



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

Отчёт по практикуму

## «Стохастический анализ и моделирование»

*Студент 415 группы*  
А. В. Бабаев

*Научный руководитель*  
доцент, к.ф.-м.н. С. Н. Смирнов

Москва, 2021

# Содержание

<b>1 Задание №1</b>	<b>4</b>
1.1 Постановка задачи . . . . .	4
1.2 Решение . . . . .	4
1.2.1 Генератор схемы Бернулли. Датчик для биномиального распределения	4
1.2.2 Геометрическое распределение . . . . .	5
1.2.3 Игра в орлянку . . . . .	5
1.3 Примеры работы программы . . . . .	6
<b>2 Задание №2</b>	<b>9</b>
2.1 Постановка задачи . . . . .	9
2.2 Решение . . . . .	9
2.2.1 Сингулярное распределения . . . . .	9
2.2.2 Свойство симметричности и самоподобия . . . . .	11
2.2.3 Вычисление математического ожидания и дисперсии . . . . .	12
2.3 Примеры работы программы . . . . .	12
<b>3 Задание №3</b>	<b>16</b>
3.1 Постановка задачи . . . . .	16
3.2 Решение . . . . .	16
3.2.1 Построение датчика экспоненциального распределения . . . . .	16
3.2.2 Реализация пуассоновского распределения на основе экспоненциального . . . . .	17
3.2.3 Датчик пуассоновского распределения как предел биномиального . . . . .	18
3.2.4 Моделирование датчика стандартного нормального распределения при помощи перехода к полярным координатам . . . . .	18
3.2.5 t-критерий Стьюдента и критерий Фишера . . . . .	19
3.3 Пример работы программы . . . . .	20
<b>4 Задание №4</b>	<b>23</b>
4.1 Постановка задачи . . . . .	23
4.2 Решение . . . . .	23
4.2.1 Датчик распределения Коши . . . . .	23
4.2.2 Метод фон Неймана . . . . .	23
4.2.3 Сравнение скоростей работы моделирования с.в. . . . .	25
4.3 Примеры работы программы . . . . .	25
<b>5 Задание 5</b>	<b>27</b>
5.1 Постановка задачи . . . . .	27
5.2 Решение . . . . .	27
5.2.1 Закон больших чисел(ЗБЧ) и центральная предельная теорема(ЦПТ) для нормального распределения . . . . .	27
5.2.2 Доверительные интервалы для среднего и дисперсии . . . . .	27
5.2.3 ЗБЧ для распределения Коши . . . . .	28
5.3 Примеры работы программы . . . . .	29

<b>6 Задание 6</b>	<b>32</b>
6.1 Постановка задачи . . . . .	32
6.2 Решение . . . . .	32
6.2.1 Метод Монте-Карло . . . . .	32
6.2.2 Метод квадратур . . . . .	34
<b>7 Задание 7</b>	<b>35</b>
7.1 Постановка задачи . . . . .	35
7.2 Решение . . . . .	35
7.2.1 Метод случайного поиска . . . . .	35
7.2.2 Метод имитации отжига . . . . .	36
7.2.3 Оценка точности вычислений . . . . .	36
<b>8 Задание 8</b>	<b>38</b>
8.1 Постановка задачи . . . . .	38
8.2 Решение . . . . .	38
8.3 Примеры работы программы . . . . .	40
<b>9 Задание 9</b>	<b>41</b>
9.1 Постановка задачи . . . . .	41
9.2 Решение . . . . .	41
9.2.1 Винеровский процесс . . . . .	41
9.2.2 Процесс Орштейна-Уленбека . . . . .	43
9.3 Примеры работы программы . . . . .	45
<b>10 Задание 10</b>	<b>46</b>
10.1 Постановка задачи . . . . .	46
10.2 Решение . . . . .	46
10.2.1 Добавлении случайной ошибки . . . . .	46
10.2.2 Фильтр Калмана . . . . .	46
10.3 Примеры работы программы . . . . .	48
<b>11 Задание 11</b>	<b>51</b>
11.1 Постановка задачи . . . . .	51
11.2 Решение . . . . .	51
11.2.1 Первая интерпретация: система массового обслуживания . . . . .	51
11.2.2 Вторая интерпретация: система массового обслуживания с циклической интенсивностью и единичными скачками . . . . .	52
11.2.3 Третья интерпретация: работа страховой компании . . . . .	53
11.3 Примеры работы программы . . . . .	54
<b>12 Список литературы</b>	<b>56</b>

# 1 Задание №1

## 1.1 Постановка задачи

1. Реализовать генератор схемы Бернулли с заданной вероятностью успеха  $p$ . На основе генератора схемы Бернулли построить датчик биномиального распределения.
2. Реализовать генератор геометрического распределения. Проверить для данного распределения свойство отсутствия памяти.
3. Рассмотреть игру в орлянку - бесконечную последовательность независимых 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  $Y(i) = S_i/\sqrt{n}$ , как функцию от номера испытания  $i = 1 \dots n$  для одной отдельно взятой траектории. Дать теоретическую оценку для  $Y(n)$  при  $n \rightarrow \infty$ .

## 1.2 Решение

### 1.2.1 Генератор схемы Бернулли. Датчик для биномиального распределения

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

1. Отсутствие взаимного влияния в испытаниях (одно на другое не оказывает влияния).
2. Воспроизводимость. Однородные испытания проводятся в схожих условиях (не одинаковых).
3. Существует признак, которые реализуется (успех) с вероятностью  $p$ , или не реализуется (неуспех) с вероятностью  $1 - p$ .

**Определение 2.** Говорят, что случайная величина  $X$  имеет распределение Бернулли, если она принимает значение 1 с вероятностью  $p$  и 0 с вероятностью  $1 - p$ . Обозначение:  $X \sim Bern(p)$ .

Реализуем генератор схемы Бернулли с заданной вероятностью успеха  $p$  следующим образом: пусть случайная величина  $\xi$  равномерно распределена на отрезке  $[0, 1]$ . Тогда случайную величину  $X \sim Bern(p)$ , можно задать в виде:

$$X = \begin{cases} 1, & \xi \in [0, p), \\ 0, & \xi \in [p, 1]. \end{cases}$$

**Определение 3.** Пусть  $X_1, \dots, X_n$  – набор независимых случайных величин с распределением Бернулли с параметром  $p$ . Тогда случайная величина:

$$Y = X_1 + \dots + X_n$$

называется случайной величиной, имеющей биномиальное распределение с параметрами  $n$  и  $p$ . Обозначение:  $Y \sim B(n, p)$ .

Случайную величину  $Y$  обычно интерпретируют, как число успехов в серии из  $n$  одинаковых независимых испытаний Бернулли с вероятностью успеха  $p$  в каждом испытании. Зададим вероятностную тройку  $(\Omega, \mathcal{F}, \mathbb{P})$ , как:

- $\Omega = \{\omega : \omega = (a_1, a_2, \dots, a_n), a_i = 0, 1, \}$
- $(P)(\omega) = p^{\sum_{i=1}^n a_i} (1-p)^{n - \sum_{i=1}^n a_i},$
- $\mathbb{F} = \{B_k\}$  где  $B_k$  событие при котором в последовательности  $\omega$  ровно  $k$  единиц.

Из этого этого следует:

$$\mathbb{P}(Y = k) = \mathbb{P}(B_k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

### 1.2.2 Геометрическое распределение

**Определение 4.** Случайная величина, равная количеству «не успехов» до появления первого «успеха» в схеме Бернулли с параметром  $p$ , называется случайной величиной, имеющей геометрическое распределение с параметром  $p$ . Обозначение:  $Geom(p)$ .

Таким образом, для  $k$  последовательных «не успехов» имеем, что:

$$\mathbb{P}(Geom(p) = k) = (1-p)^k p.$$

**Свойство 1.** Если  $Y \sim Geom(p)$ , то  $\mathbb{P}(Y > m + n | Y \geq m) = \mathbb{P}(Y > n)$  для любых целых неотрицательных  $n$  и  $m$ . Это свойство называется свойством “Отсутствия памяти” (другими словами, количество прошлых «неудач» не влияет на количество будущих «неудач»).

Рассмотрим условную вероятность того, что  $Y$  примет значение, большее  $n + m$  при условии, что  $Y \geq m$

$$\begin{aligned} \mathbb{P}(Y > m+n | Y \geq m) &= \frac{\mathbb{P}(Y > m+n \cap Y \geq m)}{\mathbb{P}(Y \geq m)} = \frac{\mathbb{P}(Y > m+n)}{\mathbb{P}(Y \geq m)} = \frac{(1-p)^{m+n+1}}{(1-p)^m} = (1-p)^{n+1} = \\ &= \mathbb{P}(Y > n). \end{aligned}$$

### 1.2.3 Игра в орлянку

Рассмотрим процесс игры в орлянку. Для этого смоделируем последовательность случайных величин  $X_1, X_2, \dots$ , где

$$X_n = \begin{cases} 1, & \xi \in [0; 0.5) \\ -1, & \xi \in [0.5; 1] \end{cases},$$

где  $\xi \sim U(0, 1)$ . Тогда

$$Y(i) = \frac{X_1 + X_2 + \dots + X_i}{\sqrt{n}}, \text{ где } i = \overline{1, n}.$$

### Центральная предельная теорема.

Пусть  $X_1, \dots, X_n, \dots$  есть бесконечная последовательность независимых одинаково

распределённых случайных величин, имеющих конечное математическое ожидание и дисперсию. Обозначим последние  $\mu$  и  $\sigma^2$ , соответственно. Пусть также

$$S_n = \sum_{i=1}^n X_i.$$

Тогда

$$\frac{S_n - \mu n}{\sigma \sqrt{n}} \rightarrow N(0, 1) \quad \text{при } n \rightarrow \infty,$$

где  $N(0, 1)$  — нормальное распределение с нулевым математическим ожиданием и стандартным отклонением, равным единице.

$$Y(n) = \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}}, \quad \text{при } n \rightarrow \infty.$$

### 1.3 Примеры работы программы

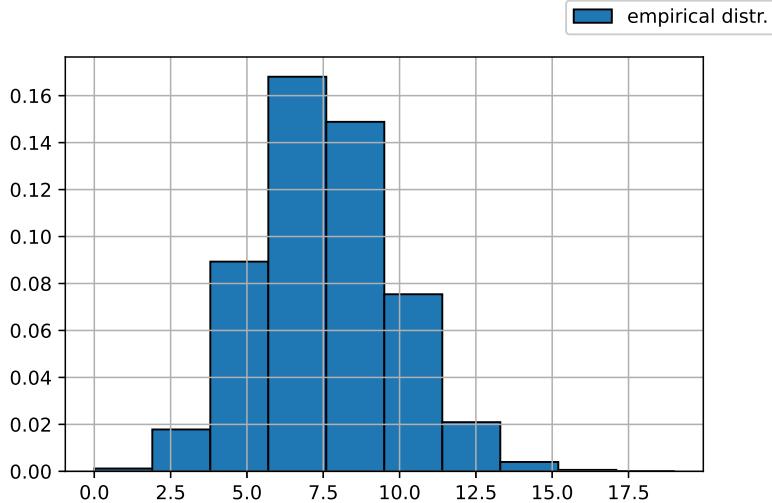


Рис. 1. Гистограмма биномиального распределения при  $p = 1/4$ ,  $n = 30$

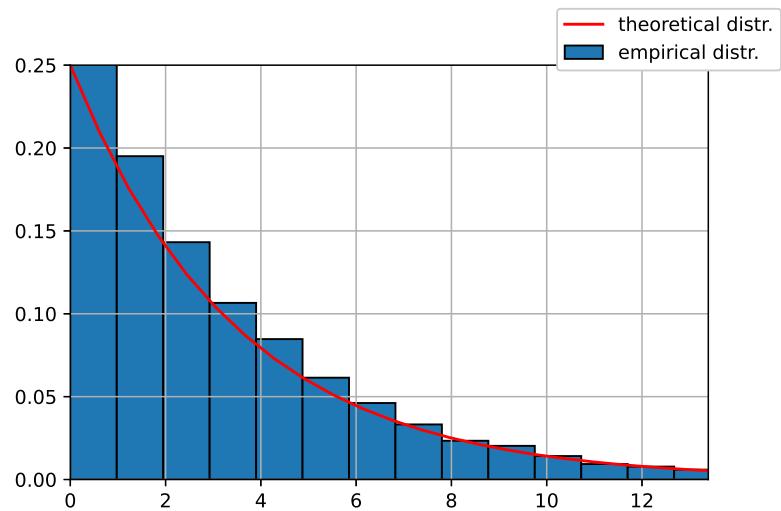


Рис. 2. Гистограмма геометрического распределения при  $p = 1/4$

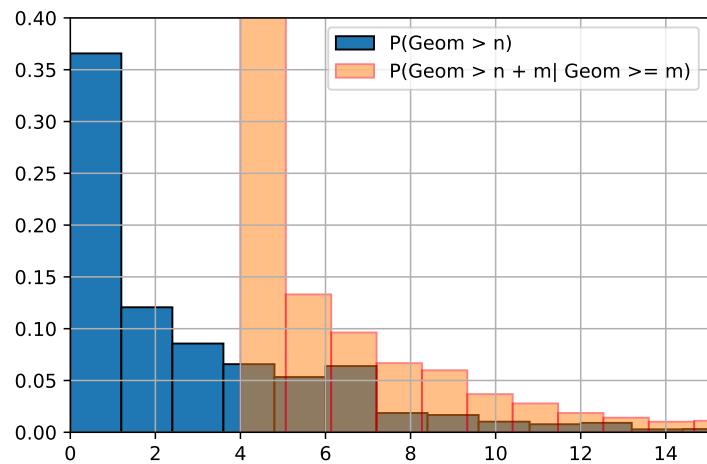


Рис. 3. Демонстрация свойства отсутствия памяти при  $m = 4, n = 0$

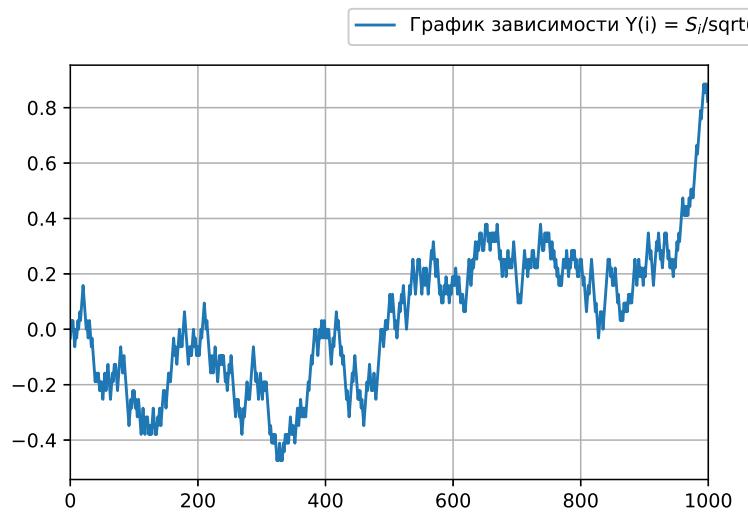


Рис. 4. Демонстрация игры в орлянку.

