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

Отчёт по практикуму

«Стохастический анализ и моделирование»

Выполнила:
студентка 415 группы
Сафонова Елизавета

Преподаватель:
к.ф.-м.н.,
доцент кафедры СА
С. Н. Смирнов

Руководитель практикума:
В. А. Сливинский

Москва
2022

Содержание

1 Задание 1

1.1 Постановка задачи

1. Реализовать генератор схемы Бернулли с заданной вероятностью успеха p. Построить на его основе датчик биномиального распределения.

2. Реализовать генератор геометрического распределения. Проверить для него свойство отсутствия памяти.

3. Рассмотреть игру в орлянку — бесконечную последовательность независимых испытаний с бросанием правильной монеты. Выигрыш S_n определяется как сумма по n испытаниям значений 1 и -1 в зависимости от выпавшей стороны. Проиллюстрировать (в виде ломаной) поведение $Y(i) = \dfrac{S_i}{\sqrt{n}}$, как функцию номера испытания $i = 1, \ldots, n$ для отдельно взятой траектории. Дать теоретическую оценку $Y(n)$ при $n \to \infty$.

1.2 Теоретические выкладки

Определение 1. Схема Бернулли – модель, состоящая из последовательности n случайных величин $\xi_1, \xi_2, \ldots, \xi_n$, принимающих значения 0 и 1 и таких, что: случайные величины ξ_i независимы и одинаково распределены (на метаязыке: отсутствие взаимного влияния и воспроизводимость, соответственно).

$$\xi_i = \begin{cases} 1, & \text{в } i\text{-ом испытании успех} \\ 0, & \text{иначе} \end{cases}$$

Определение 2. Будем говорить, что случайная величина ξ имеет распределение Бернулли, если она принимает два значения 1 и 0 с вероятностями «успеха» p и «неуспеха» $q = 1 - p$. Обозначение: $\xi \sim \mathcal{B}er(p)$.

Для генерации случайной величины, имеющей распределение Бернулли с вероятностью успеха p, будем использовать следующий алгоритм:

- Моделируем равномерно распределенную на $[0, 1]$ случайную величину $\eta \sim \mathcal{U}(0, 1)$.
- Если $\eta \leqslant p$, то датчик возвращает $\xi = 1$, В противном случае $\xi = 0$.

Определение 3. Биномиальное распределение — распределение числа появления события, имеющего вероятность p в n независимых испытаниях (т.е. распределение количества успехов в схеме Бернулли).

Определение 4. Будем говорить, что $\eta \sim \mathcal{B}in(n, p)$, если

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

Для построения датчика $\eta \sim \mathcal{B}in(n, p)$, проведем серию из n испытаний Бернулли с заданной вероятностью p, после чего посчитаем общее число «успехов»:

$$\eta = \sum_{i=1}^{n} \xi_i, \ \xi_i \sim \mathcal{B}er(p)$$

Определение 5. Геометрическое распределение — распределение, при котором проводится серия испытаний Бернулли с вероятностями успеха p и неуспеха $q = 1 - p$ до тех пор, пока не произойдет событие (пока не будет получен первый успех).

Определение 6. Будем говорить, что $\xi \sim \mathcal{G}eom(p)$, если

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

Свойство 1. (отсутствия памяти (отсутствия последействия)):

$$\mathbb{P}(X \geqslant t + s | X \geqslant s) = \mathbb{P}(X \geqslant t) > 0,\ \forall t,\ s \in \{0,\ 1,\ 2, \dots\}.$$

Доказательство.

$$\mathbb{P}(X \geqslant s) = \mathbb{P}(X = s) + \mathbb{P}(X = s+1) + \ldots = \sum_{k=s}^{\infty} p(1-p)^k =$$
$$= p(1-p)^s \left(1 + (1-p) + (1-p)^2 + \ldots\right) = p(1-p)^s \sum_{m=0}^{\infty} (1-p)^m = (1-p)^s$$

$$\mathbb{P}(X \geqslant t + s | X \geqslant s) = \frac{\mathbb{P}(X \geqslant t+s, X \geqslant s)}{\mathbb{P}(X \geqslant s)} = \frac{\mathbb{P}(X \geqslant t+s)}{\mathbb{P}(X \geqslant s)} =$$
$$= \frac{(1-p)^{t+s}}{(1-p)^s} = (1-p)^t = \mathbb{P}(X \geqslant t)$$

■

Рассмотрим игру в орлянку — бесконечную последовательность независимых испытаний с бросанием правильной монеты.

Пусть $\xi_1, \xi_2, \ldots$ – последовательность независимых одинаково распределенных случайных величин, каждая из которых принимает значение 1, если выпал орел (успех), и -1, если выпала решка (неуспех).

$$\mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{2}$$

$$S_n = \sum_{i=1}^{\infty} \xi_i$$

$$\mathbb{E}\xi_i = -1 \cdot \mathbb{P}(\xi_i = -1) + 1 \cdot \mathbb{P}(\xi_i = 1) = -\frac{1}{2} + \frac{1}{2} = 0$$
$$\mathbb{V}ar\xi_i = \mathbb{E}{\xi_i}^2 - (\mathbb{E}\xi_i)^2 = 1 \cdot \mathbb{P}(\xi_i^2 = 1) = 1$$

По классической ЦПТ:

$$\frac{S_n - \mathbb{E}(\xi_i)n}{\sigma(\xi_i)\sqrt{n}} = \frac{S_n - \mathbb{E}(\xi_i)n}{\sqrt{\mathbb{V}ar\xi_i}\sqrt{n}} = \frac{S_n}{\sqrt{n}} = Y(n) \xrightarrow[n\to\infty]{d} \nu \sim \mathcal{N}(0,1)$$

1.3 Результаты работы программы

1.3.1 Генератор распределения Бернулли

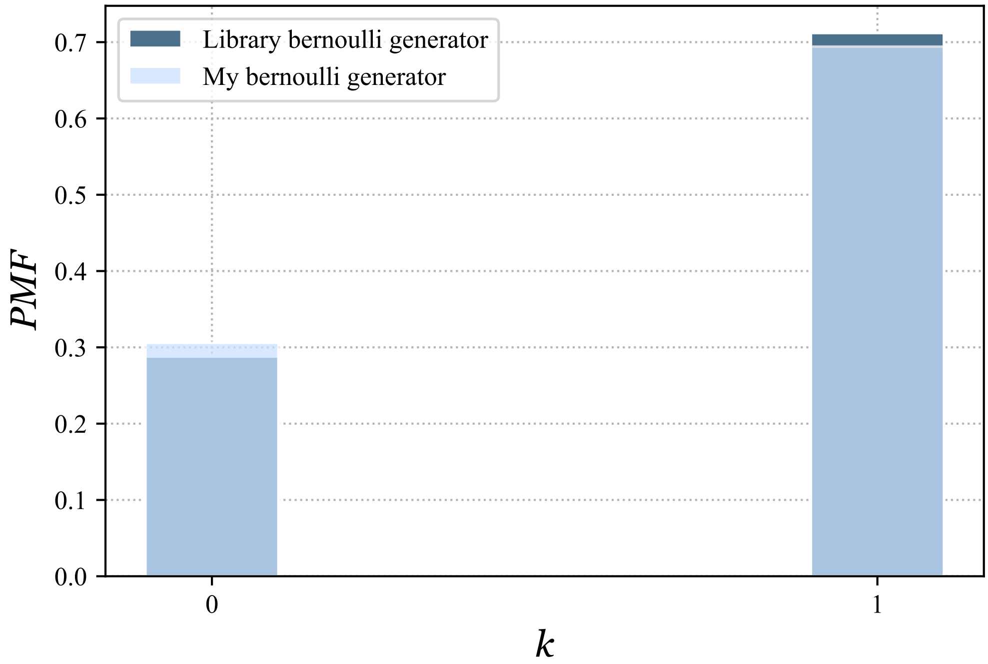

Рис. 1: Гистограмма распределения Бернулли $p = 0.7$

1.3.2 Генератор биномиального распределения

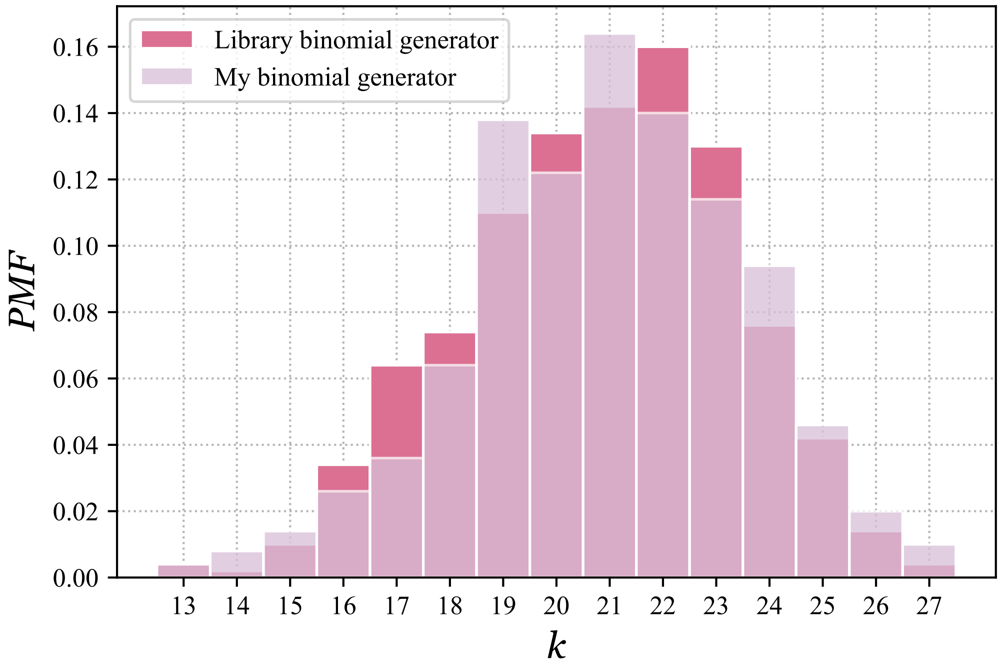

Рис. 2: Гистограмма биномиального распределения $p = 0.7$, $n = 30$

1.3.3 Генератор геометрического распределения

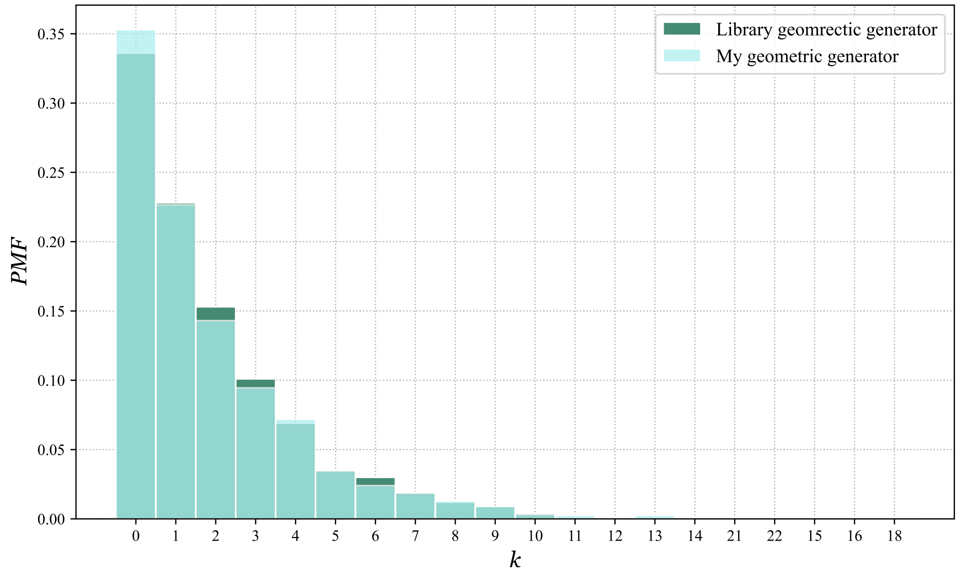

Рис. 3: Гистограмма геометрического распределения $p = 0.35$

1.3.4 Свойство отсутсвия памяти у геометрического распределения

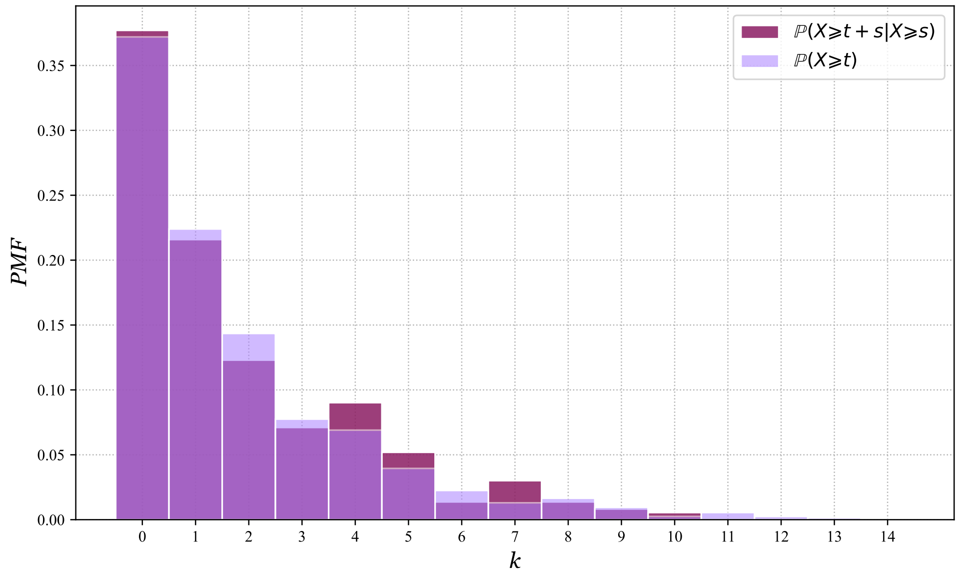

Рис. 4: Гистограмма геометрического распределения $p = 0.35$, $t = 4$

1.3.5 Пример траектории процесса при игре в орлянку

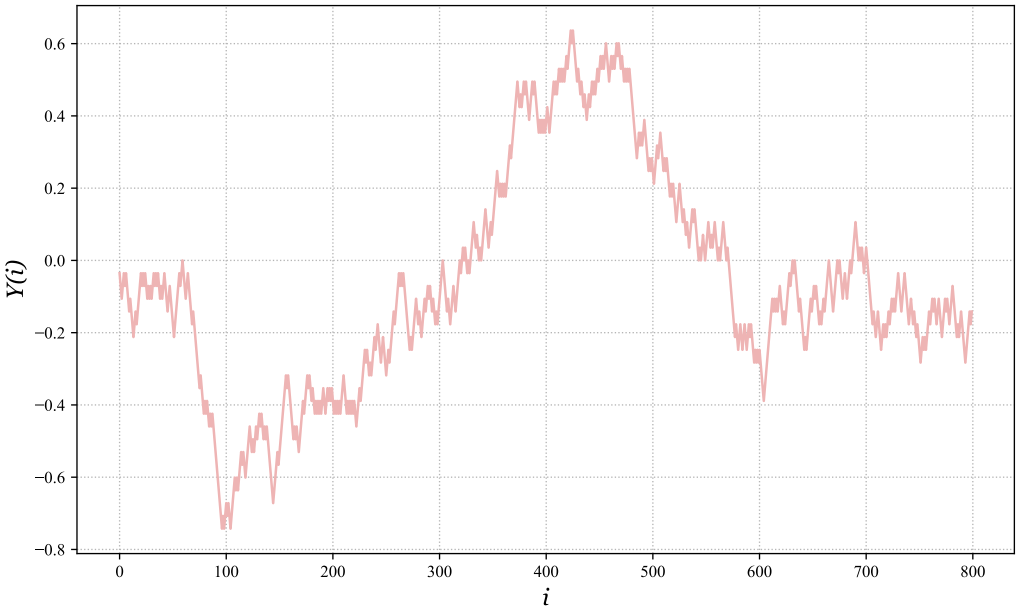

Рис. 5: Траектория $Y(i)$, при $p = 0.5$, $n = 800$

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 $Y \in [0, \frac{Y}{3}]$ совпадает с распределением $\frac{1}{3}$ с помощью критерия Смирнова.

3. Вычислить значение матожидания и дисперсии для данного распределения. Сравнить теоретические и эмпирические значения, проиллюстрировать сходимость.

2.2 Теоретические выкладки

Определение 7. Распределение называется сингулярным, если оно сосредоточено на континуальном множестве с нулевой мерой Лебега.

При построении Канторова множества $\mathcal{C}$ (мера Лебега которого равна нулю) на отрезке $[0, 1]$ мы «выбрасываем» интервалы $(\frac{1}{3}, \frac{2}{3})$, $(\frac{1}{9}, \frac{2}{9})$, $(\frac{7}{9}, \frac{8}{9}), \ldots$. То есть получается, что в троичной записи десятичной дроби будут отсутствовать единицы (Например, выбрасывая интервал $(\frac{1}{3}, \frac{2}{3})$, мы, тем самым, говорим, что на первой позиции числа после запятой $0, ? \ldots$ не может стоять единица).

Каждой точке $x = \frac{a_1}{3} + \frac{a_2}{3^2} + \frac{a_3}{3^3} + \ldots$, $x \in \mathcal{C}$ можно отнести последовательность $a_1, a_2, \ldots, a_n$, где $a_i \in \{0, 2\}$, которой, в свою очередь, соответствует последовательность вида $b_1,\ b_2, \ldots,\ b_n$, где $b_i \in \{0,\ 1\}$, генерируемая датчиком Бернулли с вероятностью $p = \frac{1}{2}$.

Пусть функция $K(x)$ – канторова лестница. В точках 0 и 1 ее значение принимается равным соответственно 0 и 1. Далее интервал $(0,\ 1)$ разбивается на три равные части $\left(0,\ \frac{1}{3}\right)$, $\left(\frac{1}{3},\ \frac{2}{3}\right)$ и $\left(\frac{2}{3},\ 1\right)$. На среднем сегменте полагаем $K(x) = \frac{1}{2}$. Оставшиеся два интервала снова разбиваются на три равные части каждый, и на средних из них $K(x)$ полагается равной $\frac{1}{4}$ и $\frac{3}{4}$.Каждый из оставшихся сегментов снова делится на три части, и на внутренних интервалах $K(x)$ определяется как постоянная, равная среднему арифметическому между соседними, уже определенными значениями $K(x)$. На остальных точках единичного отрезка определяется по непрерывности.

$$x = \sum_{i=1}^{n} \frac{2b_i}{3^i}, \quad K(x) = \sum_{i=1}^{n} \frac{b_i}{2^i}$$

Будем рассматривать частичные суммы. Для этого этого введем погрешность ε и найдем такое число n, при котором частичная сумма будет отличаться от бесконечной не более, чем на заданную погрешность.

$$\sum_{i=n+1}^{\infty} \frac{2b_i}{3^i} \leqslant 2 \sum_{i=n+1}^{\infty} \frac{1}{3^i} = \frac{1}{3^n} \leqslant \varepsilon$$

$$-\log_3 \varepsilon \leqslant n$$

Критерий Колмогорова. Пусть эмпирическая функция распределения $F_n(x)$, построенная по выборке $X = (X_1, \ldots, X_n)$, имеет вид:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{X_i \leq x},$$

где $\mathbb{I}_{X_i \leq x}$ - индикатор, указывающий попало ли наблюдение X_i в область $(-\infty, x]$.

Статистика критерия для эмпирической функции:

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

Способ проверки гипотезы — сравнение статистки $\sqrt{n} D_n$ с критическим значением K_α, то есть с верхним квантилем распределения, рассчитываемым как $K_\alpha = F_K^{-1}(1 - \alpha)$.

Этот подход требует обращения функции распределения, что аналитически невозможно. Поэтому иногда пользуются аппроксимацией функции распределения Колмогорова $F_K(x) \approx 1 - 2e^{-2x^2}$, которая легко обратима.

Гипотеза о соответствии выборки принимается с заданным уровнем значимости α при $\sqrt{n} D_n < K_\alpha$ и отвергается иначе.

$$F_K^{-1}(x) = \sqrt{-\frac{1}{2} \ln \frac{1-x}{2}}$$

$$K_\alpha = F_K^{-1}(1 - \alpha) = \sqrt{-\frac{1}{2} \ln \frac{\alpha}{2}}$$

Критерий однородности Смирнова используется для проверки гипотезы о принадлежности двух независимых выборок одному закону распределения, то есть говорит о том, что два эмпирических распределения соответствуют одному и тому же закону.

