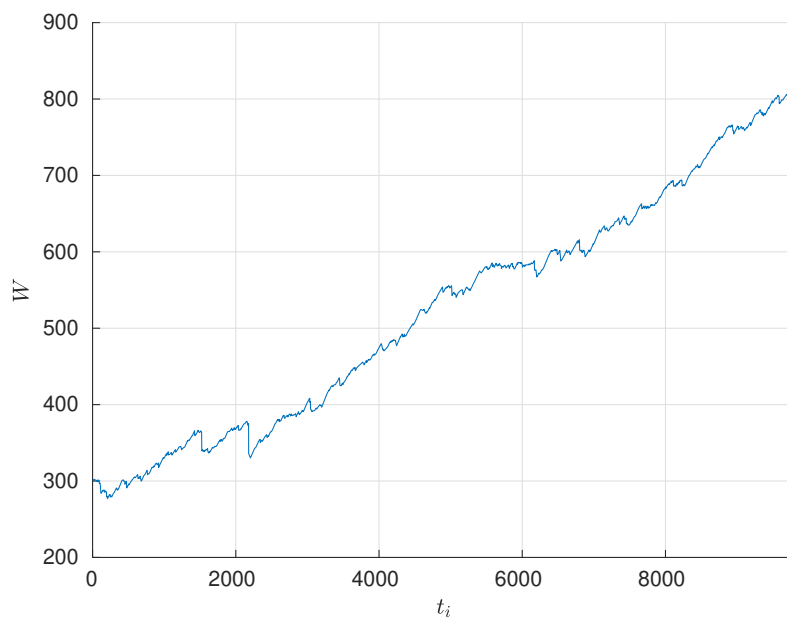


Рис. 11.7: Рост капитала при тех же параметрах, кроме  $c = 0,25$ .



## Список литературы

- [1] Ширяев А. Н. Вероятность, в 2-х кн. — 4-е изд., переработанное и дополненное — М.: МЦНМО, 2007.
- [2] Феллер В. Введение в теорию вероятностей и её приложения. — М.: Мир, 1964.
- [3] Кропачёва Н.Ю., Тихомиров А.С. Моделирование случайных величин: Метод. указания. — НовГУ им. Ярослава Мудрого — Великий Новгород, 2004.
- [4] NIST/SEMATECH, “6.3.3.1. Counts Control Charts”, e-Handbook of Statistical Methods.
- [5] Novak S.Y. Extreme value methods with applications to finance. London: CRC/ Chapman and Hall/Taylor and Francis, 2011.
- [6] Булинский А. В., Ширяев А. Н. Теория случайных процессов. — М.: «ФИЗМАТЛИТ», 2005.
- [7] Лопаткин А. С. Метод отжига — Санкт-Петербургский государственный университет, 2005.
- [8] Соболев И.М. Численные методы Монте-Карло. — Наука, 1973