## 2 Задание №2

### 2.1 Постановка задачи

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 для данного распределения. Сравнить теоретические значения с эмпирическими для разного объема выборок. Проиллюстрировать сходимость.

### 2.2 Решение

#### 2.2.1 Сингулярное распределения

**Определение 5.** Распределение называется сингулярным, если оно сосредоточено на континуальном множестве с нулевой мерой Лебега.

Напомним процесс построения канторова множества.

При построении канторова множества  $F$  на отрезке  $[0, 1]$  мы выбрасываем из него интервалы  $(\frac{1}{3}, \frac{2}{3}), (\frac{1}{9}, \frac{2}{9}), (\frac{7}{9}, \frac{8}{9}), \dots$ . В итоге получаем замкнутое множество  $F$  (как пересечение замкнутых). Оно получается из отрезка  $[0, 1]$  выбрасыванием счетного числа интервалов. Отметим, что в канторовом множестве лежат только те точки, в записи которых в троичной системе счисления нет единиц. Таким образом, канторова лестница  $F(X)$  действует на такие точки по правилу:

$$F(\{a_1, a_2, \dots, a_n, \dots\}) = \{b_1, b_2, \dots, b_n, \dots\},$$

где  $a_i$  число в троичной системе счисления, а  $b_i$  число в двоичной системе счисления.

Поэтому будем генерировать элементы данного множества следующим образом:

$$X = \sum_{k=1}^{\infty} \frac{2\xi_k}{3^k}, \text{ где } \xi_k \sim Ber(0.5).$$

Суммирование ведется до бесконечности. При построении будем суммировать до некоторого разряда  $n \in \mathbb{N}$ .

**Априорная оценка ошибки** равна  $\sum_{k=n+1}^{\infty} \left(\frac{2}{3}\right)^k = 3^{-n} \Rightarrow$  достаточно взять  $3^{-n} \leq \varepsilon$  (где  $\varepsilon$  — точность вычислений). Из этого вытекает неравенство  $n \geq -\log_3 \varepsilon$  и формула:

$$n = [-\log_3 \varepsilon].$$

Для того, чтобы убедиться в корректности работы датчика, воспользуемся **критерием Колмогорова**. При проверке согласия по критерию Колмогорова в качестве меры

расхождения между гипотетическим и истинным распределениями исследуемой случайной величины используется статистика

$$D_n = \sup_{|x|<\infty} |F_n^*(x) - F(x)|,$$

представляющая собой точную верхнюю границу абсолютной величины разности между гипотетической  $F(x)$  и эмпирической  $F_n^*(x)$ , полученной по имеющейся выборке, функциями распределения.

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{X_i \leq x}, \quad \text{где } I_{X_i \leq x} = \begin{cases} 1, & X_i \leq x; \\ 0, & X_i > x. \end{cases}$$

Статистика  $D_n$  вычисляется с помощью формул

$$\begin{aligned} D_n^+ &= \max_{1 \leq j \leq n} \left| F(x_j) - \frac{j}{n} \right|; \\ D_n^- &= \max_{1 \leq j \leq n} \left| F(x_j) - \frac{j-1}{n} \right|; \\ D_n &= \max(D_n^+, D_n^-). \end{aligned}$$

Обозначим нулевую гипотезу  $H_0$ , как гипотезу о том, что выборка подчиняется распределению  $F(X)$ . Тогда по теореме Колмогорова для введённой статистики справедливо:

$$\forall x > 0: \lim_{n \rightarrow \infty} P(\sqrt{n}D_n \leq x) = K(x) = 2 \sum_{j=1}^{+\infty} (-1)^j e^{-2j^2 x^2}.$$

Т.е. при  $n \rightarrow \infty$  закон распределения случайной величины  $\sqrt{n}D_n$  независимо от вида распределения случайной величины  $X$  стремится к закону распределения Колмогорова. Учтём, что критерий имеет правостороннюю критическую область. Если статистика  $\sqrt{n}D_n$  превышает процентную точку распределения Колмогорова  $K_\alpha$  заданного уровня значимости  $\alpha$ , то нулевая гипотеза  $H_0$  (о соответствии закону  $F(x)$  отвергается). Иначе гипотеза принимается на уровне  $\alpha$ .

Количество запусков	Гипотеза принята
100	95
1000	968
10000	9527

Таблица 1: Критерий Колмогорова

### 2.2.2 Свойство симметричности и самоподобия

Проверим свойство симметричности относительно  $\frac{1}{2}$  случайных величин  $X$  и  $1 - X$ . Рассмотрим случайную величину  $1 - X$ :

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

Очевидно, что  $\eta_k \sim Bern(0.5)$ , поэтому  $X$  и  $1 - X$  распределены одинаково.

Проверим свойство самоподобия. Рассмотрим условное распределение  $X$  при условии  $X \in [0, 1/3]$ . Из построения случайной величины  $X$  видно, что  $X \in [0, 1/3] \Leftrightarrow \xi_1 = 0$ . Если  $X \in [0, 1/3]$ , то

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

Для теоретического подтверждения данного факта воспользуемся критерием Смирнова.

**Критерий однородности Смирнова** используется для проверки гипотезы о принадлежности двух независимых выборок одному закону распределения, то есть того, что два эмпирических распределения соответствуют одному и тому же закону. Статистика критерия Смирнова измеряет различие между эмпирическими функциями распределения, построенными по выборкам

$$D_{m,n} = \sup_{|x|<\infty} |G_m(x) - F_n(x)|.$$

Если гипотеза справедлива, то при неограниченном увеличении объемов выборок,

$$\forall x > 0: \lim_{m,n \rightarrow \infty} P\left(\sqrt{\frac{nm}{n+m}} D_{m,n} \leq x\right) = K(x) = 2 \sum_{j=1}^{+\infty} (-1)^j e^{-2j^2 x^2}$$

т.е. статистика  $\sqrt{\frac{nm}{n+m}} D_{m,n}$  в пределе подчиняется распределению Колмогорова  $K(x)$ .

Если статистика  $\sqrt{\frac{nm}{n+m}} D_{m,n}$  превышает квантиль распределения Колмогорова  $K_\alpha$  для заданного уровня значимости  $\alpha$ , то нулевая гипотеза  $H_0$  (об однородности выборок) отвергается. Иначе гипотеза принимается на уровне  $\alpha$ .

Количество запусков	Гипотеза принята
100	92
1000	913
10000	8945

Таблица 2: Критерий Смирнова для симметричности

Количество запусков	Гипотеза принята
100	93
1000	928
10000	8993

Таблица 3: Критерий Смирнова для самоподобия

### 2.2.3 Вычисление математического ожидания и дисперсии

Вычислим значение математического ожидания:

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

где  $\mathbb{E}\xi_k = \frac{1}{2}$  — математическое ожидание случайной величины, имеющей распределение Бернулли с параметром  $p = 0.5$ .

Вычислим значение дисперсии:

$$\mathbb{D}X = \mathbb{D} \sum_{k=1}^{\infty} \frac{2\xi_k}{3^k} = \{\mathbb{D}[aX] = a^2\mathbb{D}[X] - \text{свойство}\} = \sum_{k=1}^{\infty} \frac{4\mathbb{D}\xi_k}{3^{2k}} = \sum_{k=1}^{\infty} \frac{1}{3^{2k}} = \frac{1}{8},$$

где  $\mathbb{D}\xi_k = pq = \frac{1}{4}$  — математическое ожидание случайной величины, имеющей распределение Бернулли с параметром  $p = 0.5$  и  $q = 1 - p = 0.5$ .

## 2.3 Примеры работы программы

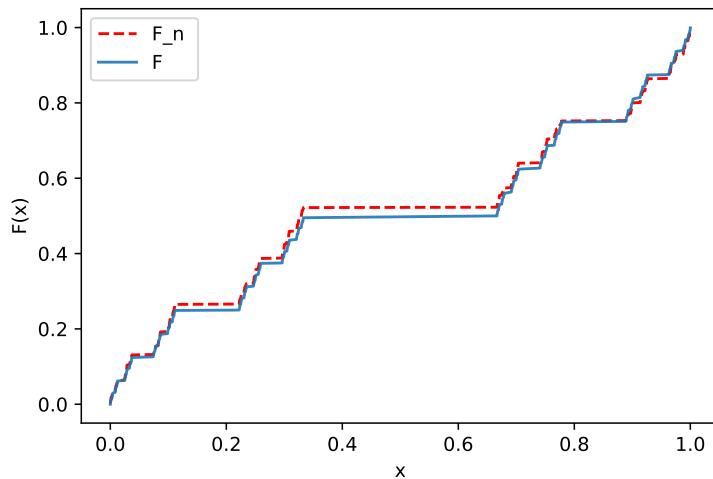


Рис. 5. Канторова лестница (теоретическая и эмпирическая) при  $n = 1000$ .

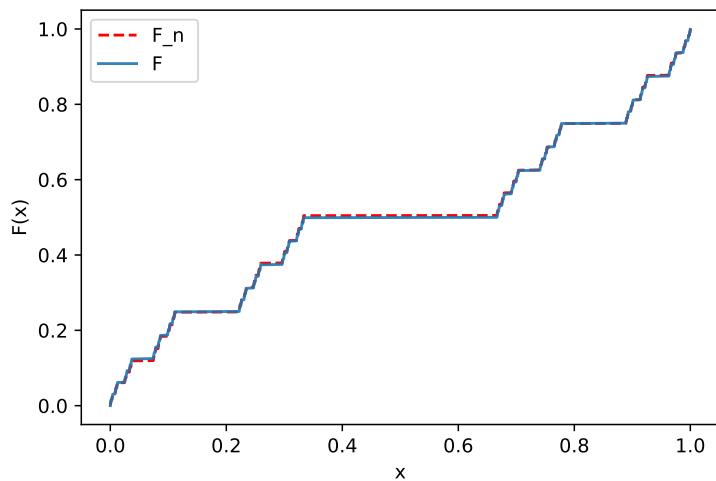


Рис. 6. Канторова лестница (теоретическая и эмпирическая) при  $n = 10000$ .

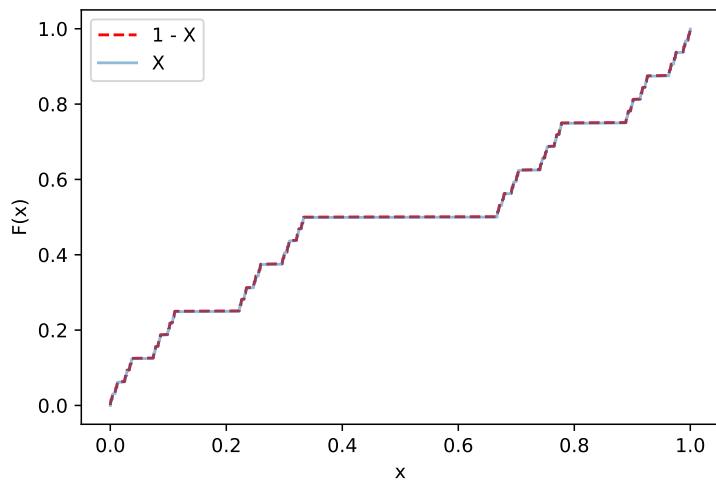


Рис. 7. Свойство симметричности  $X \sim 1 - X$ .

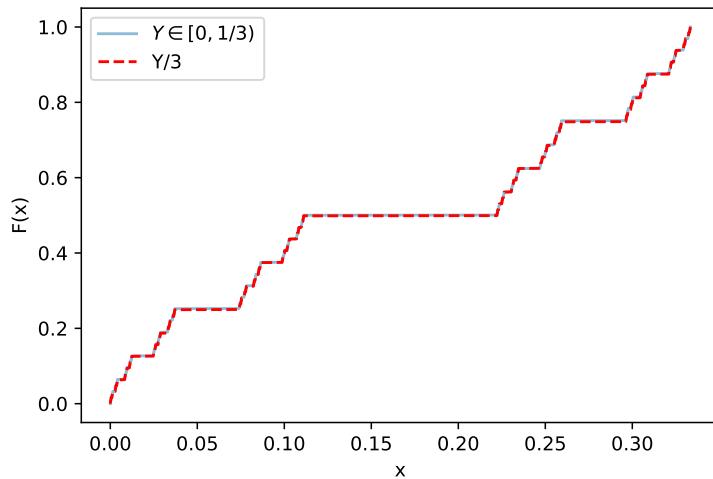


Рис. 8. Свойство самоподобия  $Y/3 \sim (Y \in [0, 1/3])$ .

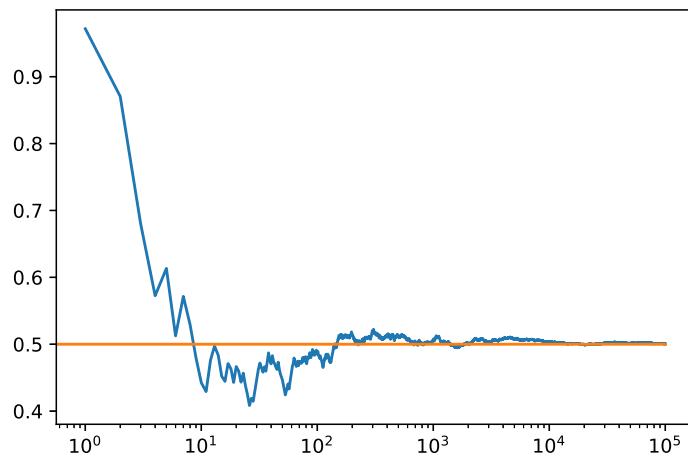


Рис. 9. Демонстрация сходимости эмпирического математического ожидания к теоретическому.

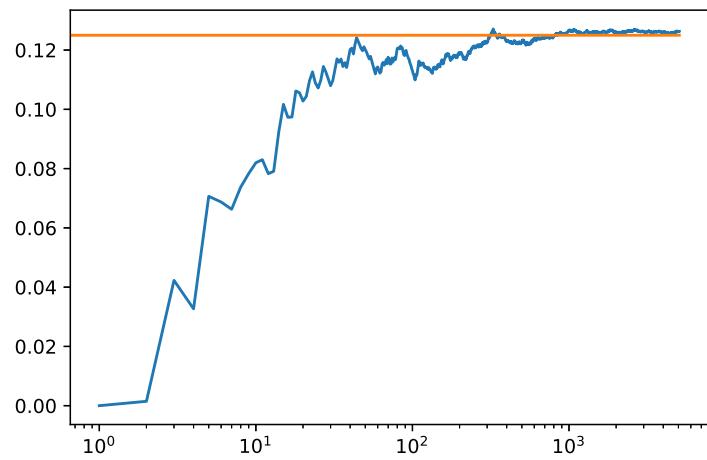


Рис. 10. Демонстрация сходимости эмпирической дисперсии к теоретической.