Критерий Смирнова. Обозначим за H_0 гипотезу о том, что две исследуемые выборки объемами n и m с эмпирическими функциями распределения $F_n(x)$ и $F_m(x)$ 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:

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

Для проверки свойства **симметричности** канторовых случайных величин рассмотрим $\xi_1 \sim \mathcal{C}antor(n)$ и $1 - \xi_2$, где $\xi_2 \sim \mathcal{C}antor(m)$. Из симметрии канторовой лестницы следует, что $F_\xi(x) + F_\xi(1-x) = 1$. Тогда имеем:

$$F_{\xi_2} = F_{1-\xi_1}(x) = P\left(1 - \xi_1 < x\right) = P\left(1 - x < \xi_1\right) =$$
$$= 1 - P\left(\xi_1 < 1 - x\right) = 1 - F_{\xi_1}(1-x) = F_{\xi_1}(x)$$

Для проверки свойства **самоподобия** канторовых случайных величин относительно деления на 3 рассмотрим $\xi_1 \sim \mathcal{C}antor(n),\ \xi_1 \in \left[0, \frac{1}{3}\right]$ и $\xi_2 = \frac{\xi_1}{3}$. Рассмотрим множество $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Имеем для ξ_1 :

$$F\left(x \mid \xi_1 \in \left[0, \frac{1}{3}\right]\right) = P\left(\xi_1 < x \mid \xi_1 \in \left[0, \frac{1}{3}\right]\right) = \frac{P\left(\xi_1 < x, \xi_1 \in \left[0, \frac{1}{3}\right]\right)}{P\left(\xi_1 \in \left[0, \frac{1}{3}\right]\right)} =$$
$$= \left\{P\left(\xi_1 \in \left[0, \frac{1}{3}\right]\right) = \frac{1}{2}\right\} = 2P\left(\xi_1 < x, \xi_1 \in \left[0, \frac{1}{3}\right]\right) = 2F_{\xi_1}(x),\ \text{при } x \in \left[0, \frac{1}{3}\right]$$

Для $\xi_2 = \frac{\xi_1}{3}$ получим:

$$F_{\xi_2} = F_{\frac{\xi_1}{3}} = P\left(\frac{\xi_1}{3} < x\right) = P\left(\xi_1 < 3x\right) = F_{\xi_1}(3x)$$

Вычислим значения математического ожидания и дисперсии для случайной величины $\xi \sim \mathcal{B}er(p)$:

$$\mathbb{E}\xi = \mathbb{P}(\xi = 1) \cdot 1 + \mathbb{P}(\xi = 0) \cdot 0 = p + (1-p) \cdot 0 = p$$
$$\mathbb{V}ar\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2 = p - p^2 = p(1-p)$$

Математическое ожидание и дисперсия для $X_n \sim \mathcal{C}antor(n)$:

$$\mathbb{E}X_n = \mathbb{E}\sum_{i=1}^{\infty} \frac{2b_i}{3^i} = \sum_{i=1}^{\infty} \frac{2}{3^i}\mathbb{E}b_i = \sum_{i=1}^{\infty} \frac{2}{3^i} \cdot \frac{1}{2} = \frac{\frac{1}{3}}{1 - \frac{1}{3}} = \frac{1}{2}$$

$X_1, \ldots, X_n$ независимы:

$$\mathbb{V}arX_n = \mathbb{V}ar\sum_{i=1}^{\infty} \frac{2b_i}{3^i} = \sum_{i=1}^{\infty} \left(\frac{2}{3^i}\right)^2 \mathbb{V}arb_i = \sum_{k=i}^{\infty} \frac{4}{9^i} \cdot \frac{1}{4} = \frac{\frac{1}{9}}{1 - \frac{1}{9}} = \frac{1}{8}$$

2.3 Результаты работы программы

2.3.1 Генератор сингулярного распределения

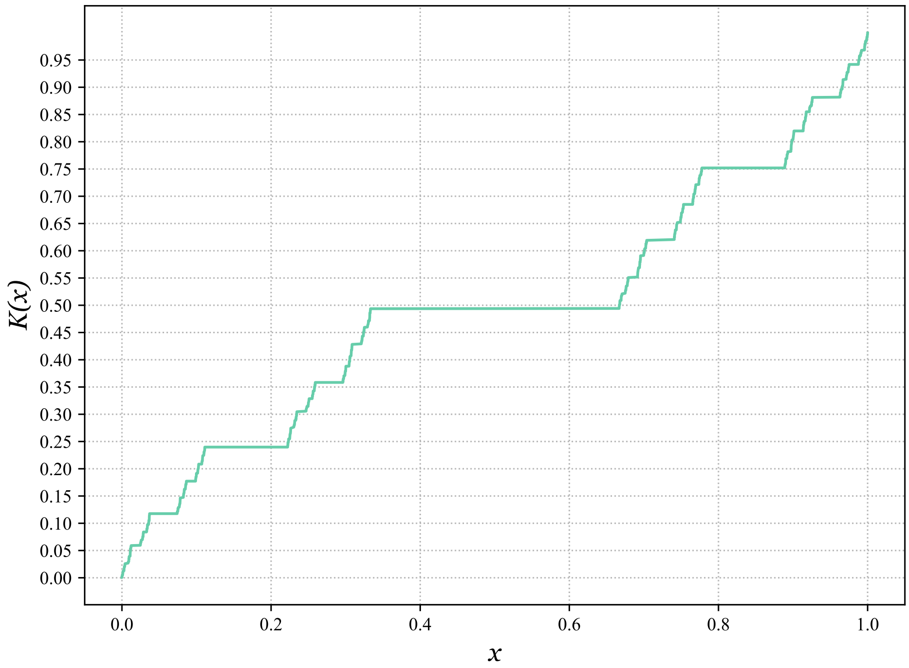

Рис. 6: Работа датчика сингулярного распределения, при $\varepsilon = 0.00001$, $n = 3000$

2.3.2 Проверка гипотез

Проверка построенного датчика с помощью критерия Колмогорова с уровнем значимости $\alpha = 0.003$ и размером выборки $n = 1000$ показала, что гипотеза принимается.

При проверке симметричности и самоподобия с помощью критерия Смирнова с теми же параметрами и $m = 1000$, гипотеза так же принимается.

2.3.3 Оценка моментов

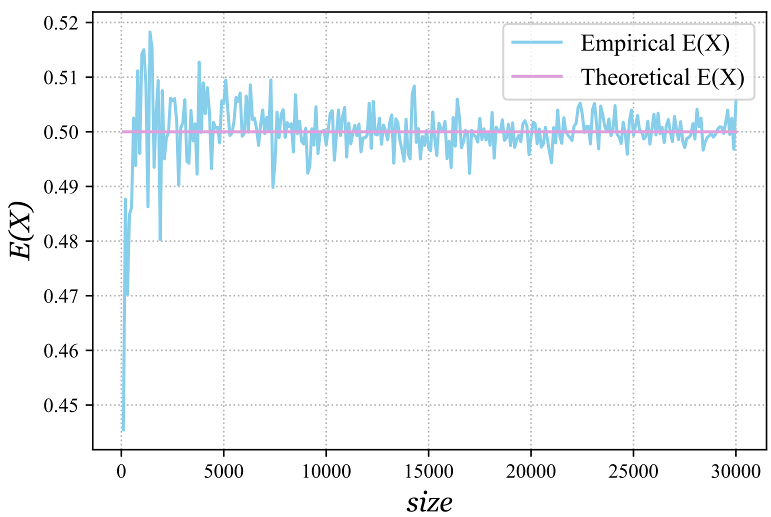

Рис. 7: Сходимость математического ожидания

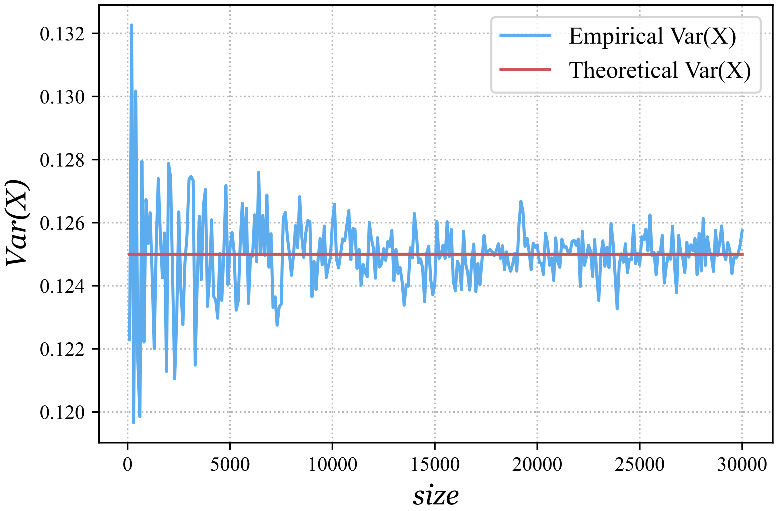

Рис. 8: Сходимость дисперсии

3 Задание 3

3.1 Постановка задачи

1. Построить датчик экспоненциального распределения. Проверить для данного распределения свойство отсутствия памяти. Пусть $X_1, X_2, \ldots, X_n$ – независимо экспоненциально распределенные с.в. с параметрами $\lambda_1, \lambda_2, \ldots, \lambda_n$ соответственно. Найти распределение случайной величины $Y = \min(X_1, X_2, \ldots, X_n)$.

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 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 Теоретические выкладки

Рассмотрим экспоненциальное распределение:

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

Теорема 1. Пусть некоторая функция распределения F имеет обратную F^{-1}. Тогда F является функцией распределения случайной величины $\eta = F^{-1}(\xi)$, где $\xi \sim \mathcal{U}[0,\ 1]$

Обращаем функцию распределения: $\tilde{F}(y) = -\frac{1}{\lambda}\ln(1-y)$. Замечаем, что если $\xi \sim U[0,1]$, то $\eta = 1-\xi$ имеет такое же распределение, что и ξ. Поэтому можем положить $x = -\frac{1}{\lambda}\ln y$, где $y-$ значение равномерно распределенной на $[0,1]$ случайной величины.

Свойство 2. Случайная величина $X \sim \mathcal{E}xp(\lambda)$ обладает свойством отсутствия памяти.

Доказательство.

$$\mathbb{P}(X \geqslant t+s | X \geqslant s) = \mathbb{P}(X \geqslant t)$$

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

■

Рассмотрим $X_1, X_2, \ldots, X_n-$ независимо экспоненциально распределенные с.в. с параметрами $\lambda_1, \lambda_2, \ldots, \lambda_n$ соответственно. Найдем распределение случайной величины

$$Y = \min\left(X_1, X_2, \ldots, X_n\right)$$

$$\begin{gathered} F_Y(x) = P(Y < x) = P\left(\min\left(X_1, \ldots, X_n\right) < x\right) = 1 - P\left(\min\left(X_1, \ldots, X_n\right) \geq x\right) = \\ = 1 - P\left(X_1 \geq x, \ldots, X_n \geq x\right) = 1 - P\left(X_1 \geq x\right) \cdot \ldots \cdot P\left(X_n \geq x\right) = \\ = 1 - \left(1 - P\left(X_1 < x\right)\right) \cdot \ldots \cdot \left(1 - P\left(X_n < x\right)\right) = 1 - e^{(\lambda_1 + \cdots + \lambda_n)x} \end{gathered}$$

Значит, случайная величина Y имеет экспоненциальное распределение с параметром $\lambda = \lambda_1 + \cdots + \lambda_n$:

$$Y \sim \mathcal{E}xp\left(\sum_{i=1}^{n} \lambda_i\right)$$

Определение 8. $\xi \sim \mathcal{P}ois(\lambda)$, если $P(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k \in \mathbb{N}_0$.

Теорема 2. Пусть $X_1, \ldots X_n-$ независимые с.в. $\sim \mathcal{E}xp(\lambda)$. Тогда можно построить с.в. $Y \sim \mathcal{P}ois(\lambda)$: $Y = \max_n\left((S_n = X_1 + \cdots + X_n) < 1\right)$, причем при $X_1 \geqslant 1$: $Y = 0$.

Таким образом, для генерации Пуассоновской случайной величины с параметром λ будем разыгрывать экспоненциальные случайные величины $X_i \sim \mathcal{E}xp(\lambda)$ до тех пор, пока их сумма S_n не превысит 1. А когда это произойдет, примем за значение $Y = n-1$, $Y \sim \mathcal{P}ois(\lambda)$.

Другой способ построения пуассоновской с.в. заключается в предельном свойстве биномиального распределения:

Биномиальное распределение сходится к распределению Пуассона при $n \to \infty$ 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 $\mathcal{B}in(n,p)$ биномиального распределения, если n достаточно велико, а p достаточно мало.

Пусть случайная величина $\xi \sim \mathcal{B}in(n,p)$:

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

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

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

Проверим с помощью критерия χ^2-Пирсона гипотезу о распределении по закону Пуассона.

Теорема 3 (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 — количество различных элементов выборки, 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 хи-квадрат распределение с $r-1$ степенью свободы.

Определение 9. $\xi \sim \mathcal{N}(0,1)$, если ее функция плотности вероятности задается формулой:

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

Построим генератор стандартной нормальной случайной величины, воспользовавшись методом Бокса-Мюллера:

$$x = r\cos\theta, \quad y = r\sin\theta$$

Для получения r и θ нужно сгенерировать две равномерно распределенные на отрезке (0, 1) случайные величины (назовем их u и v), распределение одной из которых (допустим v) необходимо преобразовать в экспоненциальное для получения радиуса. Функция экспоненциального распределения выглядит следующим образом:

$$f(v) = 1 - e^{-\lambda v}$$

Обратная к ней:

$$f(v) = -\frac{\ln(1-v)}{\lambda}$$

Так как равномерное распределение симметрично, то:

$$f(v) = -\frac{\ln v}{\lambda}$$

Из формулы распределения хи-квадрат следует, что $\lambda = 0.5$. Подставим в эту функцию λ, v и получим:

$$r^2 = -2\ln v, \quad r = \sqrt{-2\ln v}$$
$$\theta = 2\pi u$$

$$x = \sqrt{-2\ln v}\cos 2\pi u, \quad \sqrt{-2\ln v}\sin 2\pi u$$

При помощи критерия t**-критерия Стьюдента** проверим равенство математических ожиданий:

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}},$$

В нашем случае $\mu_0 = 0$. Несмещенная оценка дисперсии вычисляется следующим образом:

$$s^2 = \frac{\sum\limits_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}$$

Степень свободы равна $(n-1)$.

Критерий Фишера применяется для проверки равенства дисперсий двух выборок объемом m и n соответственно случайных величин X и Y с нормальным распределением. Статистика Фишера вычисляется следующим образом:

$$F = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2},$$

где $\hat{\sigma}_n^2$– выборочная дисперсия, вычисляемая по формуле: $\hat{\sigma}_n^2 = \frac{\sum\limits_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}$ Степени свободы определяются длиной выборок: $m-1$ и $n-1$.

