



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
Я. А. Григорьев

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

Москва, 2018

Содержание

1 Задание №1	4
1.1 Постановка задачи	4
1.2 Теория и необходимая аналитика	4
1.3 Примеры работы программы	7
2 Задание №2	10
2.1 Постановка задачи	10
2.2 Теория и необходимая аналитика	10
2.3 Примеры работы программы	14
3 Задание №3	17
3.1 Постановка задачи	17
3.2 Теория и необходимая аналитика	17
3.3 Примеры работы программы	22
4 Задание №4	27
4.1 Постановка задачи	27
4.2 Теория и необходимая аналитика	27
4.3 Примеры работы программы	30
5 Задание №5	32
5.1 Постановка задачи	32
5.2 Теория и необходимая аналитика	32
5.3 Примеры работы программы	36
6 Задание №6	39
6.1 Постановка задачи	39
6.2 Теория и необходимая аналитика	39
6.3 Примеры работы программы	42
7 Задание №7	44
7.1 Постановка задачи	44
7.2 Теория и необходимая аналитика	44
7.3 Примеры работы программы	48
8 Задание №8	49
8.1 Постановка задачи	49
8.2 Теория и необходимая аналитика	49
8.3 Примеры работы программы	53
9 Задание №9	56
9.1 Постановка задачи	56
9.2 Теория и необходимая аналитика	56
9.3 Примеры работы программы	60

10 Задание №10	63
10.1 Постановка задачи	63
10.2 Теория и необходимая аналитика	63
10.3 Примеры работы программы	65
11 Задание №11	68
11.1 Постановка задачи	68
11.2 Теория и необходимая аналитика	68
11.3 Примеры работы программы	71
Список литературы	75

1 Задание №1

1.1 Постановка задачи

1. Реализовать генератор схемы Бернулли с заданной вероятностью успеха p . На основе генератора схемы Бернулли построить датчик для биномиального распределения.
2. Реализовать генератор геометрического распределения. Проверить для данного распределения свойство отсутствия памяти.
3. Рассмотреть игру в орлянку — бесконечную последовательность 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 Теория и необходимая аналитика

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

1. Отсутствие взаимного влияния.
2. Воспроизводимость испытаний (испытания производятся в сходных условиях).
3. В каждом испытании наблюдается признак, причём вероятность его проявления (успеха) равна p .

Определение 2. Говорят, что случайная величина X имеет распределение Бернулли, если она принимает значение 1 с вероятностью p и значение 0 с вероятностью $1 - p$.

Обозначение: $X \sim \text{Bern}(p)$.

Очевидно, $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$.

В нашем распоряжении имеется датчик равномерного распределения на отрезке $[0, 1]$ (будем обозначать это распределение через $U[0, 1]$). Воспользуемся этим датчиком для построения генератора схемы Бернулли. Пусть $\xi \sim U[0, 1]$, тогда положим

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

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 (интерпретация — количество испытаний) и p .

Обозначение: $Y \sim \text{B}(n, p)$.

Распределение вероятностей в случае биномиальной случайной величины имеет вид:

$$P(X = k) = C_n^k p^k (1 - p)^{n-k}, \quad k = 0, \dots, n.$$

Очевидно, чтобы реализовать датчик биномиального распределения, возвращающий m элементов, нам достаточно воспользоваться уже реализованным генератором схемы Бернулли, а именно: с помощью этого генератора получить матрицу размеров $n \times m$, а затем взять суммы по столбцам этой матрицы.

Гистограмма, иллюстрирующая работу датчика биномиального распределения для конкретных значений параметров, представлена на рисунке 1 в пункте 1.3.

Определение 4. Случайная величина Z , равная количеству «неудач» до появления первого «успеха» в схеме Бернулли с параметром p , называется случайной величиной, имеющей геометрическое распределение с параметром p . Обозначение: $Z \sim \text{Geom}(p)$.