### 3 Задание №3

#### 3.1 Постановка задачи

1. Построить датчик экспоненциального распределения. Проверить для данного распределения свойство отсутствия памяти. Пусть  $X_1, X_2, \dots, X_n$  — независимо экспоненциально распределенные с.в. 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 датчик экспоненциального распределения построить датчик пуассонского 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 нормального распределения методом моделирования случайных величин парами с переходом в полярные координаты. Проверить при помощи критерия  $t$ -критерия Стьюдента равенство математических ожиданий, а при помощи критерия Фишера равенство дисперсий.

#### 3.2 Решение

##### 3.2.1 Построение датчика экспоненциального распределения

**Определение 6.** Случайная величина  $X$  имеет экспоненциальное распределение с параметром  $\lambda > 0$ , если ее функция распределения имеет вид:

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

**Теорема 1.** Пусть на  $\mathbb{R}$  определена функция  $F(x)$  такая, что:

- $F(x)$  непрерывна,
- $F(x)$  монотонно возрастает,
- $F(x) \rightarrow 0$  при  $x \rightarrow -\infty$ ,
- $F(x) \rightarrow 1$  при  $x \rightarrow \infty$ .

Пусть так же задана с.в.  $Y \sim U[0, 1]$ . Тогда функция  $F(x)$  является функцией распределения с.в.  $X = F^{-1}(Y)$ .

*Доказательство.* Найдем функцию распределения  $X$ :

$$F_X(x) = \mathbb{P}(X < x) = \mathbb{P}(F^{-1}(Y) < x) = P(Y < F(x)) = F(x).$$

□

Применим эту теорему для моделирования датчика экспоненциального распределения на основе датчика равномерного.

$$F_\xi(x) = 1 - e^{-\lambda x} \Rightarrow F_\xi^{-1}(x) = -\frac{1}{\lambda} \ln(1 - x).$$

Таким образом, если  $Y \sim U[0, 1]$ , то:

$$X = -\frac{1}{\lambda} \ln(1 - Y)$$

имеет экспоненциальное распределение с параметром  $\lambda$ .

Покажем, что экспоненциальное распределение обладает свойством отсутствия памяти. Для этого применим аксиому условной вероятности:

$$\mathbb{P}(X \geq s + t | X \geq t) = \frac{\mathbb{P}(X \geq s + t, X \geq t)}{\mathbb{P}(X \geq t)} = \frac{\mathbb{P}(X \geq s + t)}{\mathbb{P}(X \geq t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = \mathbb{P}(X \geq s).$$

Это выполнено для  $\forall t \neq 0$  и  $\forall s$ . Таким образом получаем:

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

Значит экспоненциальное распределение обладает свойством отсутствия памяти.

**Утверждение 1.** Пусть  $X_1, \dots, X_n$  независимые с.в., и  $X_i \sim Exp(\lambda_i)$ . Тогда  $Y = \min_{i=1, \dots, n} (X_i) \sim Exp(\sum_{i=1}^n \lambda_i)$ .

*Доказательство.* Запишем функцию экспоненциального распределения с помощью определения:  $\mathbb{P}(X < x) = 1 - e^{-\lambda x} \Rightarrow \mathbb{P}(X > x) = e^{-\lambda x}$ . Тогда  $\mathbb{P}(Y = \min(X_1, \dots, X_n) > x) = \mathbb{P}(X_1 > x, \dots, X_n > x) = \{X_1, X_2, \dots, X_n - \text{независимые с.в.}\} = \prod_{i=1}^n \mathbb{P}(X_i > x) = e^{-\lambda_1 x} \dots e^{-\lambda_n x} = e^{-\sum_{i=1}^n \lambda_i x}$ .  $\square$

### 3.2.2 Реализация пуассоновского распределения на основе экспоненциального

**Определение 7.** Случайная величина  $\xi$  имеет распределение Пуассона с параметром  $\lambda$ , если

$$\mathbb{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots$$

Обозначение:  $X \sim Pois(\lambda)$ .

**Теорема 2.** Пусть  $\eta_1, \eta_2, \dots, \eta_n, \dots$  — независимые с.в., такие, что  $X_i \sim Exp(\lambda), i = \overline{1, \infty}$ . Если

$$\xi = \max\{n : \sum_{i=1}^n \eta_i < 1\},$$

то случайная величина  $\xi \sim Pois(\lambda)$ . В случае, если  $\eta_1 > 1$ , то полагаем, что  $\xi = 0$ .

Доказательство этой теоремы можно найти в [2].

Согласно этому утверждению, можно моделировать нашу с.в. следующим образом: будем генерировать с.в.  $\eta_i$  до тех пор, пока их сумма не превзойдет единицу. Тогда с.в.  $\xi = i - 1$  будет иметь распределение Пуассона с параметром  $\lambda$ .

### 3.2.3 Датчик пуассоновского распределения как предел биномиального

**Теорема 3.** Пусть с.в.  $X \sim B(n, p)$ . Пусть  $np = \lambda = \text{const}$ . Тогда при  $n \rightarrow \infty$ :

$$\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}.$$

Доказательство этой теоремы можно найти в [2]

Критерий согласия Пирсона(критерий согласия хи-критерий) наиболее часто употребляемый критерий для проверки гипотезы о принадлежности наблюдаемой выборки  $x_1, x_2, \dots, x_n$  объемом  $n$  некоторому теоретическому закону распределения  $F(x, \theta)$ .

Будем использовать этот критерий для проверки простой гипотезы:

$$H_0 : F_n(x) = Pois(x, \lambda).$$

Исследуемое распределение принимает только целые неотрицательные значения. Обозначим за  $n_i$  количество элементов в выборке равных  $i$ . Теоретическая вероятность выпадения значения  $i$  в распределении Пуассона:

$$p_i = \frac{\lambda^i}{i!} e^{-\lambda}.$$

Пусть  $k$  — максимальное значение в выборке. Построим статистику критерия  $\chi^2$ :

$$X_n^2 = n \sum_{i=1}^n \frac{\left(\frac{n_i}{n} - p_i\right)^2}{p_i}.$$

Гипотеза о пуассоновском распределении(с.в.) будет верна, если вычисленное значение статистики не превосходит критического значения  $X_{r,\alpha}^2$ , где  $r = k - 1$ ,  $\alpha$ — заданный уровень значимости.

Таблица сравнения значений статистики при разных значениях  $n$  и  $\alpha = 0.95$ :

$n$ запусков	Гипотеза принята
100	97
1000	973
10000	9621

### 3.2.4 Моделирование датчика стандартного нормального распределения при помощи перехода к полярным координатам

**Определение 8.** Случайная величина  $\xi$  имеет нормальное распределение вероятностей с параметрами  $\mu, \sigma^2 : \xi \sim \mathcal{N}(\mu, \sigma^2)$ ( $\mu$ — математическое ожидание  $\xi$ ,  $\sigma^2$  — дисперсия  $\xi$ ), если ее плотность распределения задается формулой:

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < +\infty.$$

**Определение 9.** Нормальное распределение с параметрами  $\mu$  и  $\sigma^2 = 1$  называется стандартным нормальным распределением, и ее плотность распределения имеет вид:

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad -\infty < x < +\infty.$$

Для моделирования стандартного нормального распределения рассмотрим случайную величину  $X = \{X_1, X_2\} \sim \mathcal{N}(0, 1)$ :

$$\mathbb{P}(X_1 < x_1, X_2 < x_2) = \frac{1}{2\pi} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} e^{-\frac{\xi^2 + \eta^2}{2}} d\xi d\eta.$$

Перейдем к полярным координатам:

$$\xi = \rho \cos \phi, \quad \eta = \rho \sin \phi.$$

Учтем Якобиан:

$$J = \det \begin{pmatrix} \cos \phi & -\rho \sin \phi \\ \sin \phi & \rho \cos \phi \end{pmatrix} = \rho \neq 0.$$

Тогда

$$\mathbb{P}(X_1 < x_1, X_2 < x_2) = \int_{\rho \cos \phi < x_1; \rho \sin \phi < x_2} \frac{1}{2\pi} \frac{1}{2\pi} e^{-\frac{\xi^2 + \eta^2}{2}} d\xi d\eta = \{\omega = \rho^2\} = \int_{\sqrt{\omega} \cos \phi < x_1; \sqrt{\omega} \sin \phi < x_2} \frac{1}{4\pi} e^{-\frac{\omega}{2}} d\omega d\phi$$

Подынтегральное выражение является произведением плотностей случайных величин  $Y_1 \sim Exp(\frac{1}{2})$  и  $Y_2 \sim U[0, 2\pi]$ . Таким образом, совместное распределение случайных величин  $X_1$  и  $X_2$  совпадает с совместным распределением:

$$\{\sqrt{Y_1} \cos Y_2, \sqrt{Y_1} \sin Y_2\}, \quad Y_1 \sim Exp(\frac{1}{2}), \quad Y_2 \sim U[0, 2\pi].$$

Случайные величины  $X_1$  и  $X_2$  являются независимыми, поскольку их совместное распределение равно произведению их маргинальных распределений:

$$\mathbb{P}(X_1 < x_1, X_2 < x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi} e^{-\frac{\xi^2 + \eta^2}{2}} d\xi d\eta = \int_{-\infty}^{x_1} \frac{1}{2\pi} e^{-\frac{\xi^2}{2}} d\xi \int_{-\infty}^{x_2} \frac{1}{2\pi} e^{-\frac{\eta^2}{2}} d\eta.$$

### 3.2.5 t-критерий Стьюдента и критерий Фишера

Проверим равенство математических ожиданий построенных случайных величин  $\xi$  и  $\eta$ , используя t-критерий Стьюдента. Обозначим через  $M_1$  и  $M_2$  математические ожидания первой и второй выборки соответственно. Рассмотрим разность выборочных средних  $\Delta = \bar{\xi} - \bar{\eta}$ . Если нулевая гипотеза о равенстве математических ожиданий выполнена, то математическое ожидание  $\mathbb{E}(\Delta) = M_1 - M_2 = 0$ . Зная, что  $\mathbb{D}(\xi) = \mathbb{D}(\eta) = 1$ , то

$$\mathbb{D}(\bar{\xi}) = \frac{\mathbb{D}(\sum_{i=1}^n x_i)}{n^2} = \{X_1, \dots, X_n \text{ независимы}\} = \frac{1}{n} = \mathbb{D}(\bar{\eta}),$$

то

$$\mathbb{D}(\Delta) = \frac{2}{n}.$$

Использую несмешенную оценку дисперсии:  $s_1^2 = \frac{\sum_{i=1}^n (\xi_i - \bar{\xi})^2}{n-1}$  и  $s_2^2 = \frac{\sum_{i=1}^n (\eta_i - \bar{\eta})^2}{n-1}$ , получаем несмешенную оценку дисперсии разности выборочных средних:  $s_\Delta^2 = \frac{s_1^2 + s_2^2}{n}$ .

Значит для проверки нулевой гипотезы t-статистика равна  $t = \frac{\bar{\xi} - \bar{\eta}}{\sqrt{s_1^2 + s_2^2}} \sqrt{n}$ .

Если полученное значение статистики  $t$  превосходит критическое значение  $t_{\alpha, r}$  для заданного уровня значимости  $\alpha$ , то нулевая гипотеза отвергается.

Результаты проверки равенства математического ожидания с помощью критерия Стьюдента с  $\alpha = 5\%$  представлены в следующей таблице:

$n$ запусков	Объем выборки	Гипотеза принята
100	1000	98
100	10000	96
1000	1000	958
1000	10000	944

Теперь опишем критерий Фишера равенства дисперсий. Для ранее определенных величин  $s_1^2$  и  $s_2^2$  зададим статистику:

$$F = \frac{s_1^2}{s_2^2}.$$

Проверим равенство дисперсий с помощью критерия Фишера. Если  $F < F_{\frac{\alpha}{2}}(n - 1, n - 1)$  или  $F > F_{1 - \frac{\alpha}{2}}(n - 1, n - 1)$ , то нулевая гипотеза о равенстве дисперсий отвергается, где  $F_\alpha(n - 1, m - 1)$  есть  $\alpha$ -квантиль распределения Фишера.

Результаты проверки равенства дисперсий с помощью критерия Фишера с  $\alpha = 5\%$  представлены в следующей таблице:

$n$ запусков	Объем выборки	Гипотеза принята
100	1000	95
100	10000	96
1000	1000	948
1000	10000	946

### 3.3 Пример работы программы

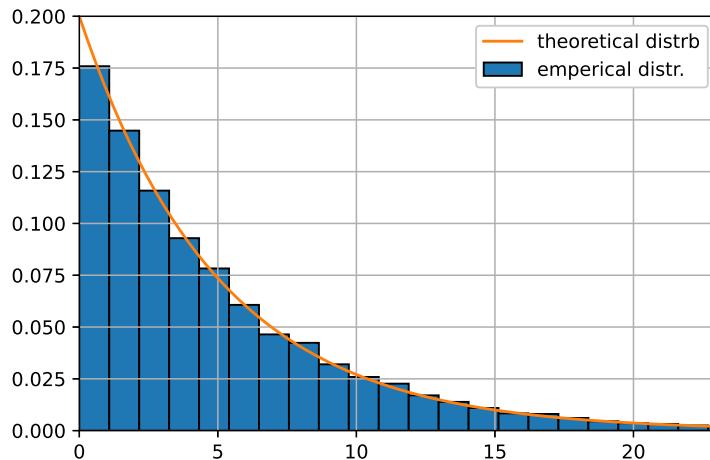


Рис. 11. Датчик экспоненциального распределения.

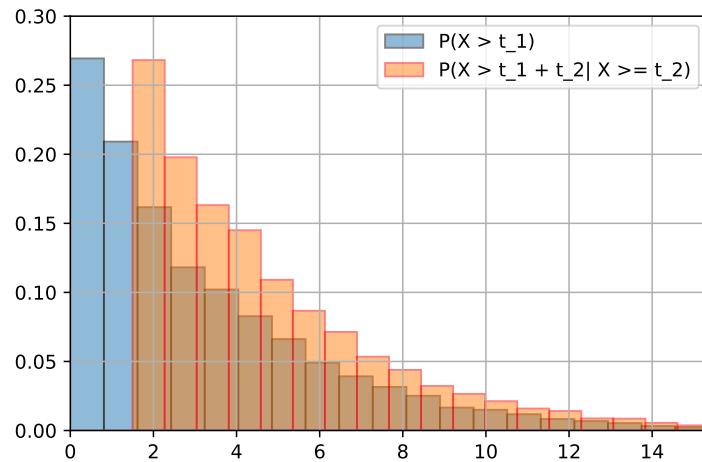


Рис. 12. Демонстрация свойства отсутствия памяти геометрического распределения.