3.3 Результаты работы программы

3.3.1 Датчик экспоненциального распределения

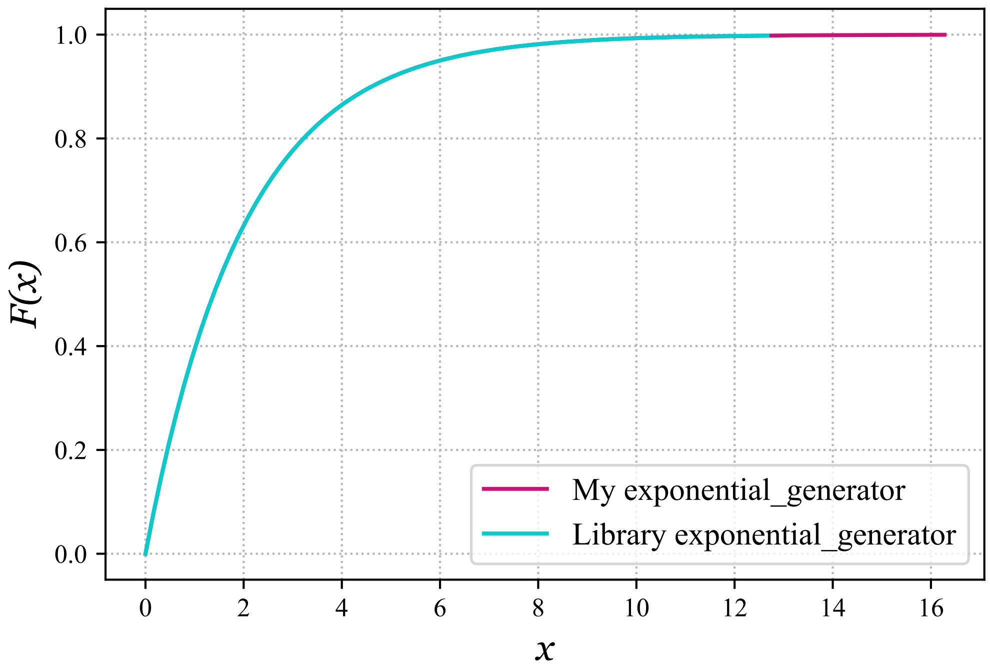

Рис. 9: Гистограмма экспоненциального распределения $\lambda = 0.5$

3.3.2 Свойство отсутсвия памяти у экспоненциального распределения

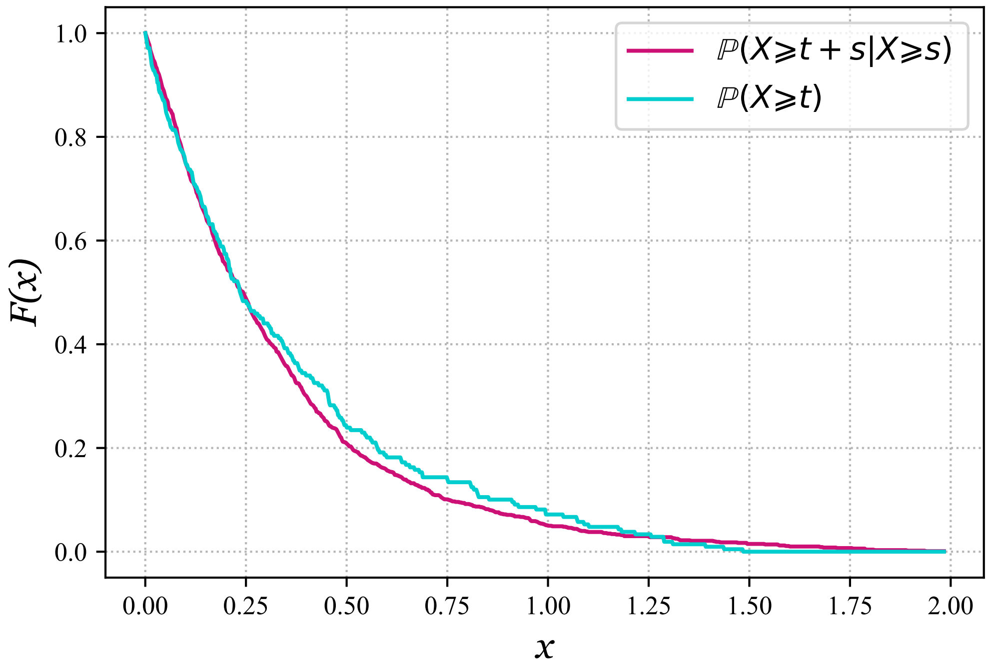

Рис. 10: График экспоненциального распределения $\lambda = 3,\ t = 0.5$

3.3.3 Датчик пуассоновского распределения

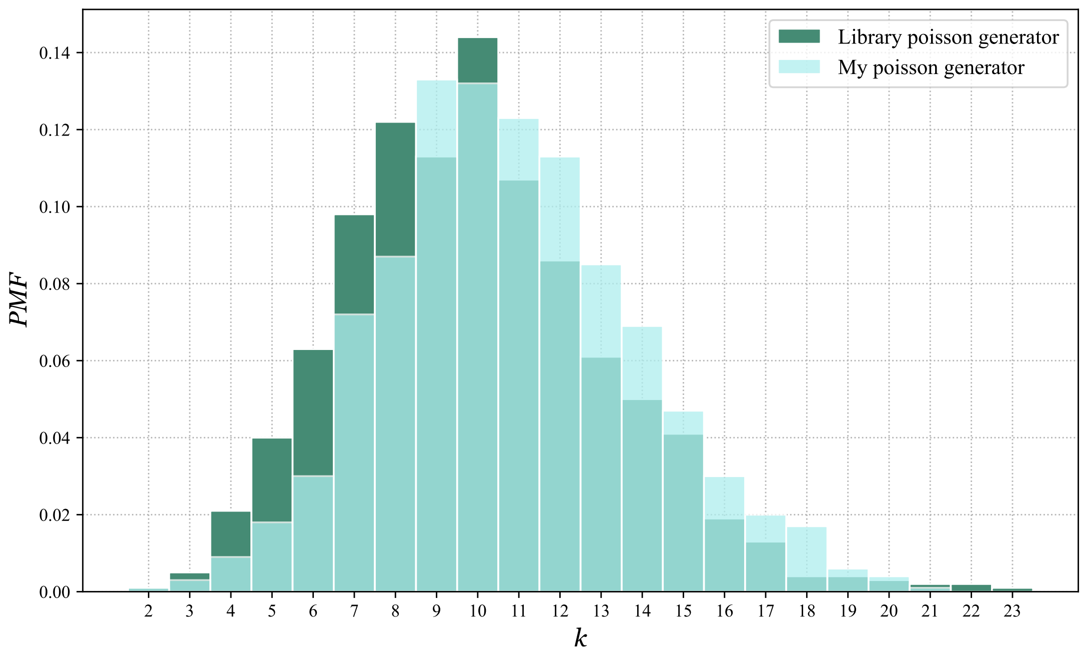

Рис. 11: Гистограмма пуассоновского распределения $\lambda = 10$

3.3.4 Датчик пуассоновского распределения как предел биномиального распределения

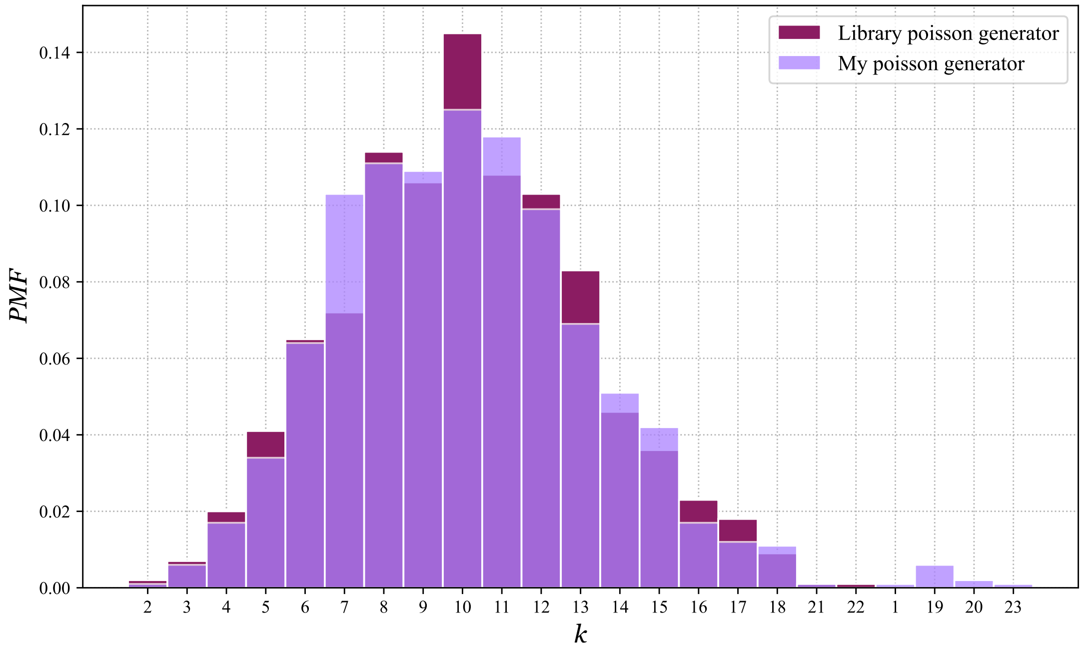

Рис. 12: Гистограмма пуассоновского распределения $\lambda = 10$

3.3.5 Датчик стандартного нормального распределения

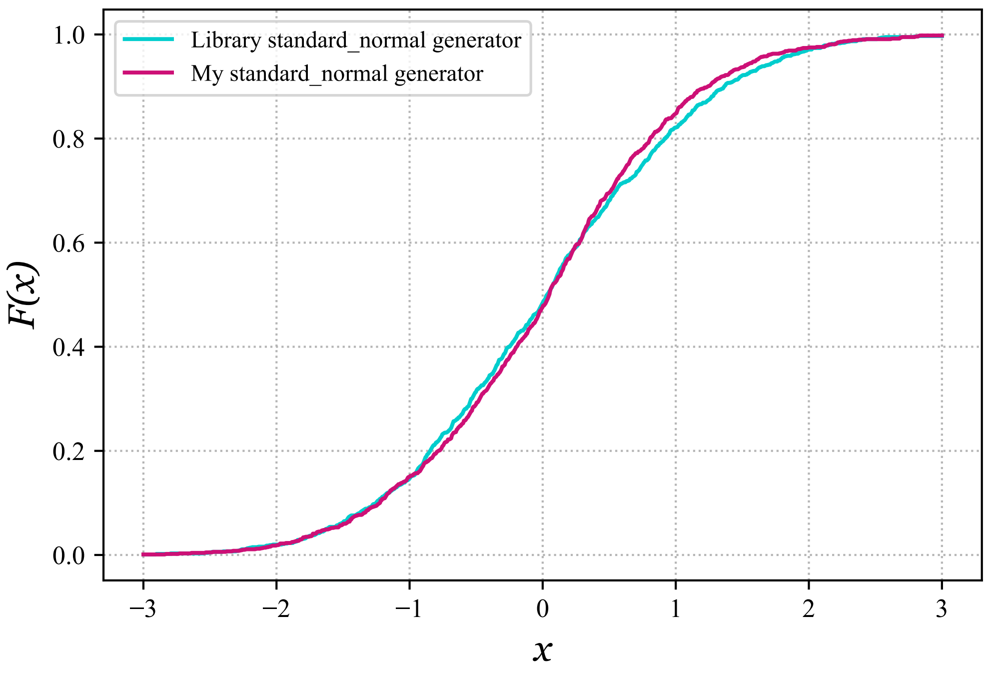

Рис. 13: График стандартного нормального распределения

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` убедиться в корректности построенного датчика и обосновать наблюдаемую линейную зависимость.

3. Сравнить скорость моделирования стандартного нормального распределения в заданиях 3 и 4.

4.2 Теоретические выкладки

Определение 10. Будем говорить, что случайная величина X имеет распределение Коши $X \sim Cauchy(a, b)$ с параметрами a и b, если ее плотность распределения задается следующей формулой:

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

где a — параметр сдвига, b — параметр масштаба.

Функция распределения Коши имеет вид:

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

Она строго возрастает и имеет обратную функцию:

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

Для описания метода фон Неймана нам потребуется вспомнить некоторые определения и утверждения. Рассмотрим измеримое пространство $(\Omega, \mathcal{F})$.

Определение 11. Пусть μ и ν - счетно-аддитивные меры, заданные на общей σ-алгере $\mathcal{F}$ подмножеста из Ω. Тогда мера μ называется абсолютно-непрерывной относительно меры ν ($\mu << \nu$), если:

$$\nu(A) = 0 \Longrightarrow \mu(A) = 0, \quad \forall A \in F$$

Теорема 4 (Радона-Никодима)**.** Пусть μ и ν — некоторые σ-конечные меры, определенные на σ-алгебре $\mathcal{F}$ подмножеств из Ω. Тогда:

$$\mu << \nu \Longleftrightarrow \exists \text{ функция } f : \mu(A) = \int\limits_A f(\omega)\mu(d\omega) = \int\limits_A f d\nu, \quad \forall A \in \mathcal{F}$$

Определение 12. Пусть $\mu << \nu$. Тогда $\mu(A) = \iint\limits_A f\mu, \forall A \in \mathcal{F}$. Функция f называется производной Радона-Никодима от меры μ по мере $\nu : f = \frac{d\mu}{dv}$.

Теперь перейдем к самому методу фон Неймана. У нас есть генератор распределения Коши с параметрами a и b. Нужно смоделировать стандартное нормальное распределение. Пусть также имеется некоторое пространство $(\mathcal{X}, \mathcal{A})$, на котором заданы вероятностные меры P и Q для стандартного нормального распределения и распределения Коши, соответственно, такие что:

$$P(A) < k \cdot Q(A), \quad \forall A \in A, k \in \mathbb{R}.$$

В нашем случае оба распределения (Коши и стандартное нормальное) абсолютно-непрерывны, т.е. они абсолютно-непрерывны относительно меры Лебега, т.е. существует производная Радона-Никодима $\frac{dP}{dQ}$, а это значит, что должно выполнятся неравенство $\frac{dP}{dQ} < k$, где dP и $dQ-$ плотности распределения соответствующих распределений.

Известно, что чем меньне константа k, тем эффективнее будет работать метод фон Неймана (плотности будут приближаться друг к другу, а это, в свою очередь, значит, что вероятность попадания в нужную нам область будет выше). Поэтому нам нужно минимизировать это вещественное число, это мокно сделать через следующую задачу минимакса:

$$k_{\min} = \min_{a,b} \max_{x \in \mathbb{R}} \left(\frac{f_n(x)}{f_c(x, a, b)} \right),$$

где f_c — плотность распределения Коши, f_n — плотность стандартного нормального распределения. Пусть для начала $a = 0$:

$$k_{\min}^0 = \min_{b} \max_{x \in \mathbb{R}} \left(\frac{f_n(x)}{f_c(x, b)} \right)$$

$$\frac{f_n(x)}{f_c(x, b)} = \frac{\sqrt{\pi}}{\sqrt{2}} e^{-\frac{x^2}{2}} \frac{x^2 + b^2}{b} = Ce^{-\frac{x^2}{2}} \left(x^2 + b^2 \right), \quad C = \frac{\sqrt{\pi}}{\sqrt{2}b} \neq 0$$

Максимум по x данного отношения достигается в точках:

$$x_{\max} = \begin{cases} \pm\sqrt{2 - b^2}, & \text{при } |b| \leq \sqrt{2}, \\ 0 & \text{при } |b| > \sqrt{2} \end{cases}$$

Тогда $k_{\min}^0$ будет выражаться следу ющим образом:

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

По аналогии рассматривается случай $a \neq 0$:

$$k_{\min} = \min_{b} \max_{x \in \mathbb{R}} \left\{ \frac{\sqrt{\pi}}{\sqrt{2}b} e^{-\frac{x^2}{2}} \left((x - a)^2 + b^2 \right) \right\} > \cdots > k_{\min}^0$$

Т.е. $\forall x_0 \neq 0, k_{\min} > k_{\min}^0$. Таким образом, для минимизации константы k нужно взять нулевой параметр $a = 0$ для распределения Коши. Теперь вспомним про вычисленное значение $x_{\max}$. Подставим его в $k_{\min}^0$:

$$k_{\min}^0 = \begin{cases} \sqrt{\frac{2\pi}{e}}, & \text{при } |b| \leq \sqrt{2} \, (b_{\min} = 1) \\ \sqrt{\pi}, & \text{при } |b| > \sqrt{2} \left(b_{\min} = \sqrt{2} \right) \end{cases}$$

Значит, нужная нам минимальная константа $k = \sqrt{\frac{2\pi}{e}}$ достигается при $b = 1, \ a = 0$

Тогда сформулируем алгоритм генерации стандартной нормальной случайной величины через распределение Коши с параметрами $a = 0, \ b = 1$ по методу фон Неймана:

1. Генерируем независимые пары случайных величин (X_i, ν_i), где X_i имеет распределение Коши с параметрами $a = 0,\ b = 1$, а ν_i– распределение Бернулли с вероятностью успеха:
$$P(\nu_i, X_i = x) = \frac{\sqrt{\pi}}{\sqrt{2}k} e^{-\frac{x^2}{2}} \left(x^2 + 1\right) = \frac{(x^2+1)}{2} e^{-\frac{x^2+1}{2}}.$$

2. Исключаем (элиминируем) все пары, где $\nu_i \neq 1$. Последовательность из о ставшихся X_i будет иметь стандартное нормальное распределение:
$$P(X \in A \mid \nu = 1) = \frac{P(X \in A, \nu = 1)}{P(\nu = 1)} = \frac{\int\limits_A P(\nu = 1, X = x) dQ(x)}{\int\limits_\Omega P(\nu = 1, X = x) dQ(x)} =$$
$$= \frac{\int\limits_A \frac{1}{k}\frac{dP}{dQ}(x) dQ(x)}{\int\limits_\Omega \frac{1}{k}\frac{dP}{dQ}(x) dQ(x)} = \frac{\int\limits_A dP(x)}{\int\limits_\Omega dP(x)} = P(A)$$

Определение 13. Случайная величина Y имеет нормальное (гауссовское) распределение с параметрами μ и $\sigma > 0\, (Y \sim \mathcal{N}(\mu, \sigma^2))$, если существует случайная величина $X \sim \mathcal{N}(0, 1)$ со стандартным нормальным распределением такая, что:
$$Y = \mu + \sigma X,$$
μ — параметр сдвига, а $\sigma > 0$ — параметр масштаба:
$$\mathbb{E}Y = \mathbb{E}(\mu + \sigma X) = \mu + \sigma \mathbb{E}X = \mu$$
$$\mathbb{V}arY = E(Y - EY)^2 = \mathbb{E}(\mu + \sigma X - \mu)^2 = \mathbb{E}(\sigma X)^2 = \sigma^2 \mathbb{E}X^2 = \sigma^2.$$

Определение 14. Пусть дано парамет рическое семейство вероятностных распределений, характеризованных их функцией распределения (или плот ностью вероят ности) $f(x; a)$, где $a \in \mathbb{R},\ a > 0$ — фиксированный параметр. Тогда этот параметр называется коэффициентом масштаба, если имеет место представление:
$$f(x; a) = \frac{1}{a} f\left(\frac{x}{a}\right),$$
где $f(x)$ - фиксированная функция распределения или плотность вероятности.

Определение 15. Пусть дано параметрическое семейство вероятностных распределений, характеризованных их функцией распределения (или плотностью вероят ности) $f(x; b)$, где $b \in \mathbb{R}$ — фиксированный параметр. Тогда этот параметр называется коэффициентом сдвига, если имеет место представление:
$$f(x; b) = f(x - b),$$
где $f(x)$ - фиксированная функция распределения или плот ность вероятности.

Т.е. для нормального распределения выполнено следующее выражение:
$$F_Y(y) = P(Y < y) = P\left(X < \frac{y - \mu}{\sigma}\right) = F_X\left(\frac{y - \mu}{\sigma}\right) =$$
$$= \int_{-\infty}^{\frac{y-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \{t = \mu + \sigma x\} = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

Другими словами, μ сдвигает график относительно графика стандартной нормальной величины, а $\frac{1}{\sigma}$– угол наклона к оси абсцисс.

4.3 Результаты работы программы

4.3.1 Датчик распределения Коши

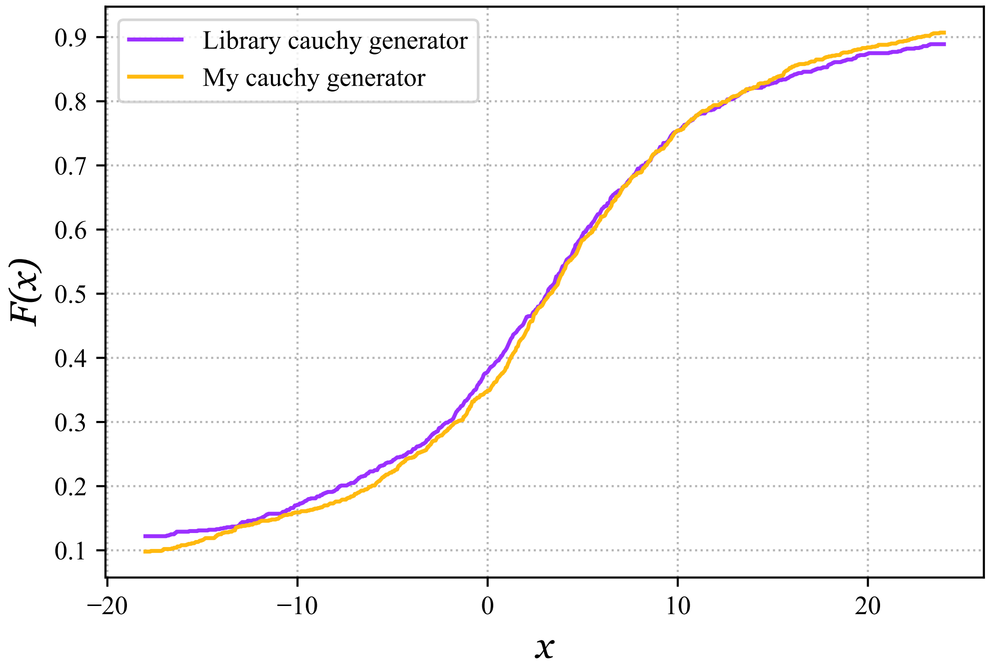

Рис. 14: График работы генератора Коши, $a = 3,\ b = 7$

4.3.2 Датчик стандартного нормального распределения по методу фон Неймана с помощью распределения Коши