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

Чтобы реализовать генератор геометрического распределения, будем строить схему Бернулли, состоящую из N элементов, где N выбирается так, чтобы с вероятностью $P = 95\%$ в этой схеме был хотя бы один «успех». Этого можно добиться, положив $N = \lceil \log_{1-p}(1 - P) \rceil$ (тогда получим, что $1 - P \geq (1 - p)^N$). Если же в полученной выборке не окажется «успехов», сгенерируем новую (и так до тех пор, пока не получим хотя бы один «успех»), отсчет уже будет идти от номера последнего элемента в предыдущей выборки.

Гистограмма, иллюстрирующая работу генератора геометрического распределения для конкретного параметра, представлена на рисунке 2 в пункте 1.3.

Из [1] известна следующая

Теорема 1. Если $Z \sim \text{Geom}(p)$, то

$$\mathbb{P}(Z > d + n | Z \geq d) = \mathbb{P}(Z > n)$$

для любых целых неотрицательных n и d .

Доказательство.

$$\begin{aligned} \mathbb{P}(Z > d + n | Z \geq d) &= \frac{\mathbb{P}(Z > d + n, Z \geq d)}{\mathbb{P}(Z \geq d)} = \frac{\mathbb{P}(Z > d + n)}{\mathbb{P}(Z \geq d)} = \\ &= \{\text{суммы бесконечной геометрической прогрессии}\} = \frac{(1 - p)^{d+n+1}}{(1 - p)^d} = (1 - p)^{n+1} = \mathbb{P}(Z > n) \end{aligned}$$

□

Данное свойство называется свойством отсутствия памяти. Гистограмма, иллюстрирующая данное свойство на выборке, приведена на рисунке 3 в пункте 1.3.

Ломаная, задающая поведение нормированной суммы $Y(i)$ при игре в орлянку, представлена на рисунке 4 в пункте 1.3.

Для дальнейших рассуждений нам понадобится классическая центральная предельная теорема (см., например, [2]):

Теорема 2. Пусть X_1, \dots, X_n, \dots есть бесконечная последовательность независимых одинаково распределённых случайных величин, имеющих конечно математическое ожидание и дисперсию. Обозначим последние через μ и σ^2 , соответственно. Пусть также

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

Тогда

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

по распределению при $n \rightarrow +\infty$, где $N(0, 1)$ — нормальное распределение с нулевым математическим ожиданием и стандартным отклонением, равным единице.

При игре в орлянку реализуются следующие случайные величины:

$$X_i = \begin{cases} 1, & p = 0.5, \\ -1, & p = 0.5. \end{cases}$$

Очевидно, что $\mathbb{E}X_i = 0$, $\mathbb{D}X_i = \mathbb{E}X_i^2 - (\mathbb{E}X_i)^2 = 1 - 0 = 1$. Тогда

$$Y(n) = \frac{\sum_{i=1}^n X_i - \mathbb{E}X_i n}{\sigma \sqrt{n}}$$

и по теореме 2 получаем, что $Y(n) \rightarrow N(0, 1)$ при $n \rightarrow +\infty$.

Иллюстрация доказанной сходимости представлена на рисунке 5 в пункте 1.3.