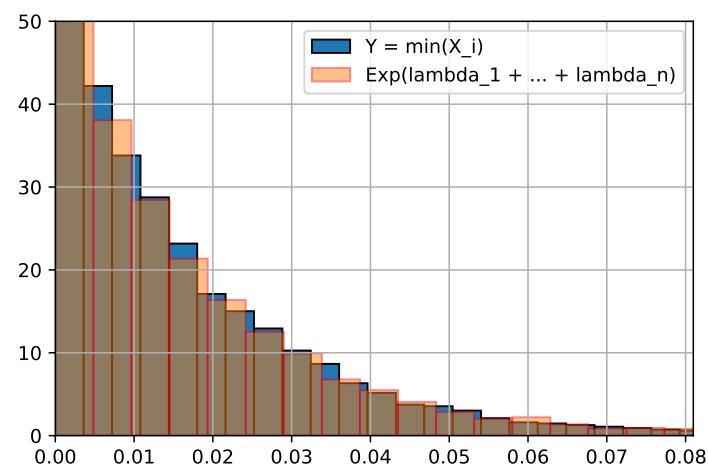


Рис. 13.  $Y = \min_{i=1,\dots,n} (X_i) \sim \text{Exp}\left(\sum_{i=1}^n \lambda_i\right)$ .

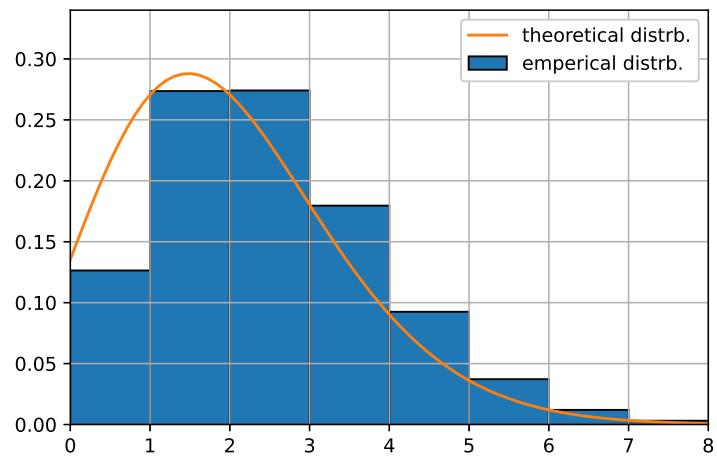


Рис. 14. Датчик распределения Пуассона на основе экспоненциального.

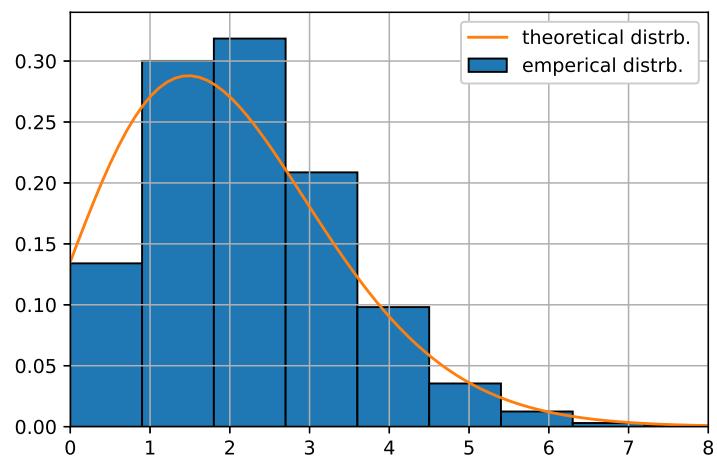


Рис. 15. Датчик распределения Пуассона как предел биномиального.

## 4 Задание №4

### 4.1 Постановка задачи

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` убедится в корректности построенного датчика и обосновать наблюдаемую зависимость.
3. Сравнить скорость моделирования стандартного нормального распределения в задания 3 и 4.

### 4.2 Решение

#### 4.2.1 Датчик распределения Коши

**Определение 10.** Случайная величина  $X$  имеет распределение Коши с параметрами  $a$  и  $b$ , если ее функция распределения имеет вид:

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

Плотность распределения Коши имеет вид:

$$p_X(x) = \frac{1}{\pi} \frac{b}{(x-a)^2 + b^2}.$$

Функция распределения с.в.  $F_X(x)$  удовлетворяет условиям теоремы 1 из задания 3. Обратная функция для данного распределения равно  $F_X^{-1}(y) = a + b \tan(\pi(y - \frac{1}{2}))$ . Значит, в качестве датчика распределения Коши можно построить датчик случайной величины  $X = F_X^{-1}(Y)$ , где  $Y \sim U[0, 1]$ .

#### 4.2.2 Метод 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 параметры  $a$  и  $b$ .

Плотность стандартного нормального распределения  $p_1(x)$  и плотность распределения Коши  $p_2(x)$  выглядят следующим образом:

$$p_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

$$p_2(x) = \frac{1}{\pi} \frac{b}{(x-a)^2 + b^2}.$$

При моделировании будем следовать алгоритму:

1. возьмем некоторое число  $k > 0$ , такое, что  $p_1(x) \leq kp_2(x)$ ,  $\forall x \in \mathbb{R}$ ,
2. рассмотрим значение случайно величины  $x = X$ , где  $X \sim Cauchy(a, b)$ ,

3. сгенерируем случайную величину  $y = Y(x) \sim Bern\left(\frac{p_1(x)}{kp_2(x)}\right)$ ,
4. если  $y = 1$ , то  $x$  — значение из распределения  $p_1(x)$ , иначе — переходим к пункту 2 алгоритма и продолжаем моделирование.

Чем ближе отношение  $\frac{p_1(x)}{kp_2(x)}$  к единице, тем быстрее работает данный алгоритм, поэтому в качестве  $k$  возьмем  $k^* = \min_{a,b} \max_x \frac{p_1(x)}{p_2(x)}$ . Рассмотрим отношение:

$$\frac{p_1(x)}{p_2(x)} = \frac{\sqrt{\pi}}{\sqrt{2}b} ((x-a)^2 + b^2) e^{-\frac{x^2}{2}}.$$

Пусть  $a = 0$ . Рассмотрим вспомогательную функцию:

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

Найдем максимум этой функции:

$$g'(x) = e^{-\frac{x^2}{2}} x (2 - b^2 - x^2) = 0$$

Корнями данного уравнения будут:  $x = 0$  при  $|b| > \sqrt{2}$  и  $x = \pm\sqrt{2-b^2}$  при  $0 < |b| \leq \sqrt{2}$ . Тогда получим:

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

Так как  $k > 0$ , то и  $b > 0$ . Найдем максимум вспомогательной функции  $h(b) = \frac{e^{\frac{b^2}{2}-1}}{b}$ :

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

Поскольку  $b > 0$ , то получим единственную точку экстремума  $b = 1$ .

Значит получим оптимальную точку при  $a^* = 0, b^* = 1$ :

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

Докажем, что  $a = 0$  — оптимальное значение параметра.

$$\begin{aligned} k^* &= \min_{a,b} \max_x \left( \frac{\sqrt{\pi}}{\sqrt{2}b} e^{-\frac{x^2}{2}} ((x-a)^2 + b^2) \right) = \min_a \left\{ \min_{b>\sqrt{2}} \frac{p_1(x)}{p_2(x)} \Big|_{x=0}, \min_{0<b\leq\sqrt{2}} \frac{p_1(x)}{p_2(x)} \Big|_{x=\pm\sqrt{2-b^2}} \right\} > \\ &> \min_a \left\{ \min_{b>\sqrt{2}} \frac{\sqrt{\pi}}{\sqrt{2}b} (a^2 + b^2), \min_{0<b\leq\sqrt{2}} (\sqrt{2-b^2} + |a|) \right\} \end{aligned}$$

Минимум выражения достигается при  $a = 0$ .

Иллюстрация работы построенного датчика, которая использует функцию normal probability plot, представлена в пункте с демонстраций работы программы. График функции распределения стандартной нормальной с.в. представляют прямую. На оси абсцисс откладываются точки выборки, на оси ординат — квантили стандартного нормального распределения.

Возьмем случайную величину  $\xi \sim \mathcal{N}(\mu, \sigma^2)$ . Тогда функция распределения:

$$F_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt.$$

Сделаем замену переменной  $s = \frac{t-\mu}{\sigma}$ . Тогда:

$$F_\xi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \exp\left(-\frac{s^2}{2}\right) ds = F\left(\frac{x-\mu}{\sigma}\right)$$

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

Таким образом, квантили различных распределений связаны между собой линейно, что означает, что любую нормальную с.в.  $\xi \sim \mathcal{N}(\mu, \sigma^2)$  можно представить в виде  $\xi = \sigma\eta + \mu$ , где  $\eta \sim \mathcal{N}(0, 1)$ , а функция normal probability plot будет прямой со сдвигом в  $\mu$  и с коэффициентом наклона  $\sigma$ , то есть корню из дисперсии.

#### 4.2.3 Сравнение скоростей работы моделирования с.в.

$n$ случайных величин	метод фон Неймана	метод перехода к полярным координатам
100	0.0327	0.00045
1000	0.2459	0.00145
10000	2.1496	0.002608

Таблица 4: Время работы методов при разных количествах с.в.

### 4.3 Примеры работы программы

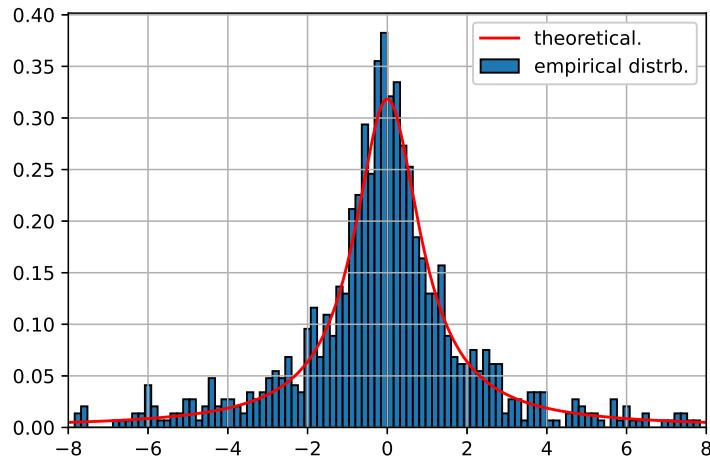


Рис. 16. Гистограмма распределения Коши.

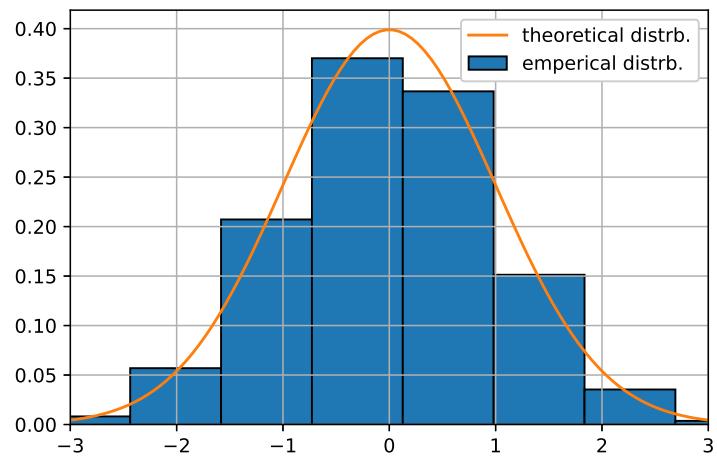


Рис. 17. Гистограмма нормального распределения, построенная с помощью метода фон Неймана, где  $\mu = 0, \sigma = 1, n = 100000$ .

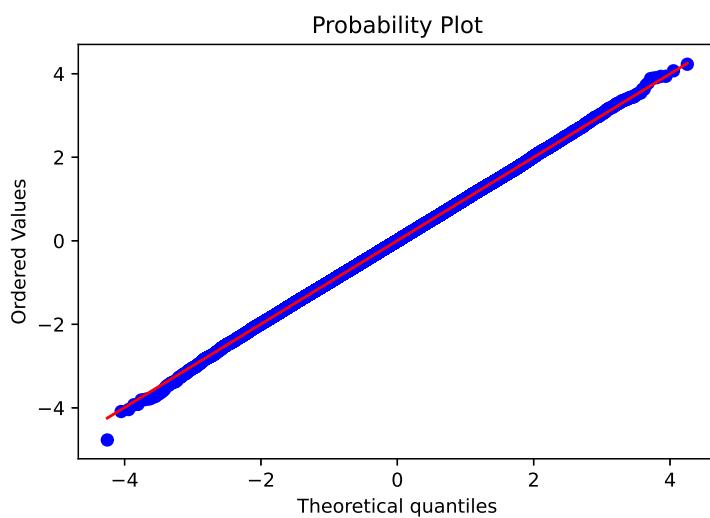


Рис. 18. Демонстрация работы функции `normal probability plot` при  $\mu = 0, \sigma = 1, n = 100000$ .

## 5 Задание 5

### 5.1 Постановка задачи

- Пусть  $X_i \sim N(\mu, \sigma^2)$ . Убедиться эмпирически в справедливости ЗБЧ и ЦПТ, т.е. исследовать поведение суммы  $S_n$  и эмпирического распределения величины

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

- Считая  $\mu$  и  $\sigma^2$  неизвестными, для пункта 1 построить доверительные интервалы для среднего и дисперсии.
- Пусть  $X_i \sim K(a, b)$  имеет распределение Коши со сдвигом  $a$  и масштабом  $b$ . Проверить эмпирически, как ведут себя суммы  $S_n/n$ . Результат работы объяснить а также найти закон распределения данных сумм.

### 5.2 Решение

#### 5.2.1 Закон больших чисел(ЗБЧ) и центральная предельная теорема(ЦПТ) для нормального распределения

Пусть  $X_i \sim N(\mu, \sigma^2)$ . Исследуем поведение суммы  $\frac{S_n}{n}$  и эмпирического распределения величины

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

**Теорема 4. (ЗБЧ)** Пусть  $X_1, X_2, \dots$  — независимые одинаково распределенные с.в.,  $\mathbb{E}X_i = \mu, \forall i \in \mathbb{N}, |\mu| < \infty, S_n = X_1 + \dots + X_n$ . Тогда

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

**Теорема 5. (ЦПТ)** Пусть  $X_1, X_2, \dots$  — независимые одинаково распределенные с.в.,  $0 < \mathbb{E}X_i^2 < \infty, \forall i \in \mathbb{N}, S_n = X_1 + \dots + X_n$ . Тогда

$$\mathbb{P}\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}}\right) \xrightarrow{n \rightarrow \infty} \Phi(x),$$

где  $\Phi(x)$  — функция стандартного нормального распределения:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

Доказательство данных теорем представлено в [2].