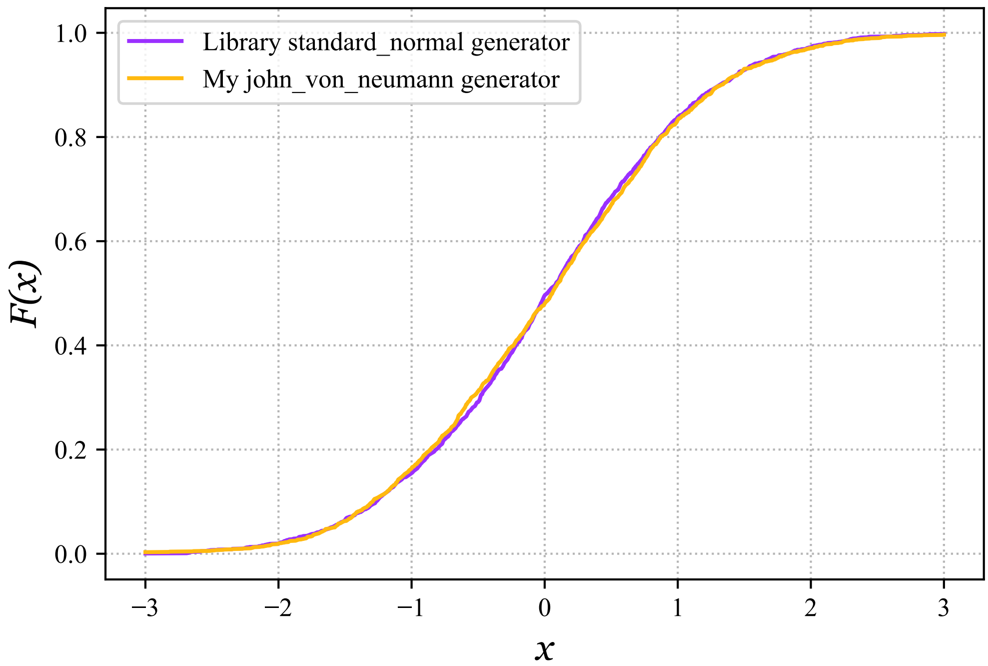

Рис. 15: График работы генератора

4.3.3 График стандартного нормального распределения

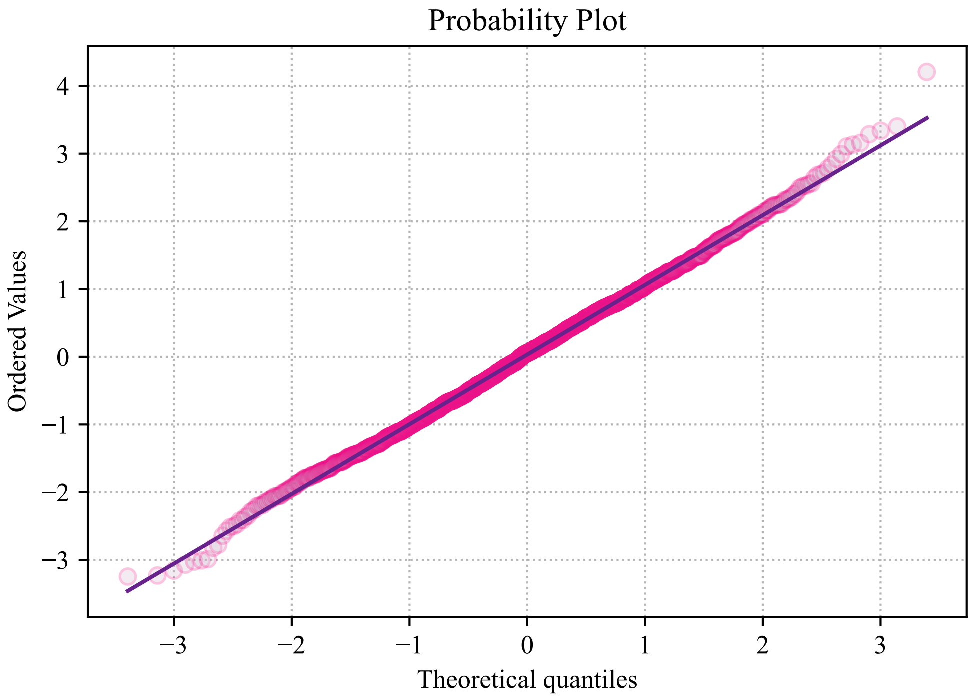

Рис. 16: Стандартное нормальное распределение, построенное по методу фон Неймана

4.3.4 Влияние параметра сдвига

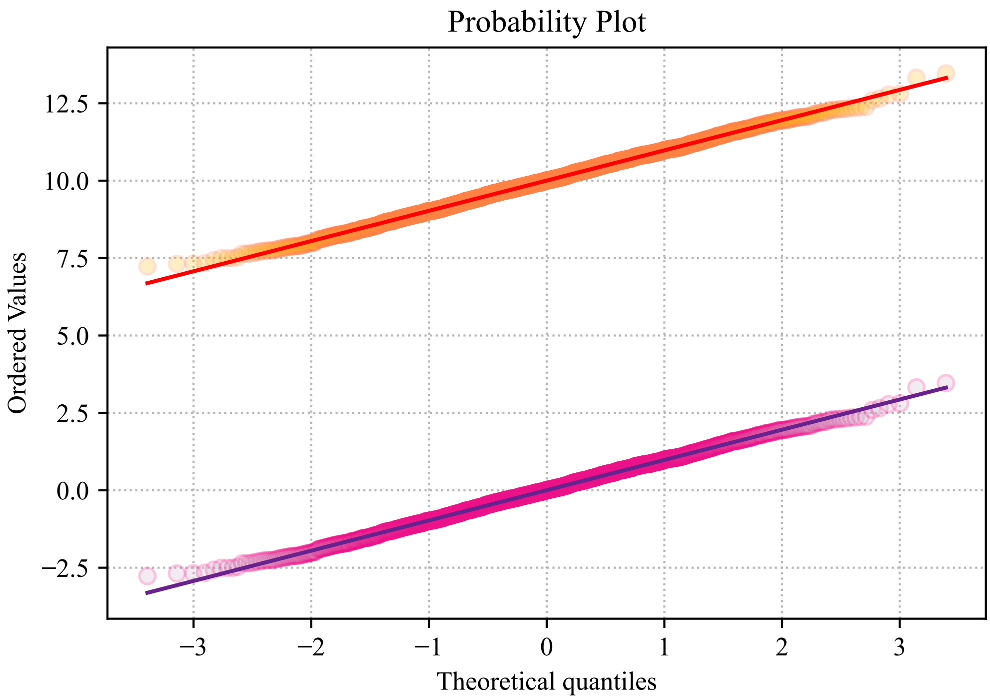

Рис. 17: $a = 10,\ b = 1$

4.3.5 Влияние параметра масштаба

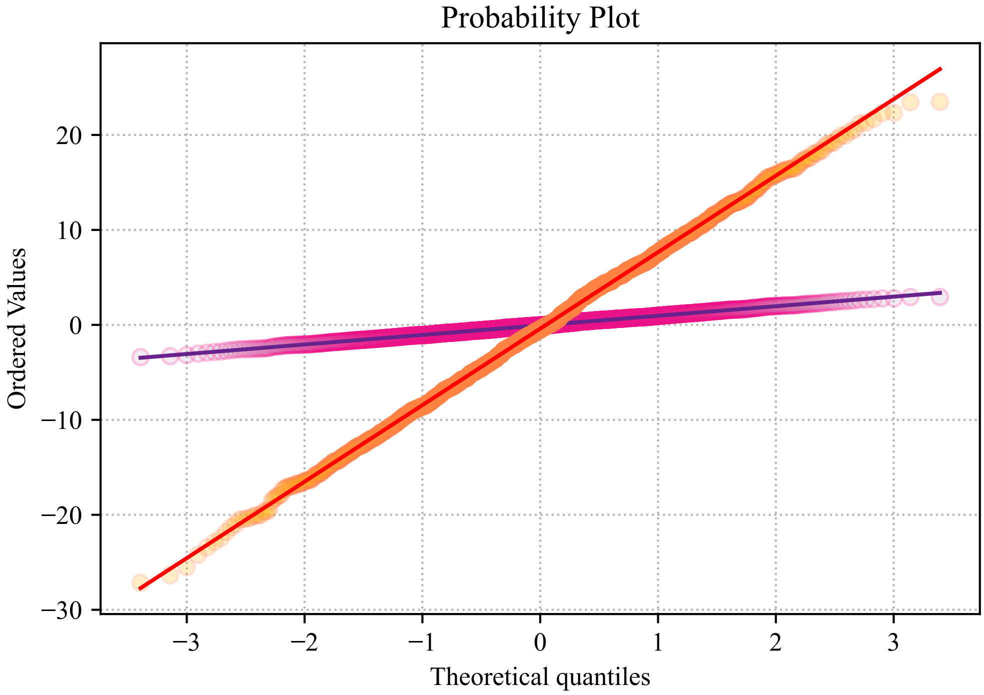

Рис. 18: $a = 0,\ \ b = 8$

4.3.6 Сравнение скорости сходимости

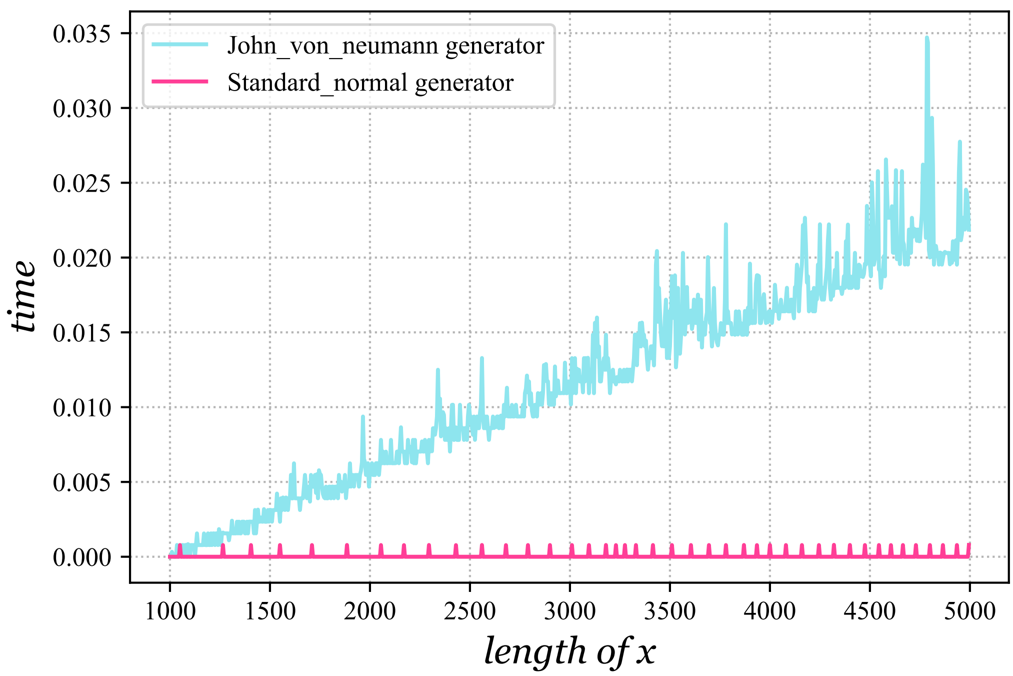

Рис. 19: Метод Бокса-Мюллера работает быстрее метода фон Неймана

5 Задание 5

5.1 Постановка задачи

1. Пусть $X_i \sim \mathcal{N}(\mu,\ \sigma^2)$. Убедиться эмпирически в справедливости ЗБЧ и ЦПТ, т.е. исследовать поведение суммы S_n и эмпирического распределения величины;
$$\sqrt{n}\left(\frac{S_n}{n} - a\right)$$
2. Считая μ и σ^2 неизвестными, для пункта 1 построить доверительные интервалы для среднего и дисперсии.
3. Пусть $X_i \sim \mathcal{K}(a,\ b)$ 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.

5.2 Теоретические выкладки

Закон больших чисел в форме Чебышёва:

Пусть $X_1,\ X_2, \ldots$ - последовательность независимых одинаково распределенных случайных величин с $\mathbb{E}X_k = a$ и $\mathbb{V}arX_k < \infty$. Тогда, если обозначить $S_n = X_1 + \ldots + X_n$, то выполнен ЗБЧ:
$$\frac{S_n - \mathbb{E}S_n}{n} \xrightarrow[n\to\infty]{\mathbb{P}} 0,$$
что равносильно:
$$\frac{S_n}{n} \xrightarrow[n\to\infty]{\mathbb{P}} a$$

Теорема 5 (ЦПТ)**.** Пусть $\xi_1, \xi_2, \ldots, \xi_n$ - последовательность случайных величин независимых в совокупности, одинаково распределенных так, что: $\mathbb{E}\xi_1^2 < \infty$, $\mathbb{V}ar\xi_1 \neq 0$. Обозначим $a = \mathbb{E}\xi_1, \sigma^2 = \mathbb{V}ar\xi_1$. Тогда справедливо выражение:
$$\frac{\xi_1 + \xi_2 + \cdots + \xi_n - na}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{d} \eta \sim \mathcal{N}(0, 1).$$

Определение 16. Пусть задано малое число $0 < \alpha < 1$. Интервал со случайными концами $(\theta_1; \theta_2)$ называется доверительным интервалом для параметра θ с уровнем доверия α, если для любого $\theta \in \Theta$
$$\mathbb{P}\left(\theta_1 < \theta < \theta_2\right) \geqslant \alpha$$

Пусть $t_{\alpha,\ n-1} - \alpha$-квантили распределения **Стьюдента**. Тогда :
$$\mathbb{P}\left(t_{\frac{1-\alpha}{2},\ n-1} \leq t \leq t_{\frac{1+\alpha}{2},\ n-1}\right) = \alpha.$$

После подстановки выражения для t-статистики, получим
$$\mathbb{P}\left(\bar{X} - t_{\frac{1+\alpha}{2},\ n-1}\frac{s}{\sqrt{n}} \leqslant \mu \leq \bar{X} + t_{\frac{1+\alpha}{2},\ n-1}\frac{s}{\sqrt{n}}\right) = \alpha.$$

Таким образом, мы получили доверительный интервал для параметра μ стандартного нормального распределения:
$$\bar{X} - t_{\frac{1+\alpha}{2},\ n-1}\frac{s}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} + t_{\frac{1+\alpha}{2},\ n-1}\frac{s}{\sqrt{n}},$$
s — несмещенное выборочное стандартное отклонение

Теорема 6 (Фишера для нормальных выборок)**.** случайная величина $H = \frac{(n-1)s^2}{\sigma^2}$ имеет распределение $\mathcal{X}^2_{n-1}$. Тогда:

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

После подстановки выражения для H, получим:

$$\mathbb{P}\left(\frac{(n-1)s^2}{\mathcal{X}^2_{\frac{1-\alpha}{2},n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\mathcal{X}^2_{\frac{1+\alpha}{2},n-1}}\right) = \alpha$$

Таким образом, доверительный интервал для параметра σ^2 стандартного нормального распределения:

$$\frac{(n-1)s^2}{\mathcal{X}^2_{\frac{1-\alpha}{2},n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\mathcal{X}^2_{\frac{1+\alpha}{2},n-1}}.$$

Рассмотрим последовательность $X_i \sim \mathcal{C}auchy(a, b)$. Для распределения Коши не существует математического ожидания (ни конечного, ни бесконечного):

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

По этой причине ЗБЧ для распределения Коши применять нельзя.

Найдем теперь как ведут себя суммы $\frac{S_n}{n}$. Для этого нам понадобится аппарат характеристических функций.

Определение 17. Пусть $\xi-$ случайная величина, а t — вещественное число. Тогда функция $\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \mathbb{E}(\cos(t\xi) + i\sin(t\xi))$ называется характеристической функцией случайной величины ξ.

Характеристическая функция обладает следующими свойствами, следующими из своего определения и свойств математического ожидания:

1. $\varphi_{a\xi+b}(t) = \mathbb{E}e^{it(a\xi+b)} = e^{itb}\mathbb{E}e^{ita\xi} = e^{itb}\varphi_\xi(at) = e^{itb}\phi_{a\xi}(t)$.

2. Если ξ и $\eta-$ независимые случайные величины, то:

$$\varphi_{\eta+\xi}(t) = \mathbb{E}e^{it(\eta+\xi)} = \mathbb{E}e^{it\eta}\mathbb{E}e^{it\xi} = \varphi_\eta(t)\varphi_\xi(t)$$

Теорема 7 (Единственности)**.** Если случайные величины ξ, таковы, что $\varphi_\xi \equiv \varphi_\eta$, то они распределены одинаково $F_\xi \equiv F_\eta$.

Характеристическая функция для распределения Коши $\xi \sim \mathcal{C}auchy(a, b)$ равна: $\varphi(t) = e^{ait-b|t|}$. Тогда найдем распределение сумм $\frac{S_n}{n} = \frac{X_1+\cdots+X_n}{n}$, где случайные величины $X_1, \ldots, X_n \sim C\left(x_0, \gamma\right)$:

$$\varphi_{\frac{s_n}{n}}(t) = \varphi_{S_n}\left(\frac{t}{n}\right) = \left(\varphi_{X_1}\left(\frac{t}{n}\right)\right)^n = \left(Ee^{\frac{ait-b|t|}{n}}\right)^n = \varphi_{X_1}(t).$$

Т.е. суммы распределены по Коши с теми же параметрами $\frac{S_n}{n} \sim \mathcal{C}auchy(a, b)$.

5.3 Результаты работы программы

5.3.1 Пример, иллюстрирующий справедливость ЗБЧ

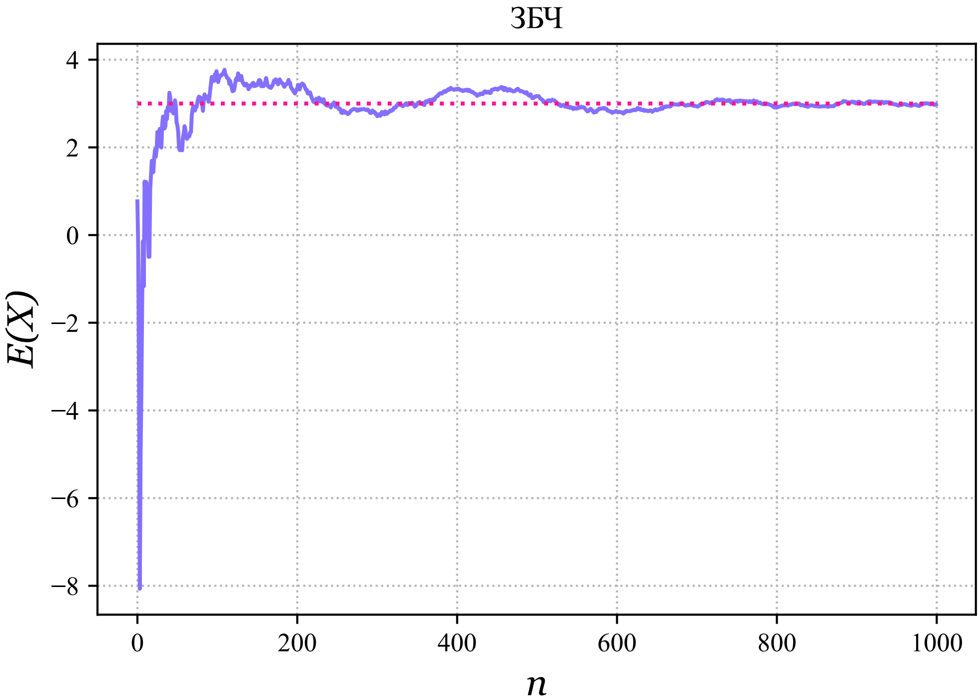

Рис. 20: Справедливость ЗБЧ для последовательности $\xi \sim \mathcal{N}(3, 9)$

5.3.2 Пример, иллюстрирующий справедливость ЦПТ

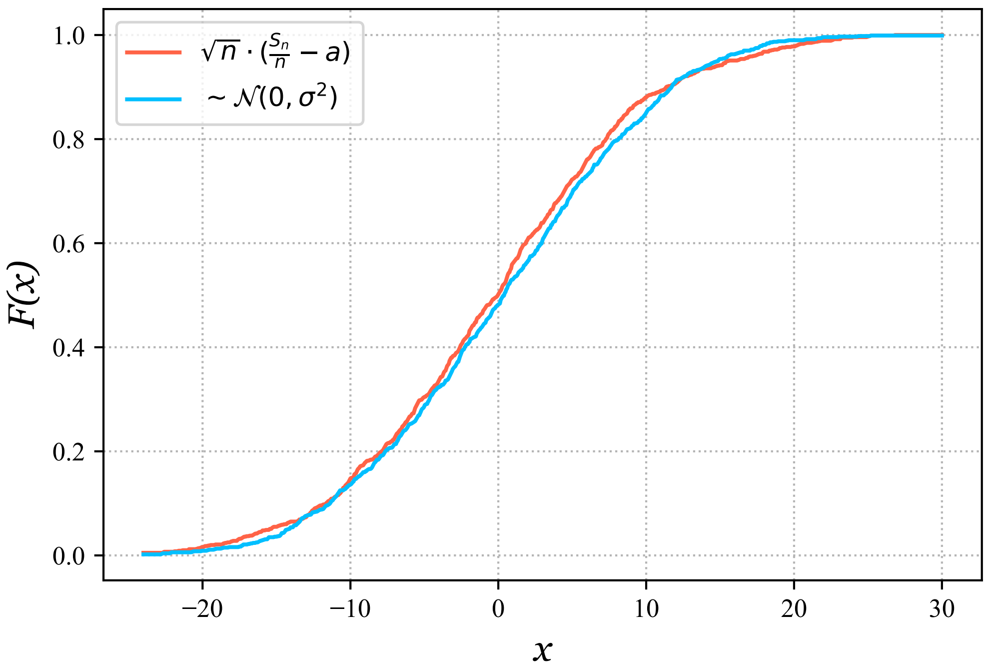

Рис. 21: Справедливость ЦПТ для последовательности $\xi \sim \mathcal{N}(3, 9)$

5.3.3 Построение доверительных интервалов

Получены результаты:

$$4.789385622834607 < \mu = 5 < 5.167903417314773$$

$$2.9429089865208087 < \sigma = 3 < 3.2127838036601357$$

5.3.4 Иллюстрация невыполнимости ЗБЧ для распределения Коши

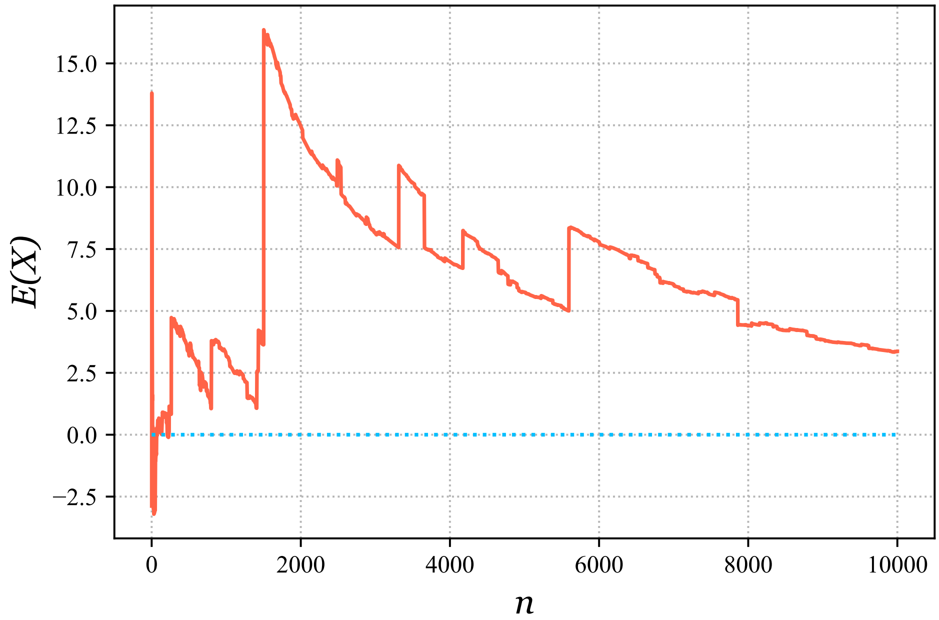

Рис. 22: ЗБЧ для последовательности $\xi \sim \mathfrak{C}auchy(0, 3)$

5.3.5 Свойство устойчивости распределения Коши

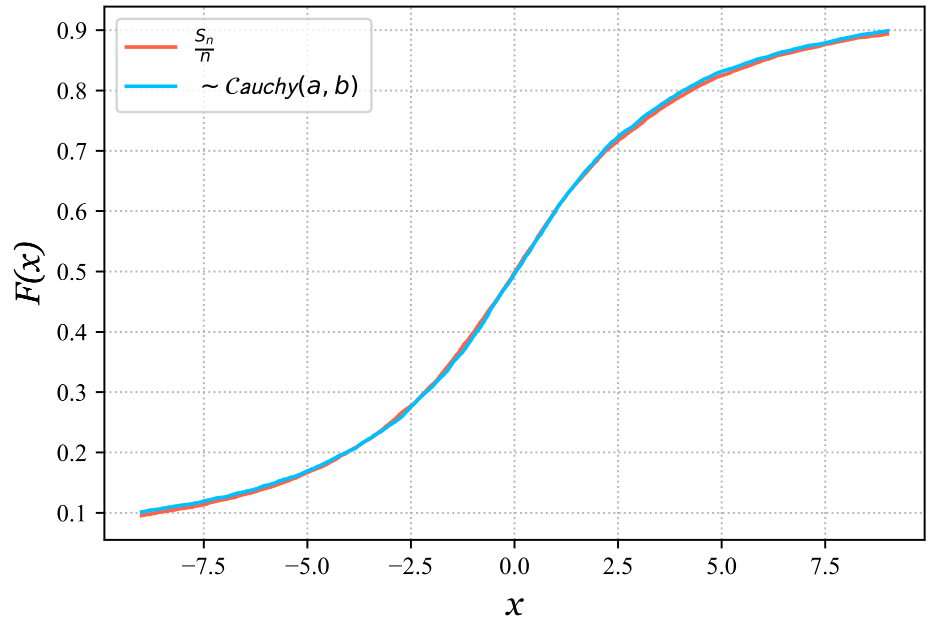

Рис. 23: $\xi \sim \mathfrak{C}auchy(0, 3)$

6 Задание 6

6.1 Постановка задачи

1. Посчитать интеграл

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\dots\int\limits_{-\infty}^{\infty}\frac{e^{-\left(x_1^2+\ldots+x_{10}^2+\frac{1}{2^7\cdot x_1^2\cdot\ldots\cdot x_{10}^2}\right)}}{x_1^2\cdot\ldots\cdot x_{10}^2}dx_1dx_2\dots dx_{10}$$

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

2. Для каждого случая оценить точность вычислений.