1.3 Примеры работы программы

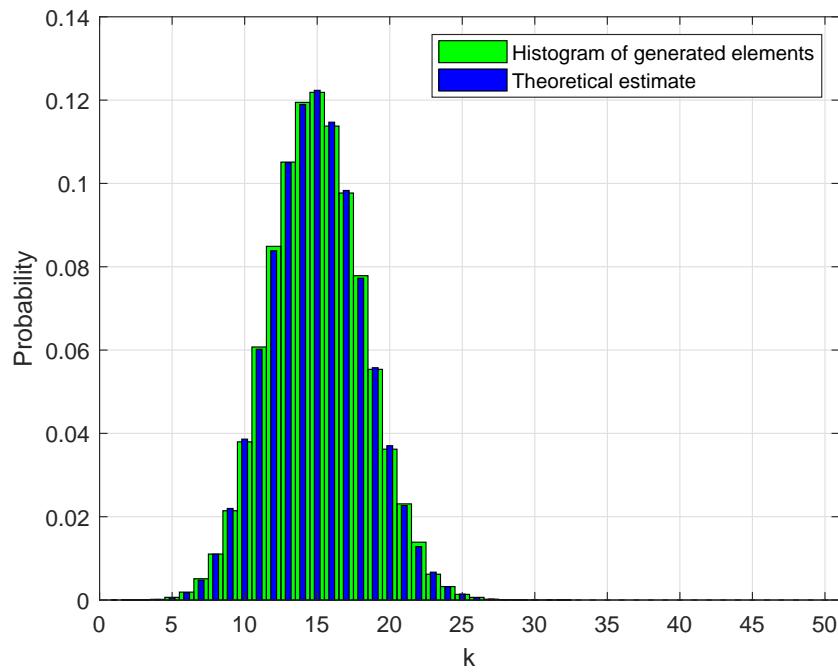


Рис. 1: Гистограмма биномиального распределения с параметрами $p = 0.3$, $n = 50$, общее количество элементов $m = 100000$

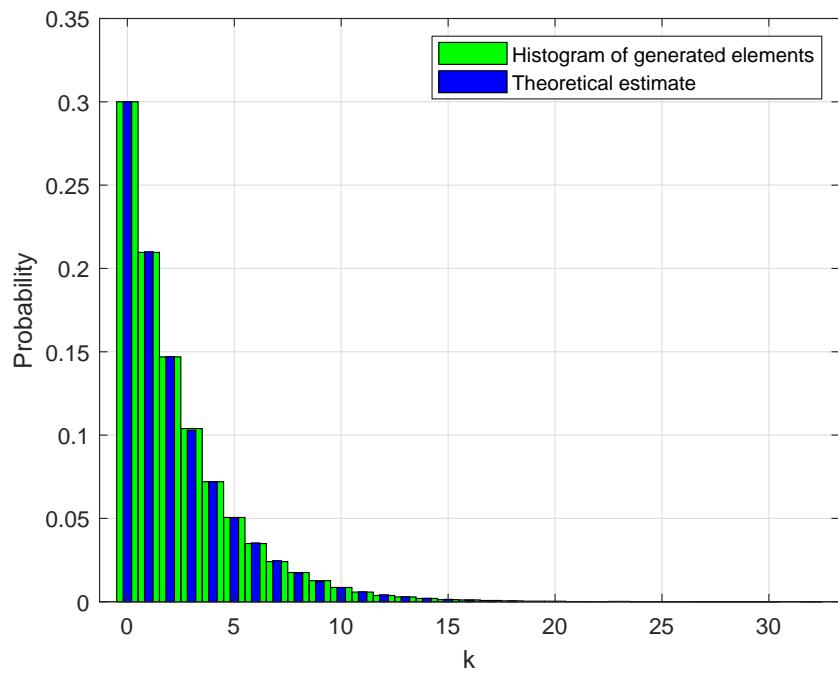


Рис. 2: Гистограмма геометрического распределения с параметром $p = 0.3$, общее количество элементов $m = 100000$