#### 5.2.2 Доверительные интервалы для среднего и дисперсии

Рассмотрим случайную величину  $X \sim N(\mu, \sigma^2)$ . Пусть  $x_1, x_2, \dots, x_n$  — ее реализации. Введем обозначения:

$$\text{Выборочное среднее} — \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i;$$

$$\text{Выборочная дисперсия} — s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Для параметров строятся следующие точные доверительные интервалы:

1. Для неизвестного среднего  $\mu$  при неизвестной дисперсии  $\sigma^2$ :

$$\bar{x} - \frac{s}{\sqrt{n}} t_\gamma < \mu < \bar{x} + \frac{s}{\sqrt{n}} t_\gamma,$$

где  $t_\gamma$  — критическая точка распределения Стьюдента (распределение Коши с параметрами  $a = 0, b = 1$ ) с  $n - 1$  степенями свободы и уровнем значимости  $\alpha = 1 - \gamma$

2. Для неизвестной дисперсии:

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}},$$

где  $\chi^2_{n-1}$  — критические точки  $\chi^2$  — распределения с  $n - 1$  степенями свободы и соответствующими уровнями значимости  $\alpha = 1 - \gamma$ .

### 5.2.3 ЗБЧ для распределения Коши

Пусть  $X_i \sim C(a, b)$ ,  $\forall i \in \mathbb{N}$ . Рассмотрим график:

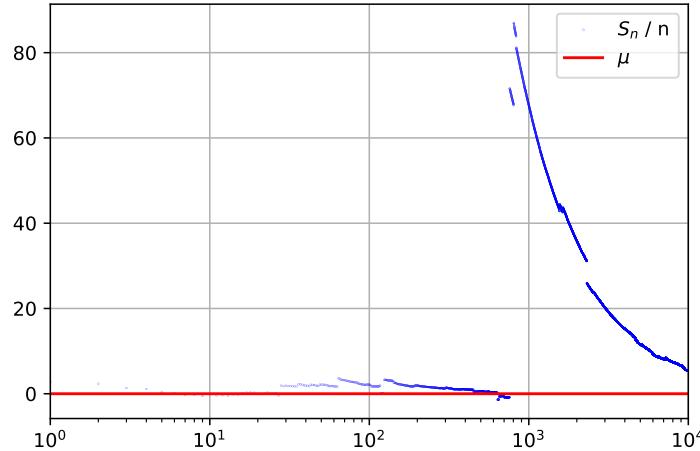


Рис. 19. Закон больших чисел для распределения Коши ( $S_n/n$ ) при  $\mu = 0, \sigma = 1$

На рисунке видно, что  $\frac{S_n}{n}$  не имеет предела, то есть закон больших чисел для распределения Коши не выполняется. Заметим, что это можно объяснить тем, что мы не можем найти математического ожидания от нашей случайной величины, то есть, мы нарушают одно из условий теоремы о ЗБЧ. Докажем, что математическое ожидание случайной величины  $X \sim C(a, b)$  не является конечным:

$$\mathbb{E}X = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{bx}{(x-a)^2 + b^2} dx = \frac{b}{2\pi} \ln((x-a)^2 + b^2) \Big|_{-\infty}^{\infty} = \infty - \infty.$$

Следовательно, эмпирический результат соответствует теории.

**Утверждение 2.** Если случайные величины  $\xi_1, \dots, \xi_n$  независимы и имеют все одно и то же распределение Коши, то среднее арифметическое  $\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$  имеет то же распределение, что и каждое  $\xi_j$ .

*Доказательство.* Докажем это утверждение с помощью характеристической функции распределения Коши:

$$\varphi(t) = \mathbb{E}e^{it\xi} = \frac{1}{\pi} \int \frac{be^{itx}}{(x-a)^2 + b^2} dx = \frac{e^{ita}}{\pi} \int \frac{e^{itbx}}{x^2 + 1} = e^{ita - |t|b}.$$

Следовательно характеристическая функция  $\frac{S_n}{n}$  имеет вид

$$\varphi_{\frac{S_n}{n}}(t) = \left( \varphi\left(\frac{t}{n}\right) \right)^n = \varphi(t),$$

$\Rightarrow \frac{S_n}{n} \sim Cauchy(a, b)$ . □

### 5.3 Примеры работы программы

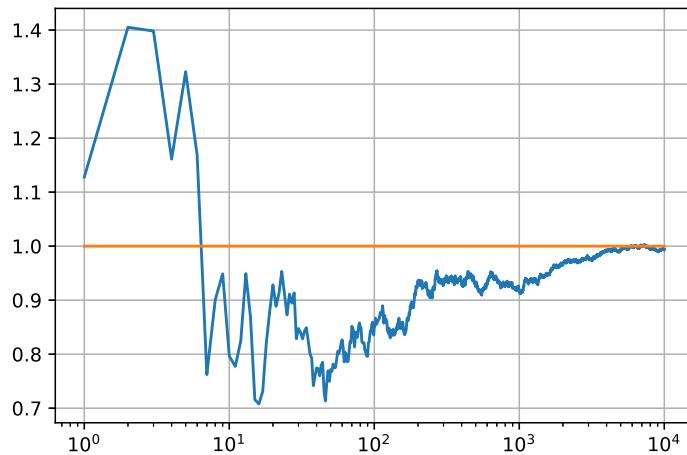


Рис. 20. Закон больших чисел для нормального распределения при  $\mu = 0, \sigma = 1$ .

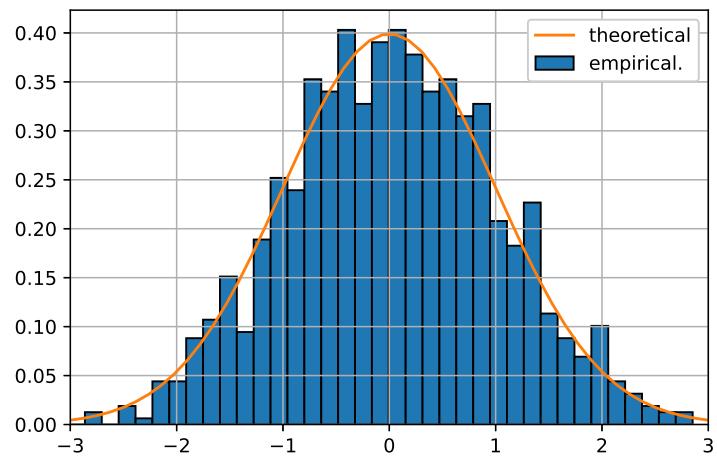


Рис. 21. Центральная предельная теорема для нормального распределения при  $\mu = 0, \sigma = 1$ .

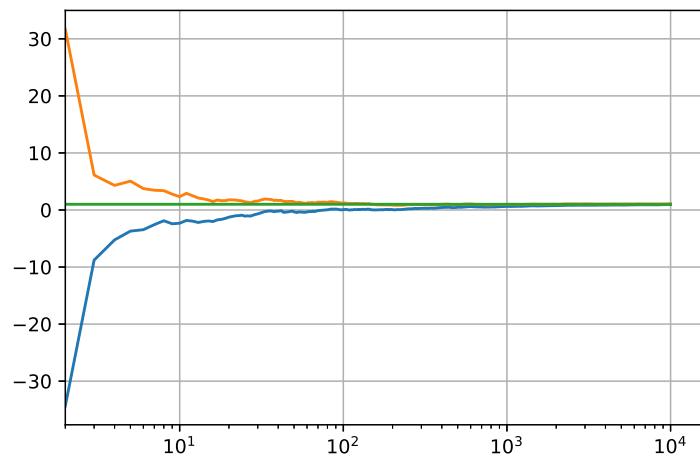


Рис. 22. Демонстрация доверительного интервала для математического ожидания при  $\mu = 1, \sigma = 3$ .

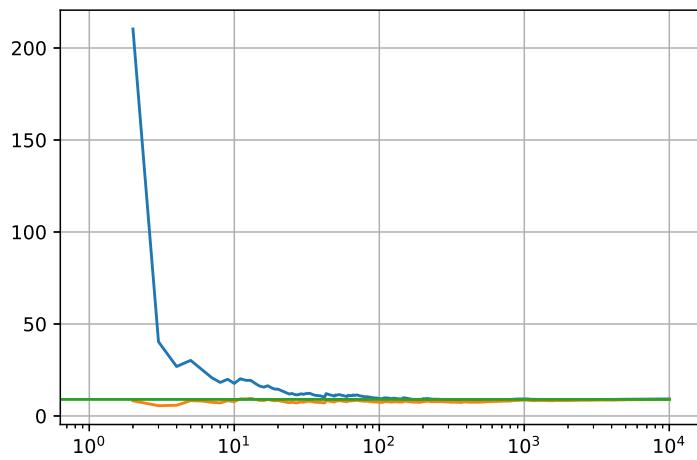


Рис. 23. Демонстрация доверительного интервала для дисперсии при  $\mu = 1, \sigma = 3$ .

## 6 Задание 6

### 6.1 Постановка задачи

1. Посчитать интеграл

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

— методом Монте-Карло

— методом квадратур, сводя задачу к вычислению собственного интеграла Римана

2. Для каждого случая оценить точность вычислений

### 6.2 Решение

#### 6.2.1 Метод Монте-Карло

Перепишем интеграл

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

в виде

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_{10}) g(x_1, \dots, x_{10}),$$

где

$$f(x) = \sqrt{\pi^{10}} \cdot \frac{e^{-\frac{1}{2^7 \cdot x_1^2 \cdots x_{10}^2}}}{x_1^2 \cdots x_{10}^2}, g(x) = \frac{1}{\sqrt{\pi^{10}}} \cdot e^{-(x_1^2 + \cdots + x_{10}^2)}.$$

Заметим, что  $g(x)$  является совместной плотностью распределения набора независимых случайных величин, имеющих нормальное распределение с параметрами 0 и  $\frac{1}{2}$ :

$$x = (x_1, \dots, x_{10}), x_i \sim N(0, \frac{1}{2}).$$

Тогда интеграл можно записать в виде:

$$I = \mathbb{E}f(x_1, \dots, x_{10}), x_i \sim N(0, \frac{1}{2}).$$

Рассмотрим выборку

$$x^i = (x_1^i, \dots, x_{10}^i), x_k^i \sim N(0, \frac{1}{2}), k = \overline{1, 10}, i = \overline{1, n}.$$

Согласно ЗБЧ выборочное среднее будет стремится к математическому ожиданию, то есть:

$$\bar{f} = \frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^n f(x^i) \xrightarrow{n \rightarrow \infty} I.$$

Оценим погрешность метода Монте-Карло с помощью центральной предельной теоремы:

$$\begin{aligned} \mathbb{P}\left(\left|\frac{S_n}{n} - I\right| < \varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n - nI}{n}\right| < \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - nI}{\sigma\sqrt{n}}\right| < \frac{\sqrt{n}}{\sigma}\varepsilon\right) = \\ &= \mathbb{P}\left(-\frac{\sqrt{n}}{\sigma}\varepsilon < \frac{S_n - nI}{\sigma\sqrt{n}} < \frac{\sqrt{n}}{\sigma}\varepsilon\right) = \Phi_0\left(\frac{\sqrt{n}}{\sigma}\varepsilon\right) - \Phi_0\left(-\frac{\sqrt{n}}{\sigma}\varepsilon\right) = \\ &= \Phi_0\left(\frac{\sqrt{n}}{\sigma}\varepsilon\right) - \left(1 - \Phi_0\left(\frac{\sqrt{n}}{\sigma}\varepsilon\right)\right) = 1 - 2\Phi_0\left(\frac{\sqrt{n}}{\sigma}\varepsilon\right) = 1 - 2\Phi_0(x_p) = \alpha, \end{aligned}$$

где

- $\Phi_0(x)$  — функция Лапласа или функция ошибок:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt,$$

- $x_p = \frac{\sqrt{n}}{\sigma}\varepsilon$  — квантиль уровня  $p$ . То есть решение уравнения:

$$\Phi_0(x_p) = p,$$

- $\alpha$  — уровень доверия.

Погрешность  $\varepsilon$  для соответствующего уровня доверия  $\alpha = 1 - 2\Phi_0(x_p)$  связана с  $x_p$  соотношением:

$$\varepsilon = \frac{\sigma x_p}{\sqrt{n}}.$$

Значение  $\sigma > 0$  используем как значение выборочной дисперсии:

$$\sigma = \frac{1}{n} \sum_{i=1}^n f^2(x_i) - \left(\frac{1}{n} \sum_{i=1}^n f(x_i)\right)^2.$$

В качестве значения уровня доверия возьмем  $\alpha = 0.99$ .

$$\mathbb{P}\left(\left|\frac{S_n}{n} - I\right| < \varepsilon\right) = \alpha = 0.99.$$

Продемонстрируем зависимость вычисления значения интеграла и полученной погрешности при разном количестве испытаний:

Количество итераций	Результат	Погрешность	Время работы
$10^2$	104.0312	256.1975	0.0042
$10^3$	177.1406	105.7845	0.06
$10^4$	106.8784	25.8910	0.3467
$10^5$	119.2904	9.1013	3.3645
$10^6$	126.8015	2.9583	25.7833
$10^7$	124.8890	0.9287	229.4606
$10^8$	124.8630	0.2935	2205.1636

### 6.2.2 Метод квадратур

Сведем задачу к вычислению собственного интеграла Римана. Для этого сделаем следующую замену переменных:

$$x_i = \tan\left(\frac{\pi}{2}t_i\right), \quad t_i \in [0, 1].$$

Таким образом по методу прямоугольников исходный интеграл приблизится значением:

$$I = \left(\frac{\pi}{2}\right)^{10} \int_{-1}^1 \cdots \int_{-1}^1 \frac{\exp\left[-\left(\sum_{k=1}^{10} \tan\left(\frac{\pi}{2}t_k\right)^2 + \frac{1}{2^7 \cdot \prod_{k=1}^{10} \tan\left(\frac{\pi}{2}t_k\right)^2}\right)\right]}{\prod_{k=1}^{10} \tan\left(\frac{\pi}{2}t_k\right)^2 \cdot \prod_{k=1}^{10} \cos\left(\frac{\pi}{2}t_k\right)^2 dt_1 \dots d_{10}}.$$

Проведем равномерное разбиение отрезка  $[0, 1]$  на  $N$  частей:

$$-1 = t_0 < t_1 < \dots < t_N = 1, \quad t_i = -1 + i \cdot \frac{1 - (-1)}{N} = i \cdot \frac{2}{N} - 1.$$

Обозначим через  $f(t_1, \dots, t_{10})$  подынтегральную функцию интеграла  $I$ . Будем использовать метод средних прямоугольников. Для этого нужно выбрать середины отрезков:

$$y_i = \frac{t_i + t_{i-1}}{2}, \quad i = \overline{1, N}.$$

Тогда интеграл можно приближенно посчитать следующим образом:

$$I_N + \left(\frac{\pi}{N}\right)^{10} \sum_{i_1=1}^N \cdots \sum_{i_{10}=1}^N f(y_{i_1}, \dots, y_{i_{10}}).$$

Оценка погрешности метода прямоугольников на равномерной сетке имеет вид:

$$\varepsilon = \frac{h^2}{24}(b-a) \sum_{i,j=1}^{10} \max \left| f''_{x_i, x_j} \right| = \frac{1}{6N^2} \sum_{i,j=1}^{10} \max \left| f''_{x_i, x_j} \right|.$$

Приведем таблицу зависимости результата от количества точек разбиения отрезка:

$N$	Результат	Время работы
3	1.2988	2.7317
4	272.7333	39.5054
5	183.4885	312.2850

## 7 Задание 7

### 7.1 Постановка задачи

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

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

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.