6.2 Теоретические выкладки

Метод Монте-Карло является единственным численным методом интегрирования по Лебегу, придуманный Джоном фон Нейманом, в качестве замены квадратурным формулам, подверженным «проклятию размерности».

Пусть у нас есть последовательность независимых случайных величин с некоторым распределением Q и со значениями из измеримого пространст ва $(E, \mathcal{E})$. Предположим, что нам необходимо посчитать интеграл:

$$I = \int\limits_E f(x)Q(dx)$$

где $f(x)$ — числовая измеримая функция, $f(x) \geqslant 0$. В качестве приближенного значения интеграла будем брать:

$$\tilde{I} = \frac{1}{n}\sum_{i=1}^{n} f\left(\xi_i\right)$$

По ЗБЧ значение $\tilde{I}$ стремится к $Ef(\xi)$, которое по определению равно:

$$\mathbb{E}f(\xi) = \int\limits_\Omega f(\xi(\omega))P(d\omega)$$

Вспомнив теорему о замене переменных под знаком интеграла, поймем, что планируем считать как раз то, что нужно:

$$\mathbb{E}f(\xi) = \int\limits_\Omega f(\xi(\omega))P(d\omega) = \int\limits_E f(x)Q(dx) = I$$

Теперь перейдем к применению данного метода к нашей задаче. Обозначим нашу подынтегральную функцию как $\tilde{F}(x) = \tilde{F}(x_1, \ \dots, x_{10})$. Представим ее как произведение $\tilde{F}(x) = \rho(x)g(x)$, где $\rho(x)$ — плотность абсолютно непрерывного распределения, а $g(x)$ — оставшаяся часть $\tilde{F}(x)$. Тогда задача вычисления интеграла сведется к поиску математического ожидания случайной величины $\eta = g(X)$:

$$\int\limits_{\mathbb{R}^{10}} \tilde{F}(x)dx = \int\limits_{\mathbb{R}^{10}} \rho(x)g(x)dx = \mathbb{E}g(X) = \mathbb{E}\eta$$

Теперь, посмотрим, плотность какого распределения лучше всего подходит в нашем случае для $\rho(x)$. «Базовым» распределением для подсчета методом Монте-Карло является равномерное, однако благодаря наличию квадратов суммы аргумента в степени экспоненты, удобнее будет воспользоваться нормальным распределением. Итак, $\rho(x)$ - плотность совместного распределения случайного вектора с 10-ю компонентами, имеющим и нормальное распределение с параметрами $\mu = 0, \sigma^2 = \frac{1}{2}$:

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

Таким образом, в качестве приближенного значения интеграла мы будем брать следующую величину:

$$\tilde{I} = \frac{1}{n} \sum_{i=1}^{n} g\left(\xi_1^i, \ldots, \xi_{10}^i\right), \quad \xi_j^i \sim \mathcal{N}\left(0, \frac{1}{2}\right), \ \forall i = \overline{1, n}, \ \forall j = \overline{1, 10}.$$

Погрешность метода Монте-Карло оценим с помощью ЦПТ:

$$\mathbb{P}(|\tilde{I} - I| \leq \varepsilon) = \mathbb{P}\left(\left|\frac{S_n}{n} - I\right| \leq \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - nI}{n}\right| \leq \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - nI}{\sqrt{n}\sigma_n}\right| \leq \varepsilon\frac{\sqrt{n}}{\sigma_n}\right) =$$
$$= \mathbb{P}\left(\frac{S_n - nI}{\sqrt{n}\sigma_n} \leq \varepsilon\frac{\sqrt{n}}{\sigma_n}\right) - \mathbb{P}\left(\frac{S_n - nI}{\sqrt{n}\sigma_n} \leq -\varepsilon\frac{\sqrt{n}}{\sigma_n}\right) \approx \Phi\left(\varepsilon\frac{\sqrt{n}}{\sigma_n}\right) - \Phi\left(-\varepsilon\frac{\sqrt{n}}{\sigma_n}\right) =$$
$$= 2\Phi\left(\varepsilon\frac{\sqrt{n}}{\sigma_n}\right) - 1 = 2F_n\left(\varepsilon\frac{\sqrt{n}}{\sigma_n}\right) - 1 = 2\left(1 - \frac{\alpha}{2}\right) - 1 = 1 - \alpha,$$

где $F_n, \Phi - \phi$ фнкция стандартного нормального распределения, $\sigma_n^2 -$ смещенная выборочная дисперсия (для η_n). Тогда получим оценку на погрешность:

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

где K_α - квантиль стандартного нормального распределения порядка α.

Для вычисления методом квадратур сделаем замену $x_i = \tan\left(\frac{\pi}{2}t_i\right), dx_i = \frac{\pi}{2}\frac{1}{\cos^2\left(\frac{\pi}{2}t_i\right)}, \ -1 \leqslant t_i \leqslant 1, \forall i = \overline{1, 10}$ и, с учетом симметричности, получим новое представление искомого интеграла:

$$I = \pi^{10} \int_0^1 \ldots \int_0^1 \frac{e^{-\left(\sum\limits_{i=1}^{10} \tan^2\left(\frac{\pi}{2}t_i\right) + \frac{1}{2^7 \prod\limits_{i=1}^{10} \tan^2\left(\frac{\pi}{2}t_i\right)}\right)}}{\prod\limits_{i=1}^{10} \left(\tan^2\left(\frac{\pi}{2}t_i\right)\cos^2\left(\frac{\pi}{2}t_i\right)\right)} dt_1 \ldots dt_{10}$$

6.3 Результаты работы программы

6.3.1 Метод-Монте Карло

```
n = 100000, epsilon = 1e-10, alpha = 0.99
```

Значение интерграла: 124.0220615848674
Погрешность: 0.04524826105723924

6.3.2 Метод квадратур

```
n = 8, h = np.pi/n
```

Значение интерграла: 122.71109951467847

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. $y = \min\limits_{x \in A} f(x)$, где
$$f(x) = x_1^3 \sin\left(\frac{1}{x_1}\right) + 10 x_1 x_2^4 \cos\left(\frac{1}{x_2}\right)$$
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
$$g(x) = (x_1 - 1)^2 + 100\left(x_2 - x_1^2\right)^2$$

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

7.2 Теоретические выкладки

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

Перейдём к полярным координатам: $x_1 = r\cos\varphi,\ x_2 = r\sin\varphi$, $r \in [0, 1]$, $\varphi \in [0, 2\pi]$.

Заметим, что $f(x_1,\ x_2) = f(x_1,\ -x_2)$, то есть функция чётна по переменной x_2. С учётом этого, получаем $\varphi \in [0,\ \pi]$.
$$r \sim U[0, 1]$$
$$\varphi \sim U[0,\ \pi]$$

Оценим точность предложенного алгоритма. Пусть $(x_1,\ x_2)$ — точка минимума, полученная описанным выше методом, а $(x_1^*,\ x_2^*)$ — реальная тока минимума. Тогда будем оценивать разность $|f(x_1,\ x_2) - f(x_1^*,\ x_2^*)|$.

Теорема 8. Пусть функция $f : \Omega \to \mathbb{R}$ непрерывно дифференцируема в выпуклой компактной области Ω пространства $\mathbb{R}^n$. Тогда $|f(x) - f(y)| \leqslant \sup\limits_{\xi \in \Omega} |\operatorname{grad} f(\xi)| \cdot |x - y|$.

Получается:

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

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

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

$$\Downarrow$$

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

Полученная оценка 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 δ-окрестности искомой точки минимума. В худшем случае точка минимума может оказаться на границе множества A (в таком случае δ-окрестность будет иметь минимальную площадь, а, значит, вероятность попадания туда равномерно распределенной случайной величины будет минимальной), поэтому оценивать будем вероятность вероятность попадания именно в граничную δ-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.

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

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:

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

Метод имитации отжига основывается на имитации физического процесса, который происходит при кристаллизации вещества. Предполагается, что атомы уже выстроились в кристалличекую решётку, но ещё допустимы переходы отдельных атомов из одной ячейки в другую. Считается, 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 процесса ищется одна или множество точек, где достигается минимум числовой функции $g(x)$, где $x = (x_1, \ldots, x_m) \in X$. Решение ищется последовательным вычислением точек x^0, x^1, … 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:

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

Здесь $t_i > 0$ — элементы произвольной убывающей, сходящейся к нулю положительной последовательности, которая задаёт аналог падающей температуры в кристалле.

Алгоритм имитации отжига не гарантирует нахождения минимума функции, однако при правильной политике генерации случайной точки в пространстве X, как правило, происходит улучшение начального приближения.

7.3 Результаты работы программы

7.3.1 Метод случайного поиска

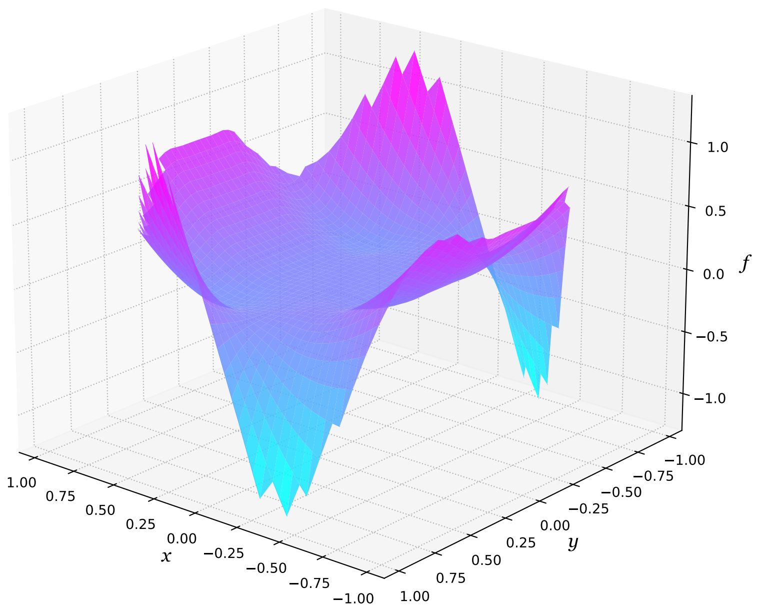

Рис. 24: $f(x)$, min значение ≈ -1.2856, погрешность ≈ 0.0221

7.3.2 Метод оджига

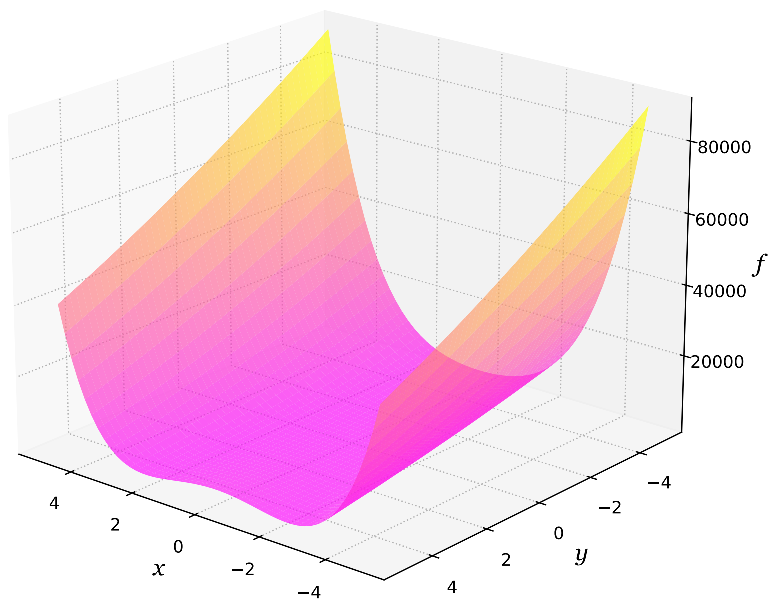

Рис. 25: $g(x)$, min значение ≈ 0.00018

8 Задание 8

8.1 Постановка задачи

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

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

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

8.2 Теоретические выкладки

Приведем **аналитическое решение**:

$$f(x,\ y) = x^2 - y^2 = r^2(\cos^2\varphi - \sin^2\varphi) = r^2\cos(2\varphi)$$

$$x^2 + y^2 = r^2(\cos^2\varphi + \sin^2\varphi) = r^2 \leqslant 1$$

То есть рассматривается задача Дирихле внутри круга $r \leqslant 1$, тогда решение представимо в виде:

$$u(r,\ \varphi) = C_2 + \sum_{k=1}^{\infty} r^k (A_k \cos(k\varphi) + B_k \sin(k\varphi))$$

$$u\big|_{x^2+y^2=1} = u\big|_{r=1} = r^2\big|_{r=1}\cos(2\varphi) = \cos(2\varphi)$$

$$k = 2,\ A_k = 1,\ C_2 = 0,\ B_k = 0$$

Ответ: $u(r,\ \varphi) = r^2\cos(2\varphi)$ или $u(x,\ y) = x^2 - y^2$.

Практическое решение:

Применим двумерный метод Монте-Карло для решения имеющейся внутренней задачи Дирихле. Для этого выберем шаг h и проведем в плоскости, содержащей D, сетку из элементарных квадратов. Будем рассматривать только узлы, попавшие в область D. Их можно разделить на два типа: внутренние - для которых 4 соседних узла лежат внутри или на границе области, и граничные — для которых число соседних внутренних узлов меньше 4. Значения функции u, заданные на границе, переносятся на граничные узлы $Q : u(Q) = f(Q)$. А значения во внутренних узлах P вычисляются из соотношения:

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

где P_1, P_2, P_3, P_4− соседние узлы к P.

Рассмотрим связанную с этой системой теоретико-вероятност ную схему. Эту схему принято объяснять в виде «задачи о пьяных». Будем рассматривать стороны решетки как городские кварталы, а узлы - как перекрестки городских улиц. Предположим, что из узла P выходит человек, который с равной вероятностью $\left(\frac{1}{4}\right)$ может попасть в любой из соседних узлов. Аналогично, попав в очередной узел (войдя на очередной перекресток), он с равной вероятностью идет по одному из примыкающих к этому перекрестку кварталу, пока не выйдет на очередной перекресток. Будем считать, что город обнесен глубоким рвом, т.е. выйдя на границу города (граничный узел решетки), "пьяный"свал ивается в ров (остается в этом узле). Возникает вопрос об отыскании вероятности $u(P, Q)$ того, что «пьяный», выйдя из узла P, окончит блуждание в граничном узле Q.

Не ограничивая время, можно утверждать, что с вероятностью, равной единице, человек, в конце концов, окажется в граничной точке. Найти искомую вероятность в явном виде сложно, однако можно вывести полезное соотношение. Заметим, что событие, заключающееся в том, что «пьяный» попадет из точки P в точку Q, равносильно тому, что он из точки P попадет в либо в точку P_1, либо в P_2, либо в P_3, либо в P_4, а оттуда в Q. Так как вероятности попадания в соседние узлы равны, то по теореме сложения вероятностей получим систему:

$$u(P, Q)=\frac{u\left(P_1, Q\right)+u\left(P_2, Q\right)+u\left(P_3, Q\right)+u\left(P_4, Q\right)}{4}$$

причем, вероятность $u(P, Q)$, удовлетворяет следующим краевым условиям для граничных узлов:

$$u\left(Q_1, Q_2\right)= \begin{cases}1 & Q_1=Q_2, \\ 0, & Q_1 \neq Q_2\end{cases}$$

Если промоделировать движение человека N раз, то заставляя его каждый раз стартовать из точки P, и сосчитать количество M испытаний, при которых его путь оканчивается в Q, то получим:

$$\frac{M}{N} \approx u(P, Q)$$

Предположим, что после того, как «пьяный» сваливается в ров в граничном узле Q, с него взимается штраф, равный $f(Q)$. Тогда величина штрафа при выходе из точки P будет случайной величиной $\xi(P)$ со значениями $\{f(Q_1),\ \ldots,\ f(Q_s)\}$, где $\{Q_1,\ \ldots,\ Q_s\}$ — совокупность всех граничных узлов. Вероятность заплатить штраф $f(Q_i)$ равна $u(P,\ Q_i)$. Тогда мат ожидание случайной величины зависит от начальной точки P и равно:

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

Полученная функция $\omega(P)$ удовлетворяет разностному уравнению

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

Тогда получим алгоритм решения внутренней задачи Дирихле методом Монте-Карло:

1. Строим сетку из элементарных квадратиков с шагом h.

2. Делим узлы на граничные и внутренние. для граничных верно, что $u(Q)=f(Q)$.