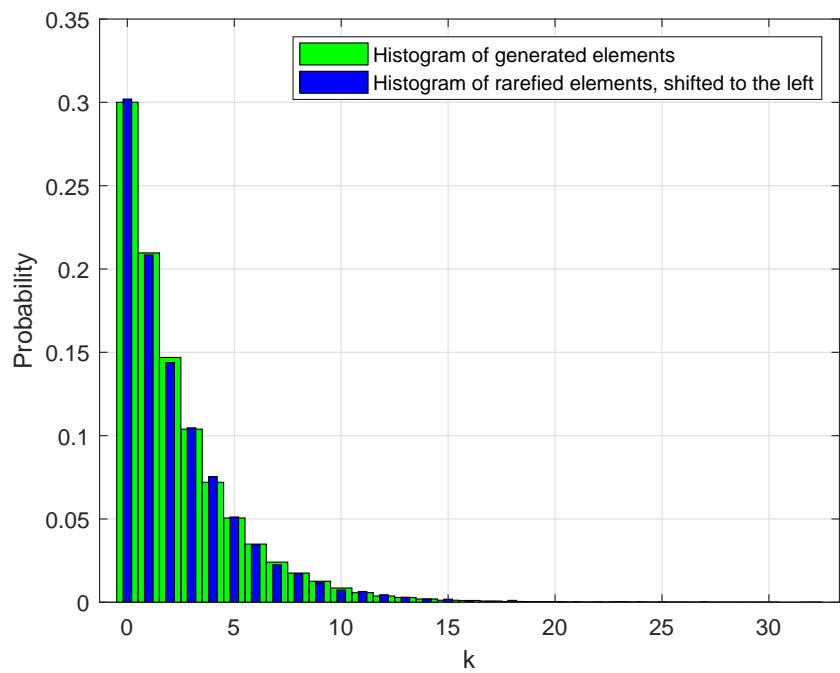


Рис. 3: Иллюстрация свойства отсутствия памяти у геометрического распределения с параметром $p = 0.3$, разделитель $d = 5$

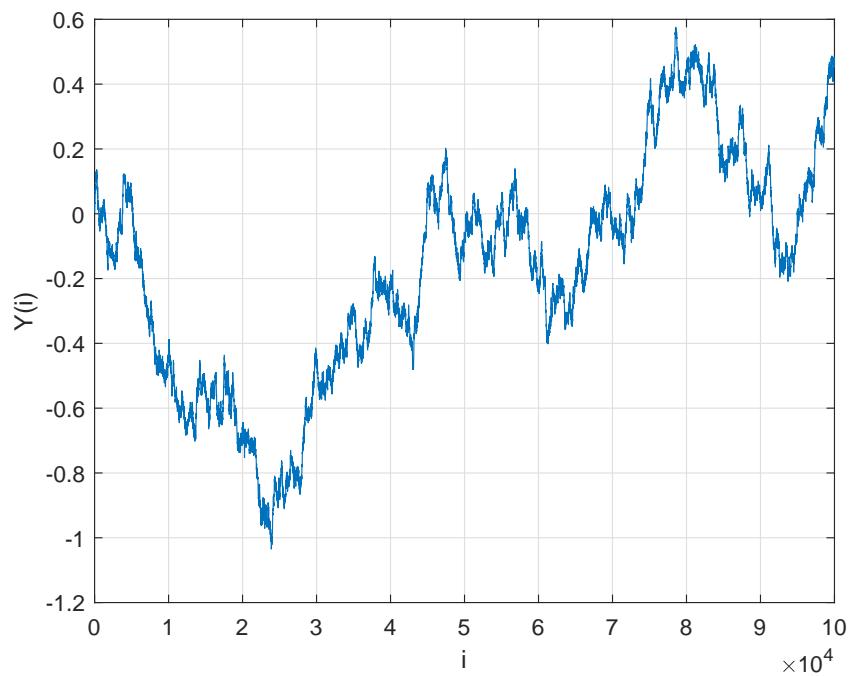


Рис. 4: Ломанная, задающая поведение нормированной суммы $Y(i)$ при игре в орлянку, количество испытаний $n = 100000$