### 7.2 Решение

#### 7.2.1 Метод случайного поиска

Возьмем единичный круг и сгенерируем на нем набор равномерно распределенных по нему точек. Найдем минимальное значение функции  $f$ .

Совместная плотность равномерного распределения случайных величин  $x_1, x_2$  на единичном круге равна:

$$f_{x_1, x_2} = \begin{cases} \frac{1}{\pi}, & x_1^2 + x_2^2 \leq 1, \\ 0, & \text{в противном случае.} \end{cases}$$

В полярных координатах:

$$\begin{cases} x_1 = r \cos \phi, & 0 \leq r \leq 1, \\ x_2 = r \sin \phi, & 0 \leq \phi \leq 2\pi \end{cases}$$

Таким образом получим:

$$\mathbb{P}((x_1, x_2) \in A) = \iint_{x_1^2 + x_2^2 \leq 1} \frac{1}{\pi} dx_1 dx_2 = \frac{1}{\pi} \int_0^1 r dr \int_0^{2\pi} d\phi = \int_0^1 dr^2 \int_0^{2\pi} \frac{1}{2\pi} d\phi. \quad (2)$$

Сделаем замену:

$$q = r^2, r = \sqrt{q}, q \in [0, 1].$$

Тогда выражение (2) примет вид:

$$\mathbb{P}((x_1, x_2) \in A) = \int_0^1 dq \int_0^{2\pi} \frac{1}{2\pi} d\phi.$$

Следовательно,  $x_1$  и  $x_2$  выражаются в виде:

$$\begin{cases} x_1 = \sqrt{q} \cos \phi, & q \sim U[0, 1], \\ x_2 = \sqrt{q} \sin \phi, & \phi \sim U[0, 2\pi]. \end{cases}$$

### 7.2.2 Метод имитации отжига

Алгоритм основывается на воссоздании физического процесса, который происходит при кристаллизации вещества, в том числе при отжиге металлов. Предполагается, что атомы уже выстроились в кристаллическую решетку, 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, причем вероятность понижается с понижением температуры. Устойчивая кристаллическая решетка соответствует минимуму энергии атомов, поэтому атом либо переходит в состояние с меньшим уровнем энергии, либо остается на месте. При помощи моделирования такого процесса ищется такая точка или множество точек, на котором достигается минимум некоторой числовой функции  $F(\bar{x})$ , где  $\bar{x} = (x_1, \dots, x_m) \in X$ . Решение ищется последовательным вычислением точек  $\bar{x}_0, \bar{x}_1, \dots$ , пространства  $X$ ; каждая точка, начиная с  $\bar{x}_1$ , «претендует» на то, чтобы лучше предыдущих приближать решение. Алгоритм принимает точку  $\bar{x}_0$  как исходные данные. На каждом шаге алгоритм вычисляет новую точку и понижает исходно положительную величину («температуру»). Алгоритм останавливается при определенных условиях, а именно — нулевая температура. Точка  $\bar{x}_{i+1}$  по алгоритму получается на основе текущей точке  $\bar{x}_i$  следующим образом. К точке  $\bar{x}_i$  применяется оператор  $A$ , который случайным образом модифицирует собственную точку, в результате чего получается новая точка  $\bar{x}^*$ . Точка  $\bar{x}^*$  становится точкой  $\bar{x}_{i+1}$  с вероятностью  $P(\bar{x}^*, \bar{x}_i)$ , которая вычисляется в соответствии с распределением Гиббса:

$$P(\bar{x}^* \rightarrow \bar{x}_{i+1} | \bar{x}_i) = \begin{cases} 1, & F(\bar{x}^*) - F(\bar{x}_i) < 0, \\ \exp\left(-\frac{F(\bar{x}^*) - F(\bar{x}_i)}{T_i}\right), & F(\bar{x}^*) - F(\bar{x}_i) \geq 0. \end{cases}$$

Здесь  $T_i > 0$  — элементы произвольной убывающей, сходящейся к нулю положительной последовательности, которая задает аналог падающей температуры в кристалле. Скорость убывания и закон убывания могут быть заданы по желанию создателя алгоритма. Результаты работы программы поиску минимума значения функции Розенброка методом отжига приведены в следующей таблице ( $x^0 = (0, 0)$ ,  $n = 100$ ):

Кол-во запусков программы	Минимальное значение
100	-1.288345
100	-1.288486
10000	-1.288488

### 7.2.3 Оценка точности вычислений

Пусть  $x = (x_1, x_2)$  — фактическая точка минимума,  $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$  — точка минимума, полученная методом случайного поиска. Оценим  $|x - \tilde{x}|$ . Рассмотрим график исследуемой функции.

Исследуемая функция четная по  $x_1, x_2$ , имеет несколько точек минимума, которые не являются граничными. Тогда

$$|x - \tilde{x}| \leq \varepsilon = \sqrt{\frac{p}{n}}.$$

Оценим  $|f(x) - f(\tilde{x})|$  через  $|x - \tilde{x}|$ .

Поскольку  $f$  — непрерывна, то  $f$  — липшицева, следовательно:

$$|f(a) - f(b)| \leq \|\nabla f\|_\infty |a - b| = \text{essup}_{a,b \in A} |\nabla f||a - b| = \max_{a,b \in A} |\nabla f||a - b|, \forall a, b \in A.$$

Оценим  $\max_{x_1, x_2 \in A} |\nabla f| :$

$$\begin{aligned} \left| \frac{\partial f}{\partial x_1} \right| &= \left| 3x_1^2 \sin \left( \frac{1}{x_1} \right) - x_1 \cos \left( \frac{1}{x_1} \right) + 10x_2^4 \cos \left( \frac{1}{x_2} \right) \right| \leq 3x_1^2 + |x_1| + 10x_2^4 \leq \sqrt{10} + 10, \\ \left| \frac{\partial f}{\partial x_2} \right| &= \left| 40x_1x_2^3 \cos \left( \frac{1}{x_2} \right) - 10x_1x_2^2 \sin \left( \frac{1}{x_2} \right) \right| \leq 10\sqrt{17}. \end{aligned}$$

Следовательно,  $|\nabla f| = \sqrt{\left( \frac{\partial f}{\partial x_1} \right)^2 + \left( \frac{\partial f}{\partial x_2} \right)^2} = \sqrt{(\sqrt{10} + 10)^2 + (10\sqrt{17})^2} \leq 34.26$ . Окончательная оценка точности вычислений:

$$|f(x) - f(\tilde{x})| \leq 34.26 \sqrt{\frac{p}{n}}.$$

## 8 Задание 8

### 8.1 Постановка задачи

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

$$\begin{cases} \Delta u = 0, (x, y) \in D \\ u|_{\delta D} = f(x, y) \\ u \in C^2(D), f \in C(\delta D) \\ D = \{x, y : x^2 + y^2 \leq 1\} \end{cases} \quad (3)$$

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

### 8.2 Решение

Для приближенного решения задачи выберем на плоскости достаточно мелкую квадратную сетку с шагом  $h$ . Координаты узлов пусть будут  $x_i = ih, y_j = jh$ , а значения  $u(x_i, y_j), f(x_i, y_j)$  обозначим  $u_{i,j}, f_{i,j}$ .

**Определение 11.** Узел  $(x_i, y_j)$  называют внутренним, если он и все четыре соседних с ним узла  $(x_{i-1}, y_j), (x_{i+1}, y_j), (x_i, y_{j-1}), (x_i, y_{j+1})$  принадлежит  $\overline{D}$ , в противном случае узел называют граничным.

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

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

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

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 узле положим

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

Представим себе частицу  $M$ , 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 одной и той же вероятностью равной  $1/4$  может переместиться в один из четырех соседних узлов  $(x_{i-1}, y_j), (x_{i+1}, y_j), (x_i, y_{j-1}), (x_i, y_{j+1})$  причем каждый такой единичный переход совершенно случаен и не зависит от положения частицы её прошлой истории. Будем считать, что блуждание заканчивается как только попадает в граничную точку. Пусть  $P(i, j, p, q)$  — вероятность того, что траектория частицы, вышедшей из узла  $(x_i, y_j)$ , закончится в граничном узле  $(x_q, y_q)$ . Так как блуждание точки неизбежно заканчивается на границе в первой же точке ее выхода на границу, то

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

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

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

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 свое случайное блуждание из внутреннего узла  $(x_i, y_j)$ , после первого шага с вероятностью, равной  $1/4$ , попадет в один из соседних четырех узлов  $(x_{i-1}, y_j), (x_{i+1}, y_j), (x_i, y_{j-1}), (x_i, y_{j+1})$ . Откуда по формуле полной вероятности

$$\begin{aligned} v_{i,j} &= \frac{1}{4} \sum_{(x_p, y_1) \in \delta D_h} (P(i-1, j, p, q) + P(i+1, j, p, q) + P(i, j-1, p, q) + P(i, j+1, p, q)) f_{pq} = \\ &= \frac{1}{4} (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}). \end{aligned}$$

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

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

По теореме о существовании решения внутренней задачи Дирихле решение задачи (3) существует. Найдем его для конкретной функции  $f(x, y) = x^2 - y^2$ . Будем искать его в виде  $u(x, y) = Ax^2 + By^2 + C$ . Подставив его в формулировку задачи, получим следующие условие на эти коэффициента:

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

То есть мы получили, что функция  $u(x, y) = x^2 - y^2$  является решением исходной задачи, причем решение единственno.

Согласно проведенным сверху выкладкам, численное решение может быть получено по следующему алгоритму:

- Построим квадратную сетку на  $[-1, 1] \times [-1, 1]$  с шагом  $\Delta$ .

2. Функцию во всех узлах, не принадлежащих кругу, положим равной  $NaN$ .
3. Все точки круга разделим на граничные и внутренние:
  - a) В граничных точках положим  $u(x, y) = f(x, y)$ .
  - b) Значение в каждой внутренней точке получим следующим образом. Попав во внутреннюю точку  $(x_i, y_j)$ , проведем серию из  $n$  случайных блужданий. Тогда

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

где  $(x_i^{(k)}, y_j^{(k)})$  — граничная точка, в которой завершилось  $k$ -е блуждание.

### 8.3 Примеры работы программы

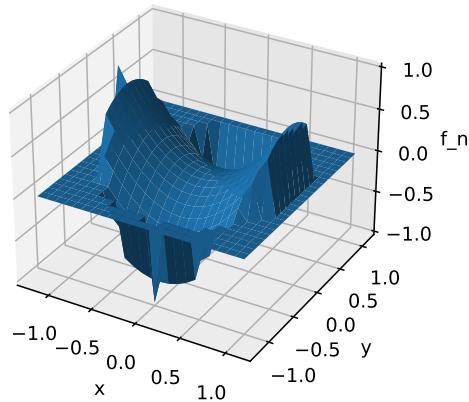


Рис. 24. Численное решение задачи.

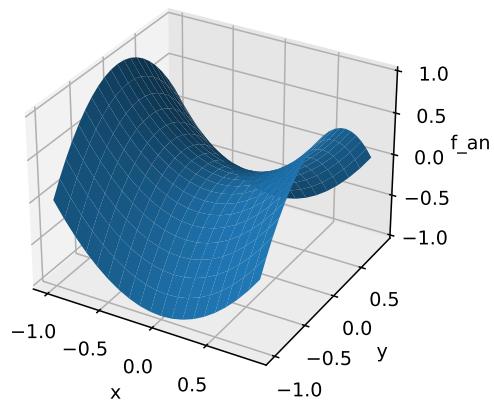


Рис. 25. Аналитическое решение задачи.

## 9 Задание 9

### 9.1 Постановка задачи

Рассмотреть два вида процессов:

- Винеровский процесс  $W(t), t \in [0, 1], W(0) = 0$ .
- Процесс Орштейна-Уленбека  $X(t), t \in [0, 1], X(0) = X_0$ , то есть стационарный марковский гауссовский процесс. Начальное значение  $X_0$  генерируется случайным образом так, чтобы полученный процесс был стационарным.

Для данных гауссовых процессов

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

### 9.2 Решение

#### 9.2.1 Винеровский процесс

**Определение 12.** Пусть дано вероятностное пространство  $(\Omega, \mathcal{F}, \mathbb{P})$ . Параметризованное семейство  $\{W_t\}_{t \in T}$  случайных величин

$$W_t(\cdot) : \Omega \rightarrow \mathbb{R}, \quad t \in T,$$

где  $T \subset [0, +\infty)$  интерпретируется как временной интервал, называется случайным процессом.

**Определение 13.** Пусть дан случайный процесс  $\{W_t\}_{t \in T}$ . Тогда он называется гауссовским, если для любых  $t_0, t_1, \dots, t_n \in T$  случайный вектор  $(W_{t_1}, W_{t_2}, \dots, W_{t_n})$  имеет многомерное нормальное распределение.

Определим винеровский процесс как гауссовский процесс в отрезке  $[0, 1]$  со средним 0 и ковариационной функцией  $cov(W(t_i), W(t_j)) = \min(t_i, t_j)$ .

Основные свойства винеровского процесса:

- $W_0 = 0$  почти наверное;
- $W_t$  является непрерывной функцией от  $t$ ;
- Приращения функции  $W(t)$  независимы и имеют нормальное распределение со средним равным 0:  $W_t - W_s \sim \mathcal{N}(0, 1)$ ,  $s < t$ .

Определим плотность  $n$ -мерного нормального распределения с невырожденной ковариационной матрицей.

**Определение 14.** Пусть  $x$  —  $n$ -мерный вектор и  $x \sim \mathcal{N}(m_x, R_x)$ . Тогда его плотность имеет вид

$$p(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{|R_x|}} e^{-\frac{1}{2}(x-m_x)^T R_x^{-1} (x-m_x)},$$

где  $R_x$  — ковариационная матрица.