3. Из каждого внутреннего узла P пустим N случайных блужданий. Обозначив за Q_i — узел, в котором закончилось i-е блуждание, получим значение функции в узле P:

$$u(P)=\frac{\sum_{i=1}^N f\left(Q_i\right)}{N}$$

8.3 Результаты работы программы

8.3.1 Аналитическое решение

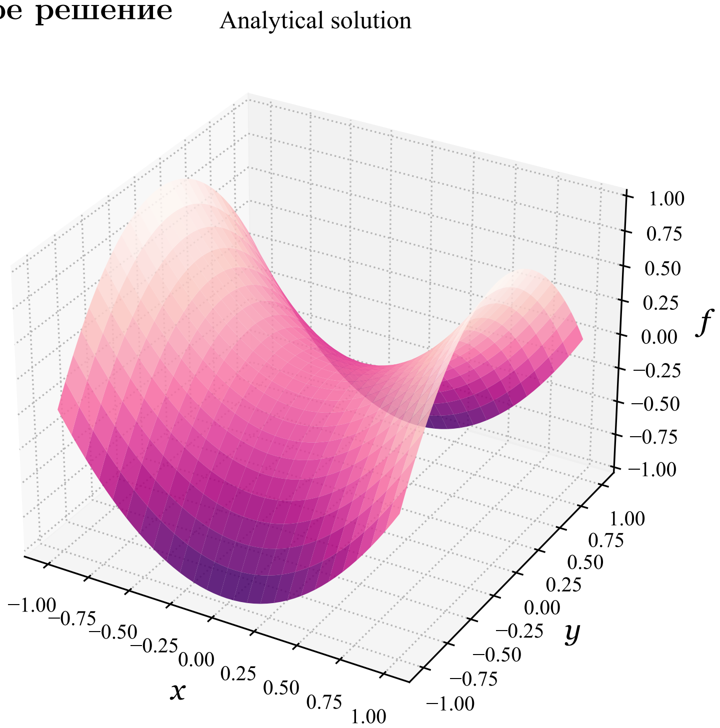

Рис. 26: $f(x,\ y) = x^2 - y^2$, $n = 50$ точек

8.3.2 Пример работы метода Монте-Карло

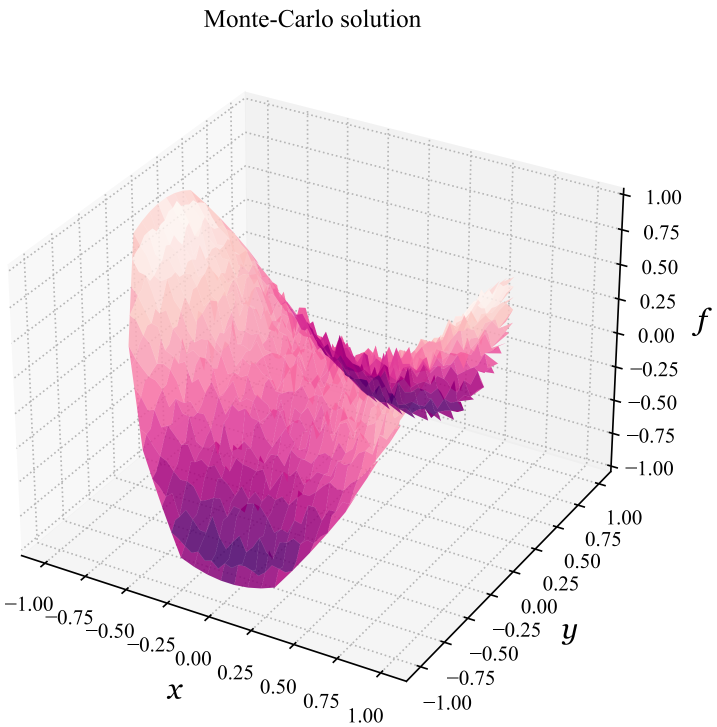

Рис. 27: Решение методом Монте-Карло

8.3.3 Разбиение точек на внутренние и граничные

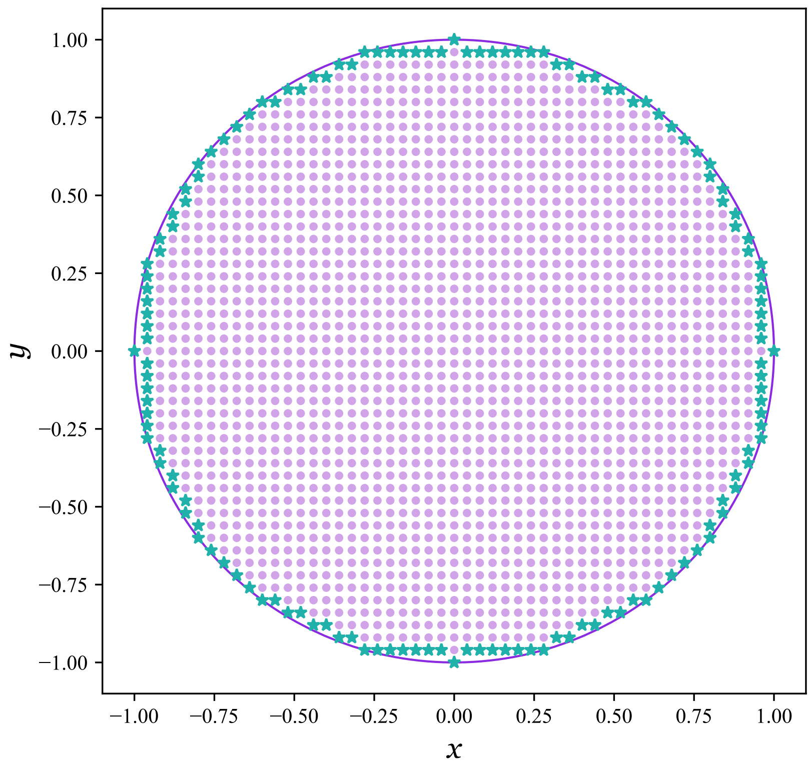

Рис. 28: $n = 50$ точек

9 Задание 9

9.1 Постановка задачи

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 марковский гауссовский процесс. Начальное значение 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 точки ломаной, с целью получения визуально непрерывной линии.

9.2 Теоретические выкладки

Определение 18. Случайным процессом $X_t, t \in T$ называется континуальное семейство случайных величин $X^T = X_t, t \in T$, определенных на общем фазовом пространстве $(\mathcal{X}, \mathcal{A})$, где множество индексов T трактуется как время.

Рассмотрим произвольный случайный процесс $X_t, t \in T = [0, +\infty)$ на $(\mathcal{X}, \mathcal{A})$.

Определение 19. Случайный процесс $X_t, t \in T$ (T произвольной мощности) называется гауссовским, если для любого конечного набора индексов $t_1, t_2, \ldots, t_n \in T$ случайный вектор $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ является гауссовским (имеет многомерное нормальное распределение).

Если существует гауссовский процесс, то для него определены функции математического ожидания и ковариации:

$$m(t) = \mathbb{E}\left(X_t\right), \quad \operatorname{cov}(t, s) = \operatorname{cov}\left(X_t, X_s\right),$$

Через последнюю определяется следующая симметричная и положительно определенная ковариационная матрица:

$$R = R\left(t_1, t_2, \ldots, t_n\right) = \begin{pmatrix} \operatorname{cov}\left(t_1, t_1\right) & \operatorname{cov}\left(t_1, t_2\right) & \ldots & \operatorname{cov}\left(t_1, t_n\right) \\ \operatorname{cov}\left(t_2, t_1\right) & \operatorname{cov}\left(t_2, t_2\right) & \ldots & \operatorname{cov}\left(t_2, t_n\right) \\ \ldots & \ldots & \ldots & \ldots \\ \operatorname{cov}\left(t_n, t_1\right) & \operatorname{cov}\left(t_n, t_2\right) & \ldots & \operatorname{cov}\left(t_n, t_n\right) \end{pmatrix}$$

В свою очередь, через полученную ковариационную матрицу и математическое ожидание выражается плотность маргинальных распределений:

$$p_{X_1, \ldots, X_n}(x) = \frac{1}{\sqrt{(2\pi)^n \det(R)}} e^{-\frac{\left\langle R^{-1}(x-m), (x-m)\right\rangle}{2}},$$

где $x = \left(x_1, \ldots, x_n\right), m = \left(m\left(t_1\right), \ldots, m\left(t_n\right)\right)$.

Определение 20. Случайный процесс $X_t, t \in T = [0, +\infty)$ называется процессом с независимыми приращениями, если для любых $0 \leq t_1 < t_2 < \ldots < t_n$ случайные элементы (со значениями в $\mathcal{X}$)$X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ независимы в совокупности.

Утверждение 1. Процесс с независимыми приращениями обладает марковским свойством.

Для процесса, обладающего марковским свойством, выполнено уравнение Колмогорова-Чепмена.

Обозначим $P_{X_{t+h}|X_t=x}(A) = P^{t,t+h}(x, A)-$ вероятность перехода из состояния x в момент времени t во множество A в момент времени $t+h$ (переходное ядро). Тогда уравнение Колмогорова-Чепмена имеет вид: $0 \leqslant s < t < u$

$$\int_{\mathcal{X}} P^{s,t}(x, dy) P^{t,u}(y, A) = P^{s,u}(x, A).$$

Определение 21. Случайный процесс $X_t, t \in T = [0, +\infty)$ называется однородным по времени, если распределения приращений $X_{t+h} - X_t$ не зависят от t при $\forall h > 0$.

Однородный по времени процесс с независимыми приращениями называется процессом с независимыми однородными приращениями.

Уравнение Колмогорова-Чепмена для процессов с независимыми однородными приращениями имеет вид: $\left(Q_h = P_{X_{t+h}-X_t}\right)$

$$Q_{h_1} * Q_{h_2} = Q_{h_1+h_2},$$

где «*» - операция свертки приращений, и $P^{t,t+h}(x, A) = Q_h(A - x)$.

Определение 22. Винеровским процессом (стандартным броуновским движением) называется процесс с независимыми однородными приращениями, имеющий непрерывные траектории, причем $\mu = 0; \sigma = 1; P(X_0 = 0) = 1; X_{t+h} - X_t \sim \mathcal{N}(0, h), \forall h > 0$.

Найдем ковариационную функцию Винеровского процесса. Для $0 = t_0 < t_1 < \ldots < t_n$ рассмотрим случайный вектор $\delta X = \left(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}\right)$ приращений процесса. Посчитаем плотность распределения этого вектора:

$$p_{\delta X}(y_1, \ldots, y_n) = \{\text{независимость приращений }\} = \prod_{i=1}^{n} p_{X_{t_i} - X_{t_{i-1}}}(y_i) =$$

$$= \left\{X_{t_i} - X_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})\right\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} e^{-\frac{y_i^2}{2(t_i - t_{i-1})}} =$$

$$= e^{-\sum_{i=1}^{n} \frac{y_i^2}{2(t_i - t_{i-1})}} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}}.$$

Вспомним про $P(X_0 = 0) = 1$:

$$\begin{aligned} &P(X_1 - X_0 = y_1, X_2 - X_1 = y_2, \ldots, X_n - X_{n-1} = y_n) = \\ &= P(X_1 = y_1, X_2 - y_1 = y_2, \ldots, X_n - y_{n-1} = y_n) = \\ &= P(X_1 = y_1, X_2 = y_1 + y_2, \ldots, X_n = y_{n-1} + y_n) \end{aligned}$$

Тогда:

$$p_{\delta X}(y_1, \ldots, y_n) = p_{X_1, \ldots, X_n}(y_1, y_1 + y_2, \ldots, y_{n-1} + y_n)$$

Таким образом:

$$p_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = p_{\delta X}(x_1, x_2 - x_1, \ldots, x_n - x_{n-1}) =$$

$$= e^{-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}}.$$

С другой стороны:

$$p_{X_1, \ldots, X_n}(x) = \frac{1}{\sqrt{(2\pi)^n \det(R)}} e^{-\frac{\left\langle R^{-1}(x-m), (x-m)\right\rangle}{2}}$$

Тогда получим :

$$m = 0, \quad R = \left(\min(t_i, t_j)\right)_{i,j=1}^{n}.$$

Т.е. ковариационная функция равна $\operatorname{cov}(t, s) = \min(s, t)$.

Независимые траектории винеровского процесса будем моделировать методом разделения отрезка $[0, 1]$ в отношении α исходя из следующих соображений:

1. В начальный момент времени $X_0 = 0$.

2. W_1 генерируем из $W_1 - W_0 \sim \mathcal{N}(0,1)$.

3. Разделим отрезок $[t_1, t_2]$ в от ношении $\alpha \in (0,1)$ точкой $t = (1-\alpha)t_1 + \alpha t_2$. Рассмотрим условную плот ность маргинального распределения:

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

   В соответствии с написанным выше: плотности маргинальных распределений, написанных справа, будут равны:

$$p_{W_{t_1}, W_t, W_{t_2}}(x_1, x, x_2) = \frac{1}{\sqrt{(2\pi)^3 \det(R_1)}} e^{-\frac{(x_1, x, x_2) \cdot R_1^{-1} \cdot (x_1, x, x_2)'}{2}},$$
$$p_{W_{t_1}, W_{t_2}}(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 \det(R_2)}} e^{-\frac{(x_1, x_2) \cdot \cdot R_2^{-1} \cdot (x_1, x_2)'}{2}},$$

   где $R_1, R_2 -$ соответствующие матрицы ковариаций. Т.к. $\operatorname{cov}(t,s) = \min(s,t)$, то находим выражения для R_1 и R_2 :

$$R_1 = \begin{pmatrix} \min(t_1, t_1) & \min(t_1, t_2) & \min(t_1, t_3) \\ \min(t_2, t_1) & \min(t_2, t_2) & \min(t_2, t_3) \\ \min(t_3, t_1) & \min(t_3, t_2) & \min(t_3, t_3) \end{pmatrix} = \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t & t \\ t_1 & t & t_2 \end{pmatrix},$$
$$R_2 = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix}$$

4. Считаем определители и обратные матрицы для R_1 и R_2 :

$$|R_1| = t_1(t - t_2)(t_2 - t), \quad |R_2| = t_1(t_2 - t_1),$$
$$R_1^{-1} = \begin{pmatrix} \frac{t}{t_1(t-t_1)} & \frac{t}{t_1 - t} & 0 \\ \frac{1}{t_1 - t} & \frac{t_2 - t_1}{(t_2 - t)(t - t_1)} & \frac{1}{t - t_2} \\ 0 & \frac{1}{t - t_2} & \frac{1}{t_2 - t} \end{pmatrix}, \quad R_2^{-1} = \begin{pmatrix} \frac{t_2}{t_1(t_2 - t_1)} & \frac{1}{t_1 - t_2} \\ \frac{1}{t_1 - t_2} & \frac{1}{t_2 - t_1} \end{pmatrix}.$$

5. Подставляя определители и обратные матрицы в формулу условной вероятности, а также вспоминая, что $t = (1-\alpha)t_1 + \alpha t_2$, получим:

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

Т.е. следующая случайная величина генерируется с нормальным распределением $W_t \sim \mathcal{N}((1-\alpha)x_1 + \alpha x_2, \alpha(1-\alpha)(t_2 - t_1))$.

Итак, сформулируем алгоритм построения траектории винеровского процесса:

1. Вводим входные данные алгоритма: $t_0 = 0, t_1 = 1, W_0 = 0$.

2. Разыгрываем $W_1 \sim \mathcal{N}(0,1)$.

3. Делим отрезок $[t_0, t_1]$ точкой $t = (1-\alpha)t_1 + \alpha t_0$, разыгрываем случайную величину $W_t \sim \mathcal{N}((1-\alpha)x_1 + \alpha x_2, \alpha(1-\alpha)(t_1 - t_0))$.

4. Рекурсивно продолжаем делить отрезок и генерировать W_t пока не достигнем нужной точности $t_{k+1} - t_k \leq \varepsilon$.

5. Дополнительно: для проверки правильности работы алгоритма можно построить доверительные интервалы по правилу трех сигм.

Определение 23. Случайный процесс $X_t, t \in T$ называется стационарным в узком смысле если $\forall t_1, \ldots, t_n \in T$ совместное распределение случайных величин одно и то же $\forall \tau$: $t_1 + \tau, \ldots, t_n + \tau \in T$.

Определение 24. Процессом Орнштейна-Уленбека называется марковский гауссовский стационарный процесс.

Процесс Орнштейна-Уленбека обладает свойством симметричности.

Для процесса Орнштейна-Уленбека работает уравнение Колмогорова-Чепмена, определенное для процессов обладающих марковскими свойствами. Тогда для коэффициента корреляции выполнено: $\rho(t+s) = \rho(t)\rho(s)$. Найдем переходную плотность данного процесса:

$$p_{X_t}(x_1 \mid X_s = x_2) = \frac{p_{X_t, X_s}(x_1, x_2)}{p_{X_s}(x_2)}.$$

Из стационарности процесса имеем:

$$EX_t = m, \operatorname{Var} X_t = \sigma^2, \operatorname{cov}(s,t) = \operatorname{cov}(|s-t|) = \rho(|t-s|) \operatorname{Var} X_t = \rho(|t-s|)\sigma^2$$

По аналогии с винеровским процессом, запишем плотности маргинальных распределений:

$$p_{X_t, X_s}(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 \det(R)}} e^{-\frac{(x_1, x_2) R^{-1} (x_1, x_2)'}{2}}, \quad p_{X_s}(x_2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_2^2}{2\sigma^2}},$$

где

$$R = \begin{pmatrix} \operatorname{cov}(0) & \operatorname{cov}(|s-t|) \\ \operatorname{cov}(|s-t|) & \operatorname{cov}(0) \end{pmatrix} = \begin{pmatrix} \sigma^2 & \operatorname{cov}(|s-t|) \\ \operatorname{cov}(|s-t|) & \sigma^2 \end{pmatrix}.$$

Т.е. переходная плотность равна:

$$p_{X_t}(x_1 \mid X_s = x_2) = \frac{1}{\sqrt{2\pi\left(\sigma^2 - \frac{\operatorname{cov}^2(|s-t|)}{\sigma^2}\right)}} e^{-\frac{\left(x_1 - \frac{\operatorname{cov}(|s-t|)}{\sigma^2} x_2\right)^2}{2\left(\sigma^2 - \frac{\operatorname{cor}^2(|s-t|)}{\sigma^2}\right)}},$$

т.е. распределение X_t при условии $X_s = x_2$ — это нормальное распределение с параметрами $\mathcal{N}\left(\frac{\operatorname{cov}(|s-t|)}{\sigma^2} x_2, \sigma^2 - \frac{\operatorname{cov}^2(|s-t|)}{\sigma^2}\right)$.

Теорема 9. Пусть функция $u(t)$ определена $\forall t > 0$ и ограничена на каждом конечном интервале. Если $u(t)$ удовлетворяет соотношению $u(t+s) = u(t)u(s)$, то или $u(t) = 0, \forall t > 0$, или $\exists \lambda > 0 : u(t) = e^{-\lambda t}$.

Т.к. $\rho(t+s)=\rho(t)\rho(s)$, то эта теорема применима к коэффициенту корреляции. Тогда, если $\rho(t)=0, \forall t>0$, то $\forall t,s>0 \operatorname{cov}(X_t,X_s)=0$, что с учетом гауссовости процесса означает независимость в совокупности случайных величин X_t. В таком случае, моделирование процесса Орнштейна-Уленбека заключается в генерации независимых сл.в. с распределением $\mathcal{N}\left(\frac{\operatorname{cov}(|s-t|)}{\sigma^2}x_2, \sigma^2-\frac{\operatorname{cov}^2(|s-t|)}{\sigma^2}\right)$.

Рассмотрим $\rho(t)=e^{-\lambda t}, \operatorname{cov}(t,s)=e^{-\lambda(t-s)}\sigma^2$. Для моделирования траекторий процесса X_t методом, таким же, как для винеровского процесса, вычислим условную плотность:

$$p_{X_t}\left(x \mid X_{t_1}=x_1, X_{t_2}=x_2\right)=\frac{p_{X_{t_1},X_t,X_{t_2}}\left(x_1,x,x_2\right)}{p_{X_{t_1},X_{t_2}}\left(x_1,x_2\right)}, \quad t=(1-\alpha)t_1+\alpha t_2.$$

Поскольку процесс гауссовский, то плотности маргинальных распределений будут равны:

$$p_{X_{t_1},X_t,X_{t_2}}\left(x_1,x,x_2\right)=\frac{1}{\sqrt{(2\pi)^3\det\left(R_1\right)}}e^{-\frac{(x_1,x,x_2)\cdot R_1^{-1}\cdot(x_1,x,x_2)'}{2}},$$

$$p_{X_{t_1},X_{t_2}}\left(x_1,x_2\right)=\frac{1}{\sqrt{(2\pi)^2\det\left(R_2\right)}}e^{-\frac{(x_1,x_2)\cdot R_2^{-1}\cdot(x_1,x_2)'}{2}}$$

$$R_1=\begin{pmatrix} 1 & e^{-\lambda(t-t_1)} & e^{-\lambda(t_2-t_1)} \\ e^{-\lambda(t-t_1)} & 1 & e^{-\lambda(t_2-t)} \\ e^{-\lambda(t_2-t_1)} & e^{-\lambda(t_2-t)} & 1 \end{pmatrix}$$

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

$$p_{X_t}\left(x \mid X_{t_1}=x_1, X_{t_2}=x_2\right)=\frac{1}{\sqrt{2\pi b^2}}e^{-\frac{(x-a)^2}{2b^2}}, \text{ где}$$

$$a=\frac{e^{-\lambda\alpha(t_2-t_1)}\left(1-e^{-2\lambda(1-\alpha)(t_2-t_1)}\right)x_1+e^{-\lambda(1-\alpha)(t_2-t_1)}\left(1-e^{-2\lambda\alpha(t_2-t_1)}\right)x_2}{1-e^{-2\lambda(t_2-t_1)}},$$

$$b^2=\sigma^2\frac{1-e^{-2\lambda(1-\alpha)(t_2-t_1)}-e^{-2\lambda\alpha(t_2-t_1)}+e^{-2\lambda(t_2-t_1)}}{1-e^{-2\lambda(t_2-t_1)}}.$$

Если положить $\mathbb{E}X_t=m$, то выражения немного изменятся:

$$a=\frac{e^{-\lambda\alpha(t_2-t_1)}\left(1-e^{-2\lambda(1-\alpha)(t_2-t_1)}\right)(x_1-m)+e^{-\lambda(1-\alpha)(t_2-t_1)}\left(1-e^{-2\lambda\alpha(t_2-t_1)}\right)(x_2-m)}{1-e^{-2\lambda(t_2-t_1)}},$$

$$p_{X_t}\left(x \mid X_{t_1}=x_1, X_{t_2}=x_2\right)=\frac{1}{\sqrt{2\pi b^2}}e^{-\frac{(x-m-a)^2}{2b^2}}$$