Смоделируем винероский процесс методом деления отрезка  $[0, 1]$ , в отношении  $\alpha$ , исходя из следующих соображений:

1. В начальный момент времени  $W_{t_0} = 0$ , по определению;
2. Генерируем  $W_{t_1} = W_{t_0} \sim \mathcal{N}(0, 1)$ ;
3. Рассмотрим отрезок  $[t_1, t_2]$ , его внутреннюю точку  $t = t_1 + \alpha(t_2 - t_1)$  и условную плотность

$$p_{W_t}(x|W_{t_1} = x_1, W_{t_2} = x_2) = \frac{p_{W_{t_1}, W_t, W_{t_2}}(x_1, x, x_2)}{p_{W_{t_1}, W_{t_2}}(x_1, x_2)}.$$

Обозначим векторы  $\bar{x} = (x_1, x, x_2)^T$  и  $\tilde{x} = (x_1, x_2)^T$  и рассмотрим плотности вероятностей этих векторов:

$$\begin{aligned} p_{W_{t_1}, W_t, W_{t_2}} &= \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{|R_1|}} e^{-\frac{1}{2}\bar{x}^T R_1^{-1} \bar{x}}, \\ p_{W_{t_1}, W_{t_2}} &= \frac{1}{(2\pi)^{\frac{3}{2}} \sqrt{|R_2|}} e^{-\frac{1}{2}\tilde{x}^T R_2^{-1} \tilde{x}}, \end{aligned}$$

где  $R_1, R_2$  — соответствующие матрицы ковариаций. Так как ковариационная функция имеет вид  $k(s, t) = \min(s, t)$ , то находим выражения для  $R_1$  и  $R_2$ :

$$R_1 = \begin{pmatrix} t_1 & t_1 & t \\ t_1 & t & t \\ t_1 & t & t_2 \end{pmatrix},$$

$$R_2 = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix}.$$

Вычислим определители и обратные матрицы для  $R_1$  и  $R_2$ :

$$R_1^{-1} = \begin{pmatrix} \frac{t}{t_1(t-t_1)} & -\frac{1}{t-t_1} & 0 \\ -\frac{1}{t-t_1} & \frac{t_2-t_1}{(t_2-t)(t-t_1)} & -\frac{1}{t_2-t} \\ 0 & -\frac{1}{t_2-t} & \frac{1}{t_2-t} \end{pmatrix},$$

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

$$|R_1| = t_1 t_2 t + t_1^2 t - t_1 t^2 - t_1^2 t_2; \quad |R_2| = t_1 t_2 - t_1^2.$$

Получим:

$$p_{W_t}(x|W_{t_1} = x_1, W_{t_2} = x_2) = \frac{1}{\sqrt{2\pi\alpha(1-\alpha)(t_2-t_1)}} e^{-\frac{(x-(\alpha x_1 + (1-\alpha)x_2))^2}{2\alpha(1-\alpha)(t_2-t_1)}}.$$

### 9.2.2 Процесс Орштейна-Уленбека

**Определение 15.** Случайный процесс  $\{W_t\}_{t \in T}$  называется стационарным, если конечномерные распределения инвариантны относительно сдвига времени.

**Определение 16.** Гауссовский процесс  $\{W_t\}_{t \in T}$  называется процессом Орштейна-Уленбека, если он является стационарным и марковским.

Из стационарности процесса Орштейна-Уленбека следует, что

$$\mathbb{E}W_t = a, \quad R(t, s) = R(|t - s|).$$

Без ограничений общности положим  $a = 0$ .

Обозначим  $\mathbb{D}W_t = \sigma^2$ , тогда  $R(t, s)$  представима в виде  $R(t, s) = \sigma^2 \rho(s, t)$ , где  $\rho(s, t)$  — коэффициент корреляции.

**Теорема 6.** Для того чтобы последовательность  $W_1, \dots, W_n$  нормально распределенных случайных величин была марковской, необходимо и достаточно, чтобы

$$\rho_{j,k} = \rho_{j,i}\rho_{i,k} \quad \forall i, j, k : j \leq i < k \leq n,$$

где  $\rho_{i,j}$  — коэффициент корреляции случайных величин  $W_i$  и  $W_j$ .

Доказательство представлено в [7] report

В силу того, что процесс  $W_t$  является марковским, получаем, что

$$\rho(s, t) = \rho(s, \tau)\rho(\tau, t). \quad (4)$$

Поскольку  $R(s, t) = R(|s - t|)$ , то  $\rho(s, t) = \rho(s - t)$ . Тогда, введя замену переменных

$$x = s - \tau,$$

$$y = \tau - t,$$

преобразуем выражение (4) к выражению

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

**Теорема 7.** Пусть функция  $u(t)$  определена при  $t > 0$  и ограничена на каждом конечном интервале. Если  $u(t)$  удовлетворяет соотношению  $u(t + s) = u(t)u(s)$ , то или  $u(t) \equiv 0$ , или  $u(t) = e^{-\lambda t}$ , где  $\lambda$  — некоторая положительная константа.

Доказательство представлено в [4]

Если  $\rho(t) \equiv 0$ , то  $\text{cov}(W_t, W_s) = 0$ , что равносильно тому, что  $W_t$  независимы в совокупности (так как процесс гауссовский), поэтому моделирование процесса Орштейна-Уленбека заключается в моделировании случайных величин, имеющих распределение  $N(a, \sigma^2)$ .

Рассмотрим случай  $\rho(s, t) = e^{-\lambda|s-t|}$ ,  $\lambda > 0$ . Ковариационная функция процесса Орштейна-Уленбека имеет вид

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

Найдем переходную плоскость

$$p_{W_t}(x_1 | W_s = x_2) = \frac{p_{W_t, W_s}(x_1, x_2)}{p_{W_s}(x_2)}.$$

Поскольку  $W_t$  — гауссовский процесс, то

$$p_{W_t, W_s}(x_1, x_2) = \frac{1}{2\pi|C|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x, C^{-1}x) \right\},$$

$$p_{W_s}(x_2) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{x_2^2}{2\sigma^2} \right\},$$

где  $x = (x_1, x_2)$ . Ковариационная матрица  $C$  имеет вид

$$C = \begin{pmatrix} \sigma^2 & R(t, s) \\ R(t, s) & \sigma^2 \end{pmatrix}.$$

Тогда

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

Поэтому

$$p_{W_t}(x_1 | W_s = x_2) = \frac{1}{(2\pi(\sigma^2 - \frac{R^2(s,t)}{\sigma^2}))^{\frac{1}{2}}} \exp \left\{ -\frac{(x_1 - \frac{R(t,s)}{\sigma^2}x_2)^2}{2(\sigma^2 - \frac{R^2(t,s)}{\sigma^2})} \right\},$$

то есть

$$F(W_t | W_s = x_2) \sim \mathcal{N}(x_2 e^{-\lambda|t-s|}, \sigma^2(1 - e^{-2\lambda|t-s|})).$$

Для непосредственного моделирования процесса рассмотрим (как и в предыдущем случае) случайные величины  $W_{t_1}, W_{t_2}$ . Мы можем сгенерировать случайную величину  $W_t$ , где  $t_1 < t < t_2$ . Будем моделировать  $W_t$  способом аналогичным предыдущему заданию. Для упрощения положим, что  $\alpha = 1/2$ . Найдем условную вероятность:

$$p_{W_t}(x | W_{t_1} = x_1, W_{t_2} = x_2) = \frac{p_{W_{t_1} W_t W_{t_2}}(x_1, x, x_2)}{p_{W_{t_1}} p_{W_{t_2}}},$$

где  $t = \frac{t_1+t_2}{2}$ , поскольку  $W_t$  — гауссовский

$$p_{W_{t_1} W_t W_{t_2}}(x_1, x, x_2) = \frac{1}{(2\pi)^{\frac{3}{2}} |R_1|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x_1, x, x_2)^T R_1^{-1}(x_1, x, x_2) \right\},$$

$$p_{W_{t_1}} p_{W_{t_2}} = \frac{1}{2\pi |R_2|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}(x_1, x_2)^T R_2(x_1, x_2) \right\}$$

где

$$R_1 = \sigma^2 \begin{pmatrix} 1 & e^{-\lambda(t_2-t_1)} & e^{-\lambda(t_2-t_1)} \\ e^{-\lambda(t_2-t_1)} & 1 & e^{-\lambda(t_2-t_1)} \\ e^{-\lambda(t_2-t_1)} & e^{-\lambda(t_2-t_1)} & 1 \end{pmatrix}, \quad R_2 = \sigma^2 \begin{pmatrix} 1 & e^{-\lambda(t_2-t_1)} \\ e^{-\lambda(t_2-t_1)} & 1 \end{pmatrix}$$

Используя символьные вычисления получим

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

В качестве начальных значений возьмем  $W_0 \sim N(0, \sigma^2)$ ,  $W_1 \sim N(x_0 e^{-\lambda T}, \sigma^2(1 - e^{-2\lambda T}))$ .

### 9.3 Примеры работы программы



Рис. 26. Демонстрация винеровского процесса.

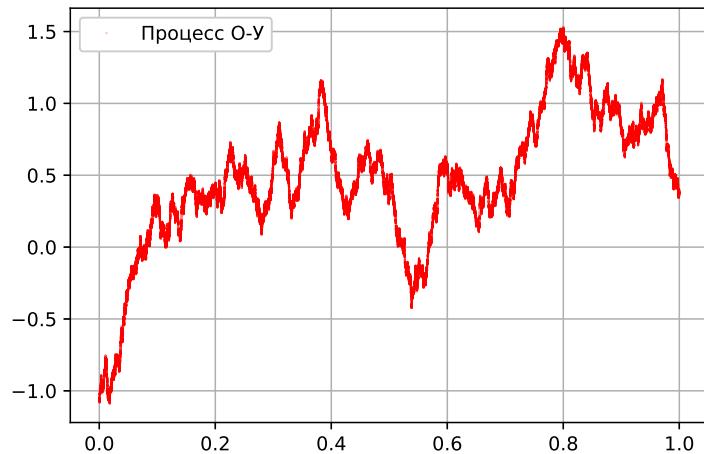


Рис. 27. Демонстрация процесса Орнштейна-Уленбека.

## 10 Задание 10

### 10.1 Постановка задачи

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 одномерного 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.

### 10.2 Решение

#### 10.2.1 Добавление случайной ошибки

**Определение 17.** Дискретным белым шумом называется последовательность  $\varepsilon_1, \dots, \varepsilon_n, \dots$  независимых одинаково распределенных случайных величин.

Рассмотрим соотношение

$$x_{k+1} = f(x_k) + \omega(k),$$

где  $\omega(k)$  — случайная помеха,  $x_k, \omega(k)$  независимы,  $f(x_k) = \mathbb{E}(x_{k+1}|x_k)$ . Пусть рассматривается марковский процесс, тогда совместная плотность по всем моментам времени

$$\begin{aligned} p(x_k, \dots, x_0) &= p(x_k|x_{k-1}, \dots, x_0) \cdot p(x_{k-1}|x_{k-2}, \dots, x_0) \cdots \cdots p(x_1|x_0) \cdot p(x_0) = \\ &= \{\text{марковский процесс}\} = p(x_k|x_{k-1}) \cdot p(x_{k-1}|x_{k-2}) \cdots \cdots p(x_1|x_0) \cdot p(x_0). \end{aligned}$$

Обратим внимание, что в случае, когда шум имеет распределение Коши, фильтрацию провести не получится, потому что распределение Коши не имеет математического ожидания. Далее будем рассматривать случай, когда шум является гауссовским.

#### 10.2.2 Фильтр Калмана

Рассмотрим линейное стохастическое уравнение

$$x_{k+1} = A_k x_k + \omega_k.$$

Поскольку случайные величины гауссовские, то для их полного описания достаточно знать их первые и вторые моменты.

Пусть имеет следующую систему:

$$\begin{cases} x_{k+1} = A_k x_k + \omega_k, \\ y_{k+1} = C_{k+1} x_{k+1} + v_{k+1}, \end{cases} \quad (5)$$

причем  $x_0, \omega_0, \dots, \omega_{N-1}, v_0, \dots, v_{n-1}$  независимы в совокупности.  $Y_{N-1} = (y_0, \dots, y_{N-1})^T$  — все наблюдения, а  $X_{N-1} = (x_0, \dots, x_{N-1})$  — исходный процесс, который надо найти. Для этого воспользуемся фильтром Калмана, а именно его схемой «шагаем-мерим», общий вид которой совпадает с (5).

Обозначим  $\mathbb{E}x_0 = \bar{x}_0, \mathbb{D}x_0 = S, \mathbb{E}\omega_k = \mathbb{E}v_k = 0, \mathbb{D}\omega_k = M_k, \mathbb{D}v_k = N_k > 0$ .

Фильтра Калмана для схемы «шагаем-мерим» имеет вид:

$$\begin{cases} \tilde{x}_{k+1|k} = A_k \tilde{x}_{k|k}, \\ \tilde{x}_{k+1|k+1} = \tilde{x}_{k+1|k} + R_{k+1|k} C_{k+1}^T (C_{k+1} R_{k+1|k} C_{k+1}^T + N_{k+1})^{-1} (y_{k+1} - C_{k+1} \tilde{x}_{k+1|k}), \\ R_{k+1|k} = A_k R_{k|k} A_k^T + M_k, \\ R_{k+1|k+1} = R_{k+1|k} - R_{k+1|k} C_{k+1}^T (C_{k+1} R_{k+1|k} C_{k+1}^T + N_{k+1})^{-1} C_{k+1} R_{k+1|k}, \\ \tilde{x}_{0|0} = \bar{x}_0, \\ R_{0|0} = S. \end{cases}$$

В нашей задаче  $x_k$  — процесс Орнштейна-Уленбека с параметрами  $\sigma_W$  и  $\lambda$ ,  $y_{k+1} = x_{k+1} + v_{k+1}$ , где  $v$  — белый шум. Пусть  $\sigma_n^2$  — его дисперсия. Тогда получаем, что  $N_k = \sigma_n^2$ , а  $C_k = 1$ . Осталось найти  $A_k$  и  $M_k$ . Будем считать, что  $t_{i+1} - t_i = \Delta t$  независимо от  $i$ . Так как мы рассматриваем одномерный процесс Орнштейна-Уленбека, то  $A_k, C_k$  являются скалярами, и от их транспонирования ничего не меняется. Обозначим  $\mathbb{D}x_k = V_k$ .

С одной стороны имеем

$$\mathbb{D}x_{k+1} = A_k^2 \mathbb{D}x_k + \mathbb{D}\omega_k = A_k^2 V_k + M_k,$$

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

С другой стороны так как ковариационная функция процесса Орнштейна-Уленбека имеет вид  $R(t, s) = \sigma_W^2 e^{-\lambda|t-s|}$ , то получим следующую систему уравнений:

$$\begin{cases} A_k^2 V_k + M_k = \sigma_W^2, \\ A_k V_k = \sigma_W^2 e^{-\lambda \Delta t}, \\ V_k = \sigma_W^2. \end{cases}$$

Получаем, что  $V_k = \sigma_W^2, A_k = e^{-\lambda \Delta t}, M_k = \sigma_W^2 (1 - e^{-2\lambda \Delta t})$ . Обратим внимание, что когда мы в 9 задании вводили процесс Орнштейна-Уленбека, то считали, что  $\mathbb{D}x_k = \sigma_W^2$ , что согласуется с тем, что мы получили.