9.3 Результаты работы программы

9.3.1 Винеровский процесс

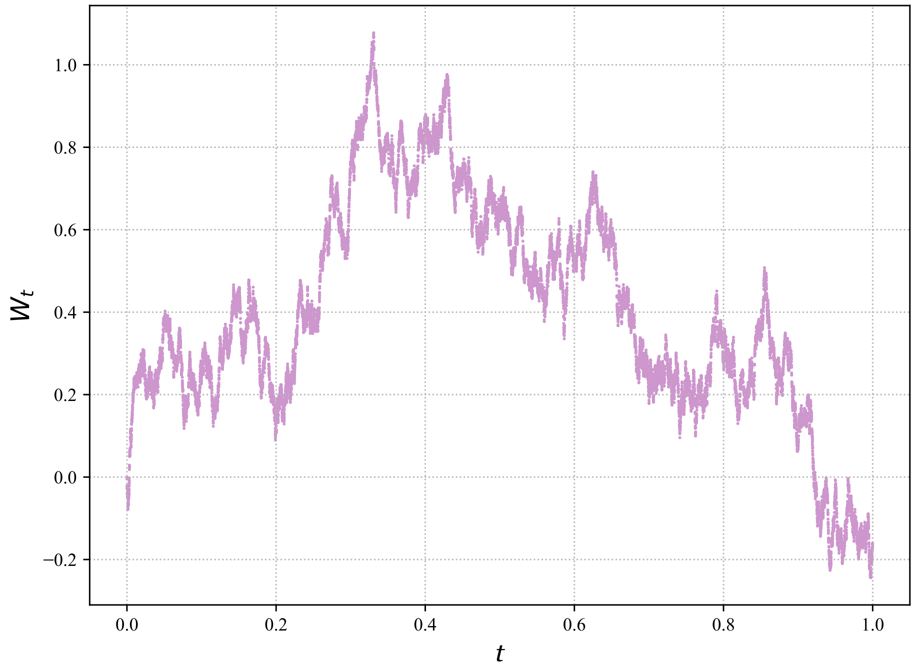

Рис. 29: График траектории винеровского процесса $\alpha = 0.2, \varepsilon = 0.0001$

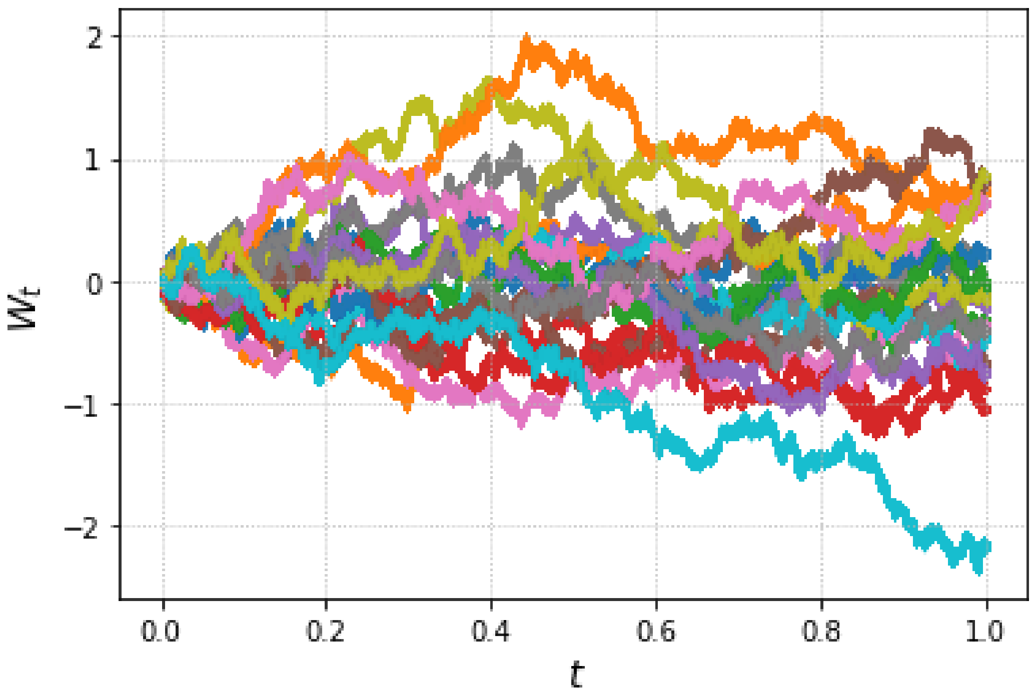

Рис. 30: График траекторий 20 винеровских процессов $\alpha = 0.37, \varepsilon = 0.000001$

9.3.2 Процесс Орнштейна-Уленбека

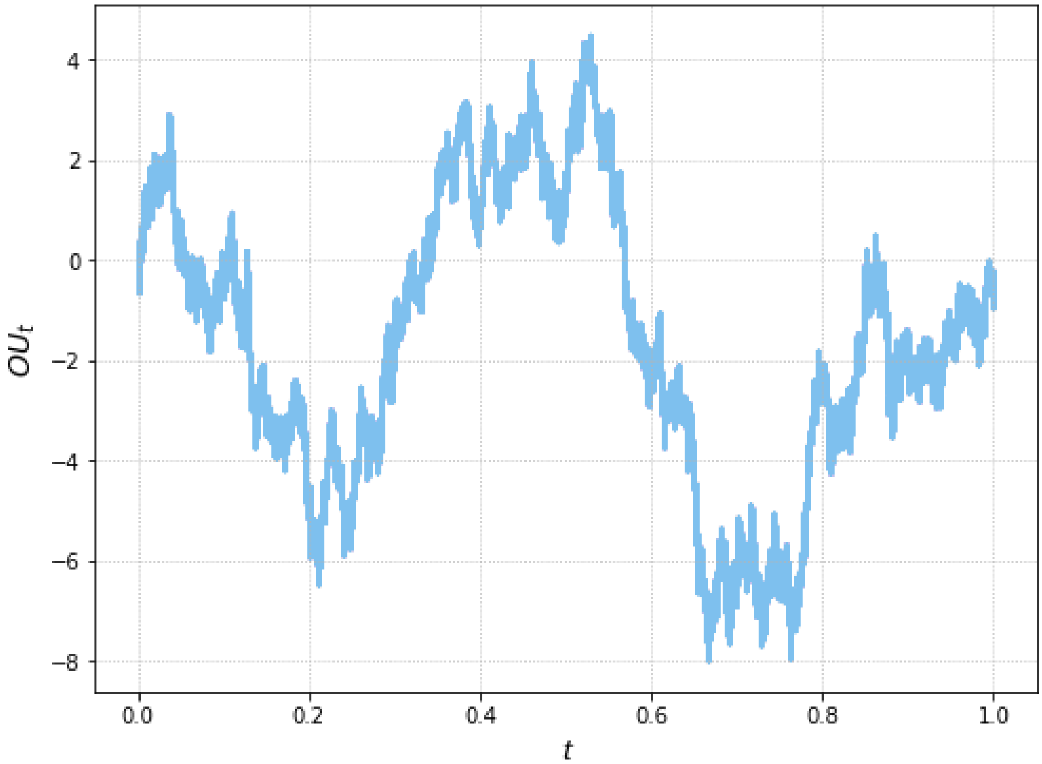

Рис. 31: График траектории процесса Орнштейна-Уленбека, $\sigma = 3, \ \lambda = 6, \ \varepsilon = 0.0001$

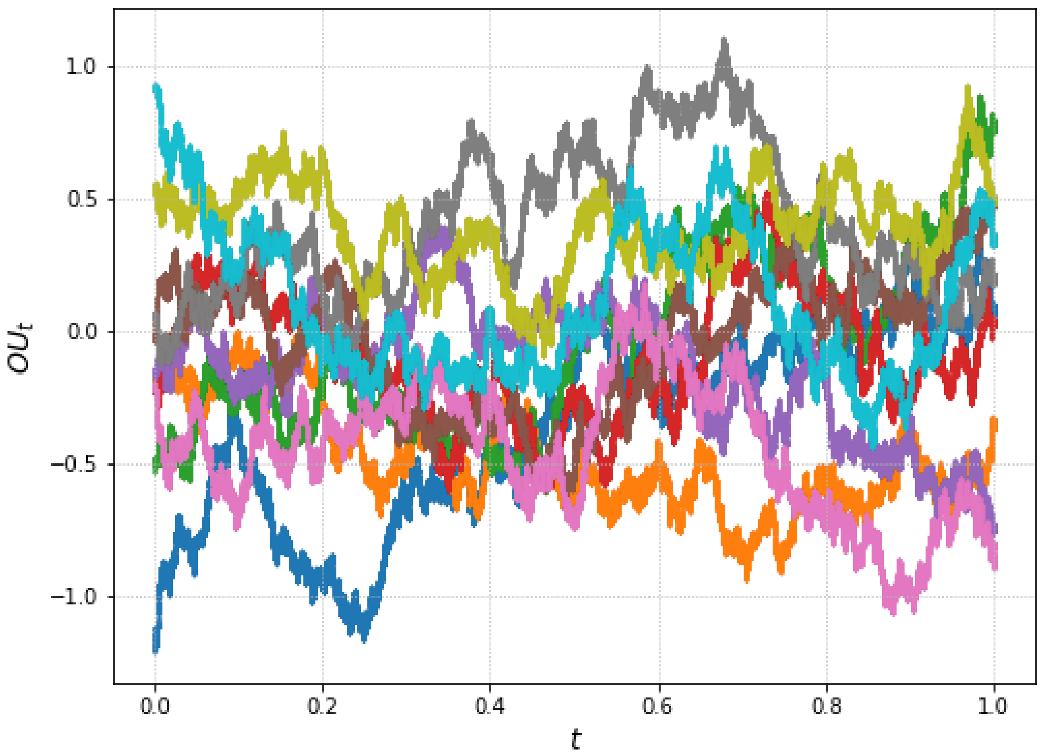

Рис. 32: Траектории 10 процессов Орнштейна-Уленбека, $\sigma = 0.5, \ \lambda = 2, \ \varepsilon = 0.00001$

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 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 Теоретические выкладки

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 $x_0, w_0, \ldots, w_{N-1}, v_0, \ldots, v_{N-1}$ независимые в совокупности случайные величины. $Y_{N-1} = (y_0, \ldots, y_{N-1})^T$ — все наблюдения, а $X_{N-1} = (x_0, \ldots, x_{N-1})$ — исходный процесс, который надо найти. Для этого воспользуемся фильтром Калмана, а точнее его схемой «шагаем-меряем», общий вид которой совпадает с вышеприведенной системой. Фильтр Калмана для схемы «шагаем-меряем» имеет вид:

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

Замечание 1. 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 σ и λ, $y_{k+1} = x_{k+1} + v_{k+1}$. То есть сразу имеем, что $C = 1$, а N_k — дисперсия белого шума (обозначим эту дисперсию за σ_v^2). В системе нам пока неизестны 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

$$\mathbb{V}ar\, x_{k+1} = A_k^2 \mathbb{V}ar\, x_k + \mathbb{V}ar\, w_k = A_k^2 \mathbb{V}ar\, x_k + M_k,$$

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

Замечание 2. Здесь и далее считаем распределение помехи v гауссовским. В случае распределения Коши матожидание помехи не определено и фильтрацию получить не получится.

Ковариационная функция процессса Орнштейна–Уленбека $R(t,s) = \sigma^2 e^{-\lambda|t-s|}$. Тогда:

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

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

Получаем фильтр Калмана для нашей задачи:

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

10.3 Результаты работы программы

10.3.1 Процесс Орнштейна-Уленбека

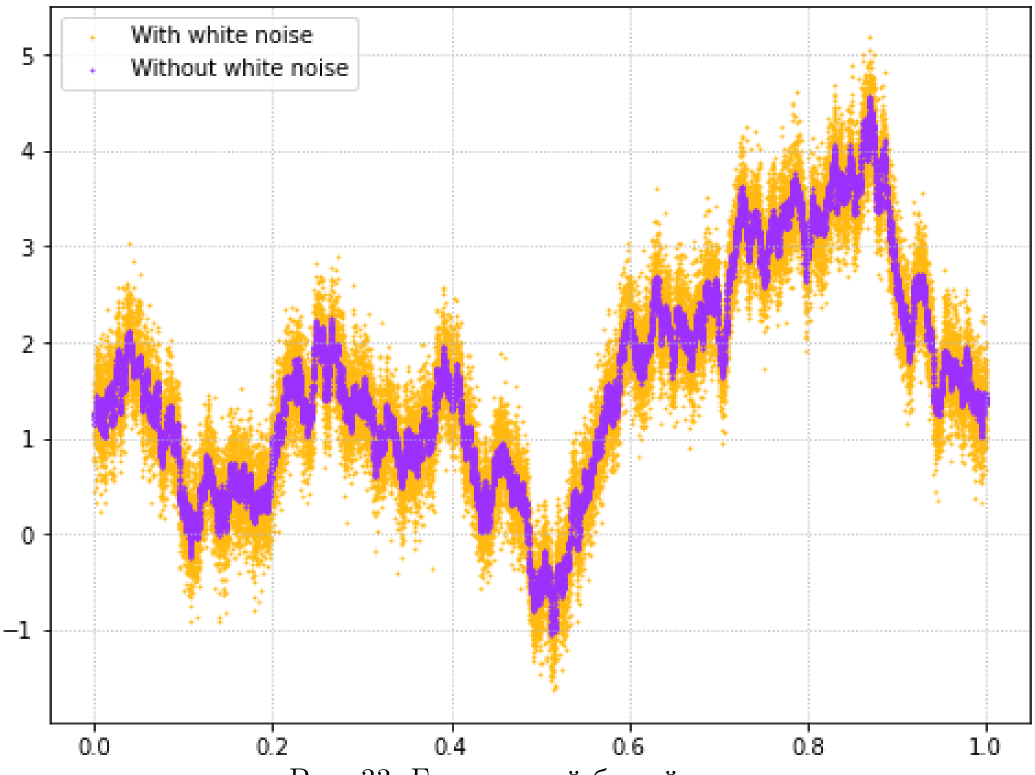

Рис. 33: Гауссовский белый шум

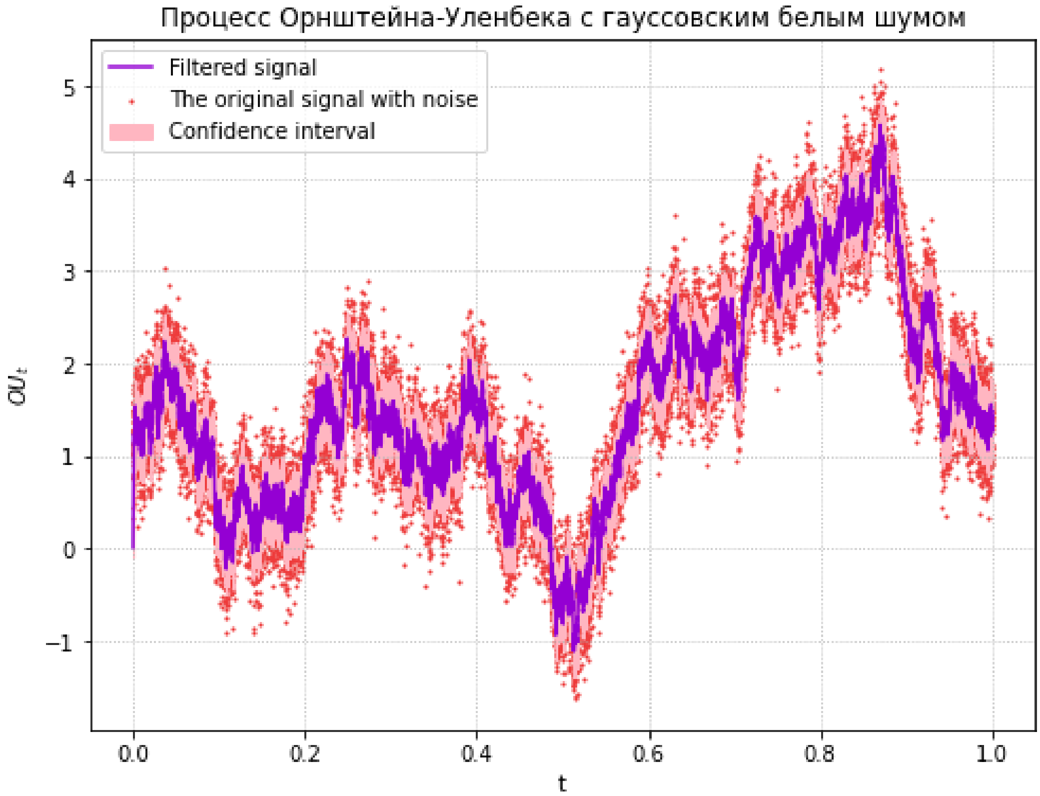

Рис. 34: Результат работы фильтра Калмана для гауссовского белого шума

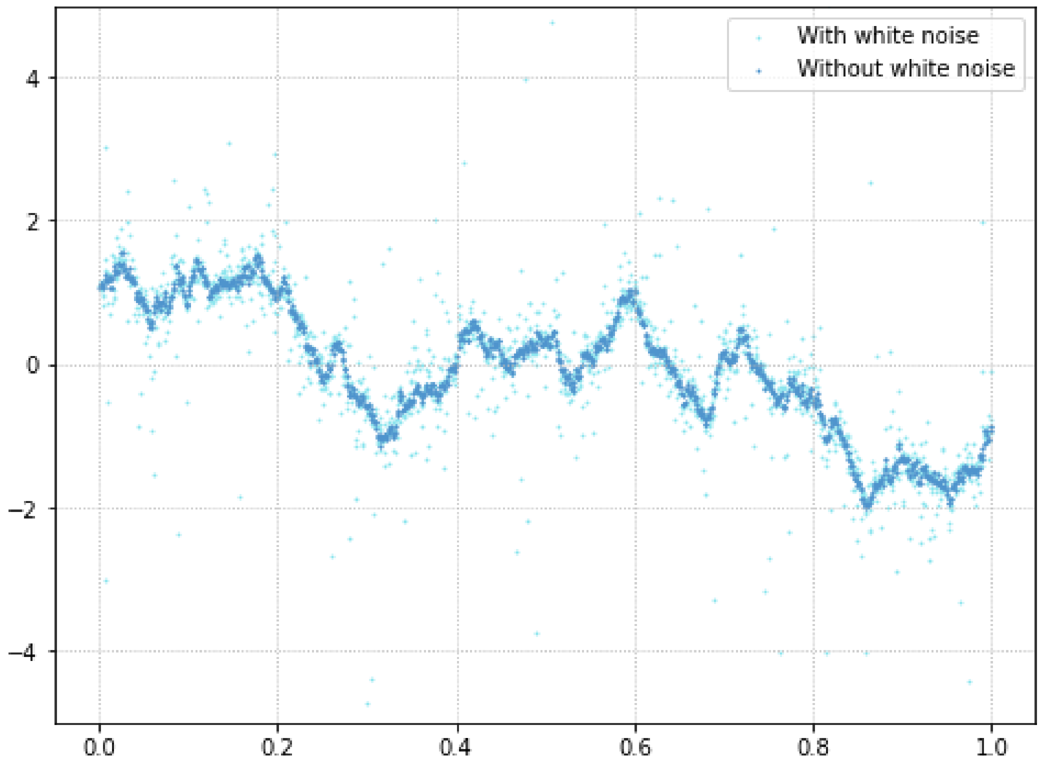

Рис. 35: Белый шум имеет распределение Коши

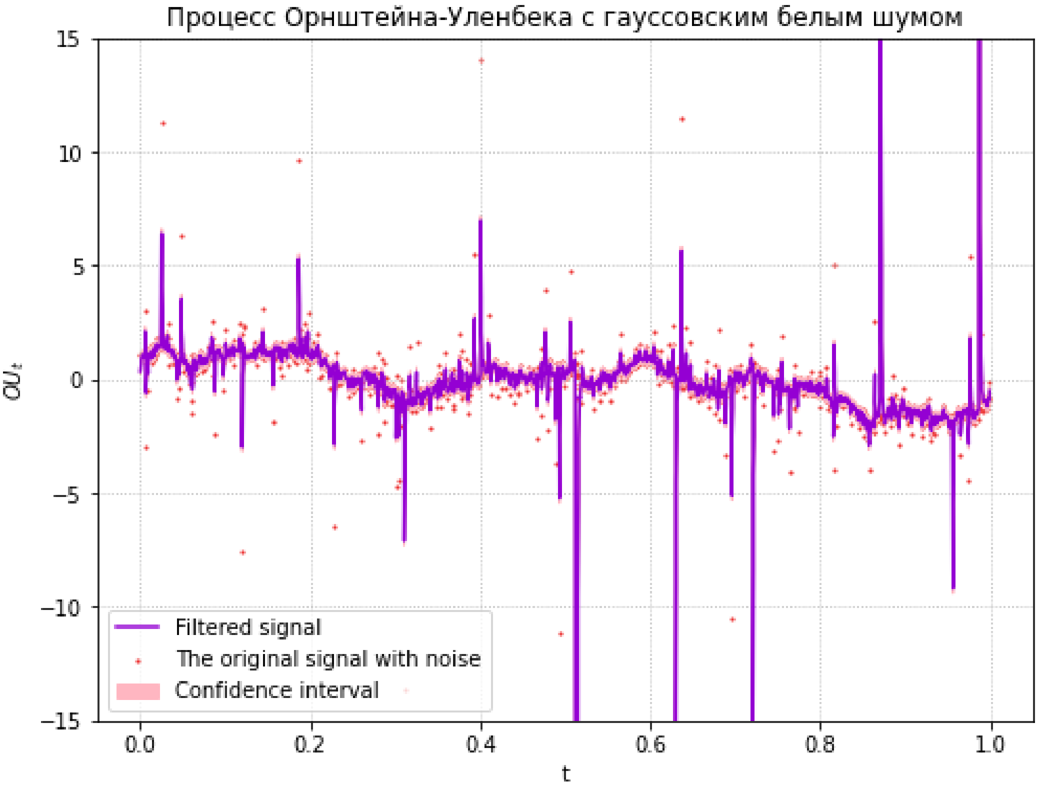

Рис. 36: Результат работы фильтра Калмана для белого шума с распределением Коши

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 χ^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 Шедлеара 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 Теоретические выкладки

Для моделирования системы массового обслуживания (СМО) будем генерировать времена поступления заявок (для пуассоновского поля с интенсивностью потока заявок λ): каждая i-я заявка поступает во время t_i, причем $\delta_{t_i} = t_i - t_{i-1}$. Время обработки каждой i-й заявки (обозначим s_i) имеет распределен ие $\mathcal{X}^2(10)$ с 10-ю степенями свободы. Для каждой заявки также есть время ее обработки Q_i. При этом все заявки обрабатываются последо вательно. Если к моменту посту пления i-й заявки очереди на исполнение нет $(Q_{i-1} < t)$, то $Q_i = t_i + s_i$ (предыду щая $(i-1)$ заявка обработана), если же есть о чередь $Q_i \geq t_i$, то $Q_i = Q_{i-1}$. В общем виде полу чается: $Q_i = t_i + \max(0, Q_{i-1} - t_i) + s_i$. Для каждой i - й заявки считается число заявок в очереди n_i : если во время ее поступления очередь пуста, то $n_i = 0$, если же нет, то числом заявок в очереди будет количество таких Q_k, что $k < i$ и $Q_k > t_i$. Среднее время обработки равно $E(s_i) = 10$, а средний интервал между временами поступления заявок равен $E(\delta_{t_i}) = \frac{1}{\lambda}$. Очередь будет наблюдаться при $E(s_i) \geq E(\delta_{t_i})$,т.е. очередь будет неограниченно расти при $\lambda > 0.1$ и убывать при $\lambda < 0.1$

Теперь рассмотрим систему массового обслуживания с циклической интенсивностью $\lambda(1 + \cos(t))$ и единичными скачками. Теперь время t_i– время наступление некоторого события, а $N(t_i, t_j)$ - число событий наступивших в отрезок времени $[t_i, t_j]$. Также теперь δ_{t_i} имеет распределение $F(x) = 1 - e^{(\Lambda(t) - \Lambda(t+x))}$, где $\Lambda(t) = \int_0^t \lambda(\tau) d\tau, x \leq 0$. В нашем случае $\Lambda(t) = \lambda(t + \sin(t))$. Теорема 11.1 Пусть существует одномерный неоднородный пуассоновский процесс с параметром $\lambda^*(t)$, тогда количество точек $N^*(T)$ на фиксированном интервале $(0, T]$ имеет распределение Пуассона с параметром $\mu^* = \Lambda^*(T) - \Lambda^*(0)$. Пусть $X_1^*, \ldots, X_{N(T)}^*$– точки процесса на рассматриваемом интервале. Пусть $0 \leq t \leq T, \lambda(t) \leq \lambda^*(t). \forall i = \overline{1, n}$ удаляем все X_i^* с вероятностью $1 - \frac{\lambda(X_i^*)}{\lambda^*(X_i^*)}$. Тогда оставшиеся точки образуют одномерный неоднородный пуассоновский процесс $\{N(x) : x \geq 0\}$ с параметром $\lambda(x)$, на интервале $(0, T]$.

Алгоритм:

1. $\lambda^* = 2\lambda$.

2. Генерируем n случайных величин $X_1, \ldots, X_n : X_i - X_{i-1} \sim \mathrm{Exp}(\lambda^*)$. Положим $i = 1, k = 0$.

3. Генерируем $U_i \sim U[0, 1]$. Если $U_i \leq \frac{\lambda(X_i)}{\lambda^*(X_i)}$, то увеличиваем k на единицу и принимаем $X_k = X_i$.

4. Увеличиваем i на единицу. Если $i \leq n$ переходим к шагу 3.

5. Возвращаем $X_1, \ldots, X_k$.

Определение 25. Случайная величина имеет распределение Парето с параметрами x_m, k, если ее функция распределения имеет вид:

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

Для построения генератора случайной величины, распределенной по Парето с параметрами x_m, k воспользуемся методом обращения:

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

Тогда, получим сл.в. ξ, распределенную по Парето:

$$X \sim U[0,1] \Longrightarrow \xi = \frac{x_m}{(1-X)^{\frac{1}{k}}}.$$

Пусть времена наступления страховых случаев $t_1, \ldots, t_{\vdots} \delta_{t_i} \sim \mathrm{Exp}(\lambda))$, где $\lambda > 0$ — интенсивность потока страховых случаев.

Пусть ущерб страхового случая s_i в момент времени t_i, яляется случайной величиной, распределенной по Парето. Ее математическое ожидание будет вычисляться следующим образом:

$$Es_i = \begin{cases} \frac{kx_m}{k-1}, & \text{при } k > 1, \\ +\infty, & \text{иначе.} \end{cases}$$

Введем величину капитала в момент времени t, как результат формулы:

$$W(t) = W(0) + ct - S(t), \quad S(t) = \sum_{i:t_i \le t} s_i,$$

где $S(t)$− сумма страховых выплат до момента времени t. Тогда время разорения зададим условием:

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

Найдем среднюю скорость прироста капитала:

$$(EW(t))' = (W(0) + ct - ES(t))' = c - \left(E \sum_{i:t_i \le t} s_i\right)' = c - (\lambda t Es_i)' = c - \frac{\lambda k x_m}{k-1}.$$

Из найденной средней скорости видно, что капитал будет увеличиваться, при $c(k-1) > \lambda k x_m$, будет уменьшаться, при $c(k-1) < \lambda k x_m$, и будет оставаться в относительном равновесии при $c(k-1) = \lambda k x_m$.

11.3 Результаты работы программы

11.3.1 СМО

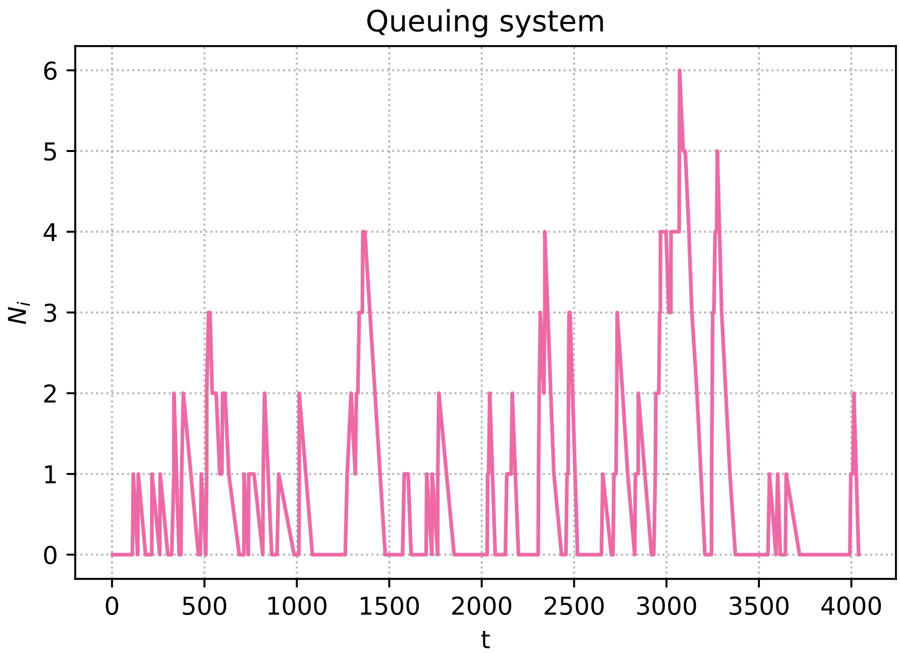

Рис. 37: СМО $\lambda = 0.05$

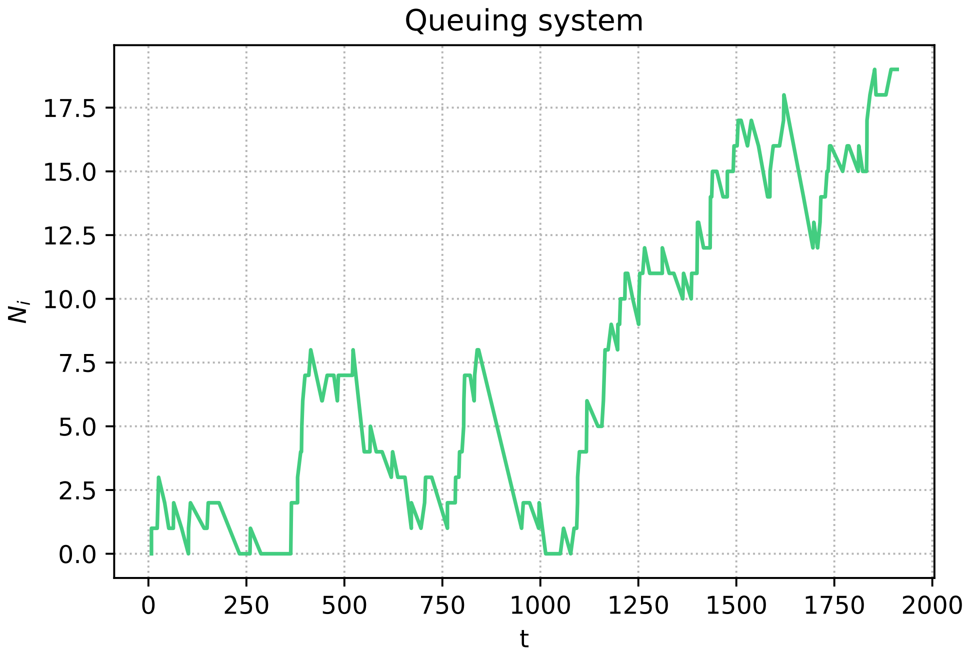

Рис. 38: СМО $\lambda = 0.1$

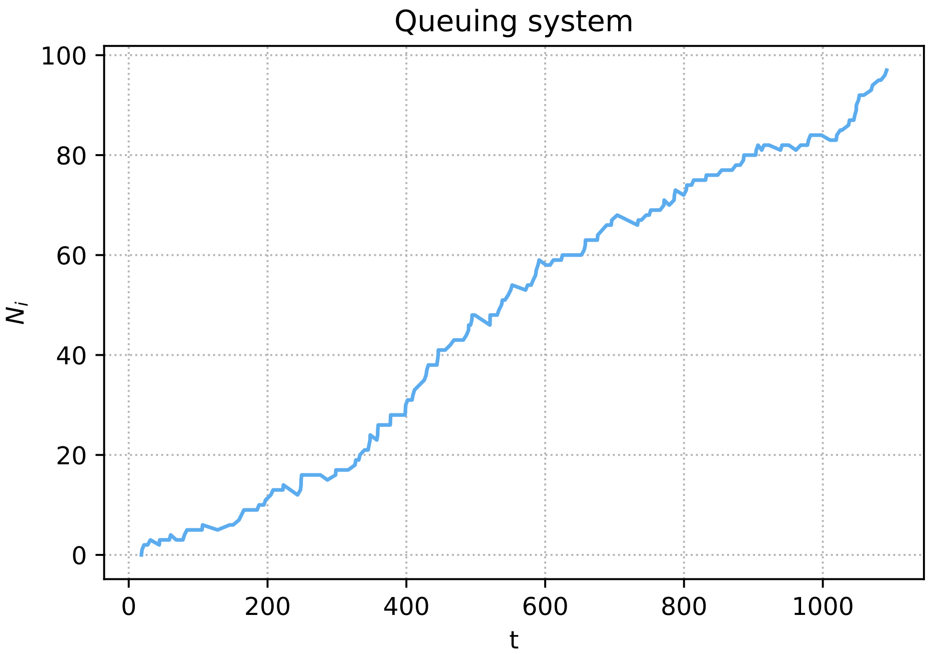

Рис. 39: СМО $\lambda = 0.2$

11.3.2 СМО с циклической интенсивностью

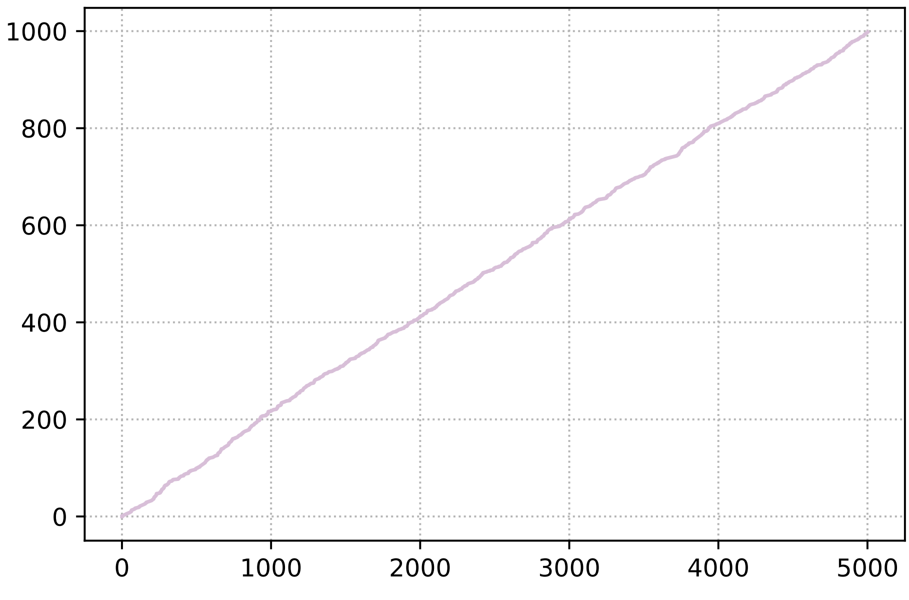

Рис. 40: СМО $\lambda = 0.5$

11.3.3 Работа страховой компании

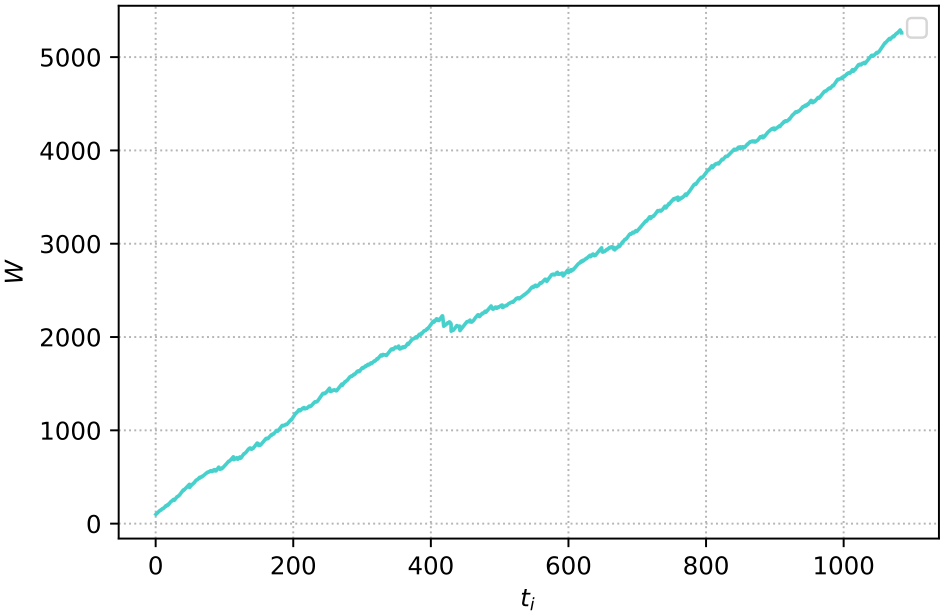

Рис. 41: $\lambda = 0.9,\ k = 3,\ x_m = 5,\ c = \frac{\lambda k x_m}{(k-1)+5}$

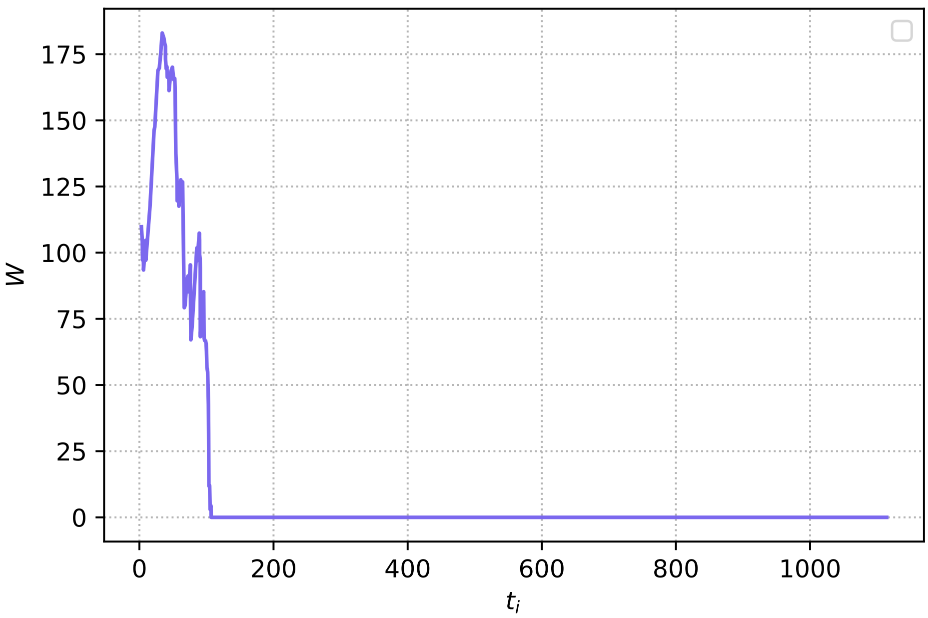

Рис. 42: $\lambda = 0.9,\ k = 3,\ x_m = 5,\ c = \frac{\lambda k x_m}{(k-1)-1}$

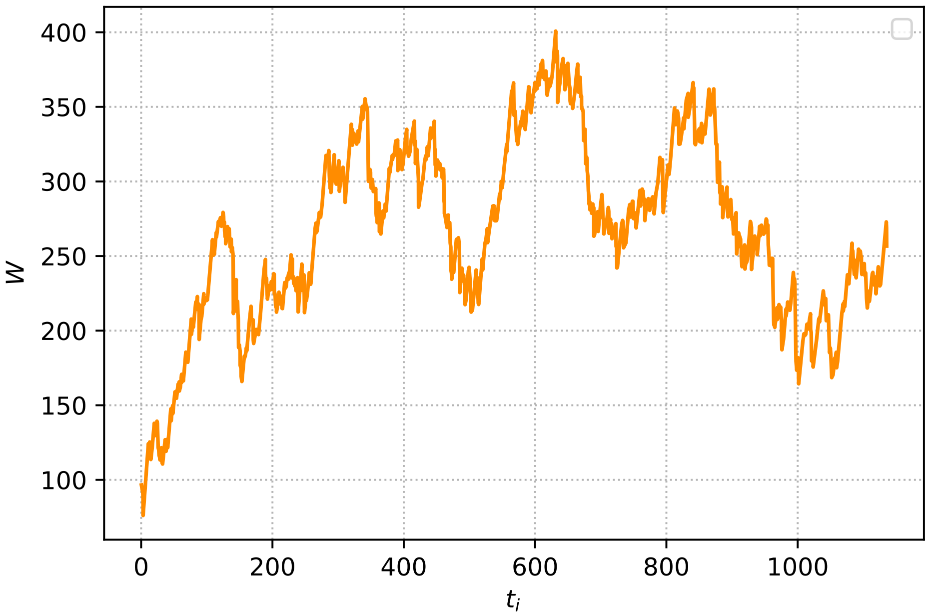

Рис. 43: $\lambda = 0.9,\ k = 3,\ x_m = 5,\ c = \frac{\lambda k x_m}{(k-1)}$

Список литературы

[1] С. Н. Смирнов. Лекции по стохастическому анализу и моделированию. 2022

[2] А. В. Арутюнов. Лекции по выпуклому и многозначному анализу. 2021

[3] К. Ю. Острем. Введение в стохастическую теорию управления. М.: Мир, 1973

[4] Ж. Неве. Математические основы теории вероятностей. 1969