Следовательно, фильтр Калмана для нашей задачи будет иметь вид:

$$\begin{cases} \tilde{x}_{k+1|k} = e^{-\lambda \Delta t} \tilde{x}_{k|k}, \\ \tilde{x}_{k+1|k+1} = \tilde{x}_{k+1|k} + R_{k+1|k} (R_{k+1|k} + \sigma_n^2)^{-1} (y_{k+1} - \tilde{x}_{k+1|k}), \\ R_{k+1|k} = e^{-2\lambda \Delta t} R_{k|k} + \sigma_W^2 (1 - e^{-2\lambda \Delta t}), \\ R_{k+1|k+1} = R_{k+1|k} - R_{k+1|k} (R_{k+1|k} + \sigma_n^2)^{-1} R_{k+1|k}, \\ \tilde{x}_{0|0} = 0, \\ R_{0|0} = \sigma_W^2. \end{cases}$$

Обозначив  $h = R_{k+1|k}(R_{k+1|k} + \sigma_n^2)^{-1}$ , получим итоговую систему:

$$\begin{cases} \tilde{x}_{k+1|k} = e^{-\lambda\Delta t}\tilde{x}_{k|k}, \\ \tilde{x}_{k+1|k+1} = (1 - h)\tilde{x}_{k+1|k} + hy_{k+1}, \\ h = R_{k+1|k}(R_{k+1|k} + \sigma_n^2)^{-1}, \\ R_{k+1|k} = e^{-2\lambda\Delta t}R_{k|k} + \sigma_W^2(1 - e^{-2\lambda\Delta t}), \\ R_{k+1|k+1} = (1 - h)R_{k+1|k}, \\ \tilde{x}_{0|0} = 0, \\ R_{0|0} = \sigma_W^2. \end{cases}$$

В данном случае доверительный интервал

$$x_k + k_{1-\frac{\alpha}{2}}[-\sqrt{R_{k|k}}, \sqrt{R_{k|k}}]$$

где  $k_\alpha$  — квантиль стандартного нормального распределения.

### 10.3 Примеры работы программы

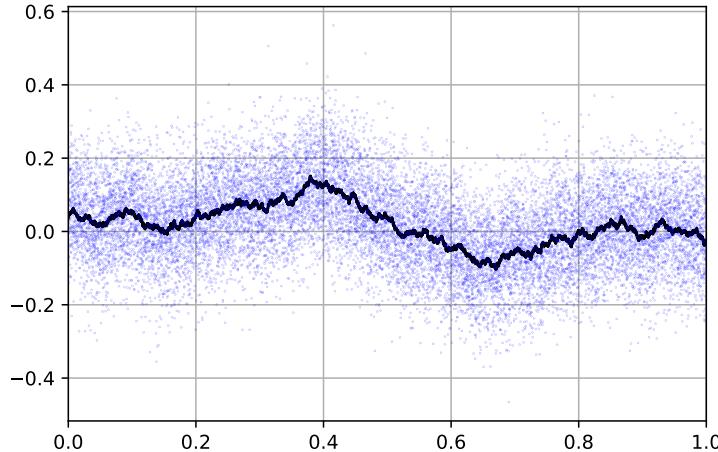


Рис. 28. Процесс Орнштейна-Уленбека с гауссовским шумом.



Рис. 29. Результат применения работы фильтра Калмана к процессу Орнштейна-Уленбека с гауссовским шумом .

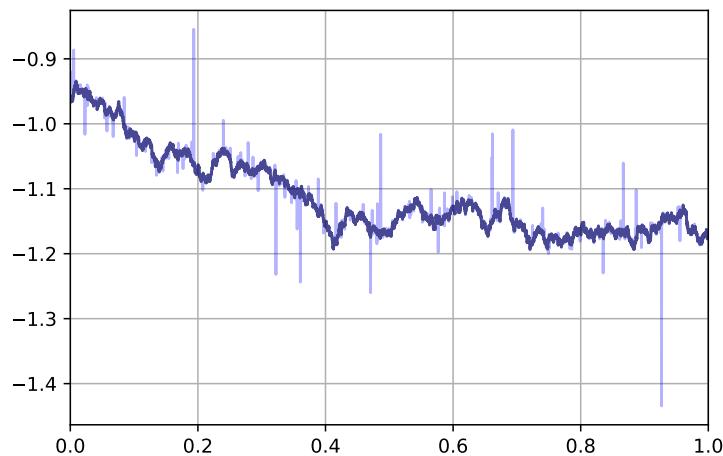


Рис. 30. Процесс Орнштейна-Уленбека с шумом, имеющим распределение Коши.



Рис. 31. Результат применения работы фильтра Калмана к процессу Орнштейна-Уленбека с шумом, имеющим распределение Коши.

## 11 Задание 11

### 11.1 Постановка задачи

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 ее обслуживания (распределение  $\chi^2$  с 10 степенями свободы).
2. Вторая интерпретация: система массового обслуживания с циклической интенсивностью  $\lambda(1 + \cos(t))$  и единичными скачками. Свести данную задачу моделирования неоднородного пуассоновского процесса при помощи метода Льюиса и Шеддера 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.

### 11.2 Решение

#### 11.2.1 Первая интерпретация: система массового обслуживания

Пусть  $\lambda$  — интенсивность пуассоновского поля. Времена поступления заявок генерируются так, что  $\Delta t_i = t_i - t_{i-1} \sim Exp(\lambda)$ .

**Определение 18.** Распределением  $\chi^2$  с  $k$  степенями свободы называется распределение суммы квадратов  $k$  независимых стандартных нормальных случайных величин.

Время обслуживания каждой заявки  $s_i$  независимы и генерируются как случайные величины с распределением  $\chi^2(10)$ .

Поскольку все заявки обрабатываются последовательно, время окончания обработки заявки, поступившей в момент времени  $t_i$  можно найти следующим образом:

- если к моменту поступления заявки предыдущая заявка уже обработана, то нужно к времени поступления текущей заявки прибавить время ее обработки:

$$Q_i = t_i + s_i.$$

- если предыдущая заявка еще не обработана, то нужно прибавить к времени конца обработки предыдущей заявки время обработки текущей:

$$Q_i = Q_{i-1} + s_i.$$

Обобщая вышесказанное, имеем:

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

Для каждой заявки будем считать количество людей в очереди.

- если во время поступления  $i$ -й заявки очереди не было то положим  $n_i = 0$ .
- если предыдущая заявка еще не обработана, то:

$$n_i \neq Q_k : k < i \text{ и } Q_k > t_i,$$

т.е. количество еще не выполненных к моменту времени  $t_i$  заявок.

Поскольку время обработки еще одной заявки в среднем равно 10, а средний интервал между поступлениями заявок равен  $\mathbb{E}\Delta_i = \frac{1}{\lambda}$ , то при  $\lambda < 0.1$  очереди практически не будет, а при  $\lambda > 0.1$  очередь будет неограниченно расти.

### 11.2.2 Вторая интерпретация: система массового обслуживания с циклической интенсивностью и единичными скачками

Пусть  $T_1, \dots, T_n, \dots$  — времена наступления некоторых событий, а  $N(t_1, t_2)$  — количество событий, произошедших в промежуток  $[t_1, t_2]$ . Заметим, что  $T_{n+1} - T_n$  имеет функцию распределения  $F(x) = 1 - \exp\{-\Lambda(t+x) - \Lambda(t)\}, x > 0$ , где

$$\Lambda(t) = \int_0^t \lambda(u)du = \lambda(t + \sin t)$$

неограниченно возрастает с ростом  $t$ .

$T_{n+1}$  распределено как  $T_n + F^{-1}(U)$ , где  $U$  равномерно распределена на  $[0, 1]$ . Заметим, что если записать  $U$  как  $1 - \exp\{-E\}$ , где  $E$  — экспоненциальная случайная величина с параметром  $\lambda_E = 1$ , то  $T_{n+1}$  распределена как  $\Lambda^{-1}(E + \Lambda(T_n))$ .

Будем искать обратную функцию  $\Lambda^{-1}(y)$  численно, так как аналитически это не представляется возможным ( $\Lambda'(t) = \lambda_0(1+\cos(t))$  почти всюду положительна, то есть функция возрастает). Такой метод моделирования неоднородного процесса Пуассона называется методом Льюиса-Шеддера.

Чтобы не искать обратную функцию, можно воспользоваться следующей модификацией метода Льюиса-Шеддера. Пусть имеется переменная  $t$ , в которой хранится текущее время (но не обязательно событие произошло строго в это время).

- На каждом шаге генерируем случайную величину  $\xi \sim Exp(2\lambda_0)$ .
- Прибавляем к переменной  $t$  величину  $\xi$  и генерируем случайную величину  $\eta = Bern(\frac{1+\cos t}{2})$ .
  - если она приняла значение 1, то полагаем  $T_{i+1} = t$  и  $i = i + 1$
  - иначе ничего не делаем и повторяем процесс заново

### 11.2.3 Третья интерпретация: работа страховой компании

**Определение 19.** Случайная величина  $X$  имеет распределение Парето с параметрами  $x_m$  и  $k$ , если ее функция распределения имеет вид:

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

Для моделирования случайной величины, имеющей распределение Парето, снова воспользуемся методом обратной функции.

Обратная функция для данной функции распределения имеет вид:

$$F_X^{-1}(x) = \frac{x_m}{(1-x)^{\frac{1}{k}}}.$$

Сгенерируем времена наступления страховых случаев на временном интервале  $[0, T]$ :

$$0 \leq t_1 \leq \dots \leq t_n < T,$$

причем  $t_i - t_{i-1} \sim Exp(\lambda)$ ,  $\lambda > 0$  — интенсивность потока страховых случаев.

Величину ущерба  $s_i$  страхового случая в момент времени  $t$  будем генерировать с помощью распределения Парето с параметрами  $x_m$  и  $k$ . Случайную величину, распределенную по Парето, будем генерировать, воспользовавшись методом обратных функций:

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

Учтем, что если  $Y \sim U[0, 1]$ , то и  $(1 - Y) \sim U[0, 1]$ . Тогда случайная величина

$$X = x_m Y^{-\frac{1}{k}}, \quad Y \sim U[0, 1]$$

имеет распределение Парето с параметрами  $x_m$  и  $k$ .

Величина капитала компании в момент времени  $t$  выражается как

$$W_t = W_0 + ct - s(t),$$

где  $s(t)$  — сумма величин ущерба страховых случаев, произошедших в моменты времени  $t_i$  такие, что  $t_i \leq t$ . Время разорения — случайная величина, задаваемая следующим условием:

$$T = \min\{t > 0 | W_t < 0\}.$$

Выведем зависимость функции  $W(t)$  от параметров  $\lambda, x_m, k, W_0, c$ . Будем считать, что  $k > 1$ . Тогда

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

Таким образом:

- $c(k-1) > \lambda k x_m$  капитал растет.
- $c(k-1) = \lambda k x_m$  система находится в положении равновесия.
- $c(k-1) < \lambda k x_m$  капитал уменьшается.

### 11.3 Примеры работы программы

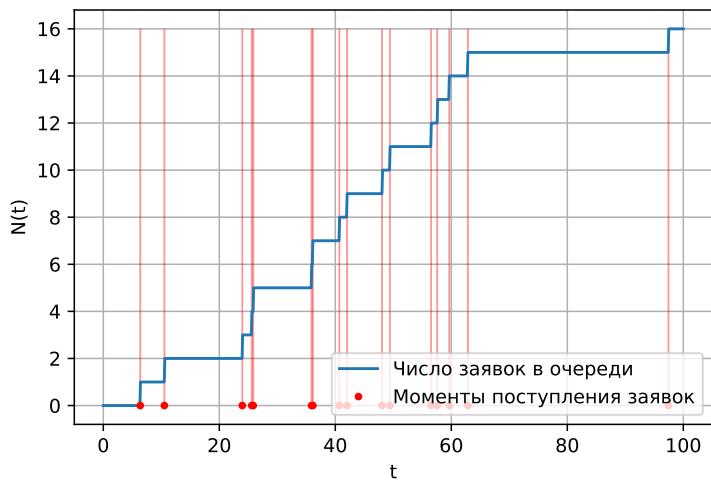


Рис. 32. Моделирование СМО.

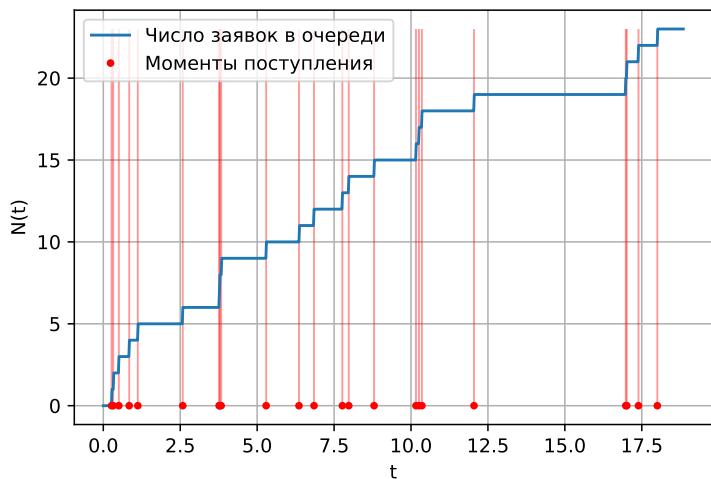


Рис. 33. Моделирование СМО с циклической активностью.

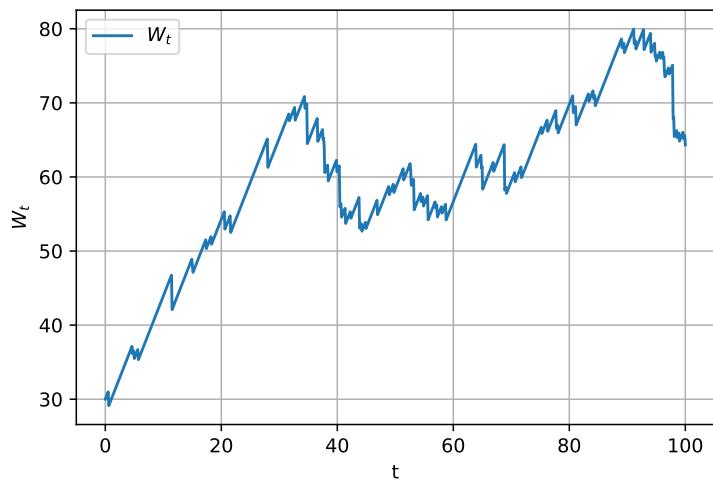


Рис. 34. Моделирование работы страховой компании.

## **12 Список литературы**

- [1] Смирнов С.Н. Лекции по курсу «Стохастический анализ моделирование», 2021.
- [2] Ширяев А.Н. Вероятность, Наука. М.:1989.
- [3] Кропачёва Н.Ю., Тихомиров А.С. Моделирование случайных величин: Метод указания, НовГУ им. Ярослава Мудрового, 2004
- [4] Феллер В. Введение в теорию вероятностей и её приложения, в 2—х томах. Т.1, М., Мир, 1984.