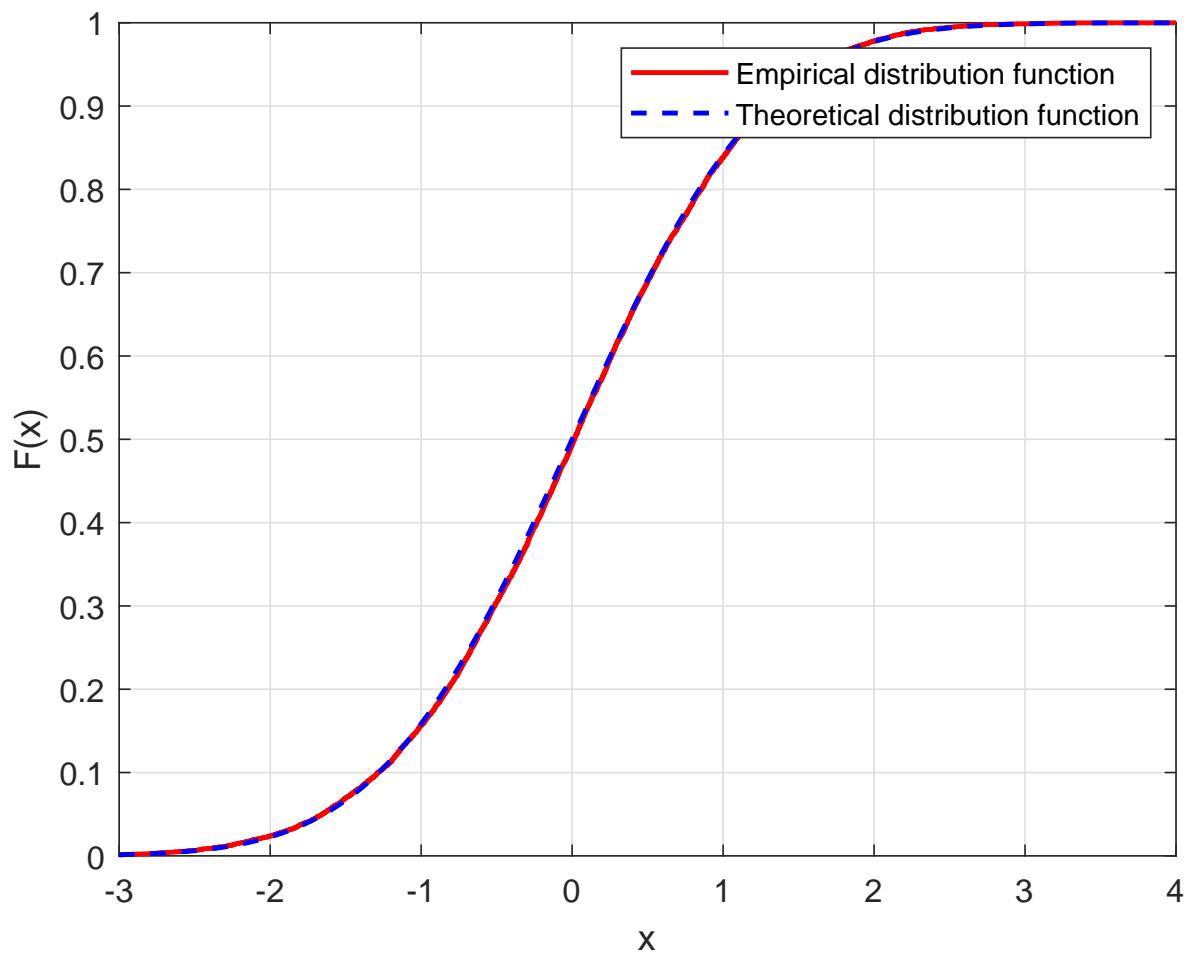


Рис. 5: Эмпирическая функция распределения, построенная на выборке $Y(n)$ длины m , где $n = 100000$, $m = 10000$, и теоретическая функция стандартного нормального распределения

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 $Y \in [0, 1/3]$ 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 Теория и необходимая аналитика

Определение 5. Пусть ξ — случайная величина (то есть действительная измеримая функция от элементарного исхода). Тогда её функцией распределения называется функция вещественного аргумента

$$F_\xi(x) = \mathbb{P}(\xi \leq x).$$

Определение 6. Функция распределения некоторой случайной величины называется сингулярной (относительно меры Лебега), если распределение этой случайной величины сосредоточено на континуальном множестве с нулевой мерой Лебега.

Определение 7. Из единичного отрезка $C_0 = [0, 1]$ удалим интервал $(1/3, 2/3)$. Оставшееся точечное множество обозначим через C_1 . Множество $C_1 = [0, 1/3] \cup [2/3, 1]$ состоит из двух отрезков; удалим теперь из каждого отрезка его среднюю третью, и оставшееся множество обозначим через C_2 . Повторив эту процедуру опять, удаляя средние трети у всех четырех отрезков, получаем C_3 . Далее таким же образом получаем последовательность замкнутых множеств $C_0 \supset C_1 \supset C_2 \supset \dots$ Пересечение

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

называется канторовым множеством.

Обозначим его через K .

Из построения канторова множества вытекает, что любую точку, принадлежащую этому множеству, в троичной системе можно записать, не используя единицы (то есть используя только нули и двойки). Более того, объединение всех таких точек и образует канторово множество.

Очевидным образом построив взаимнооднозначное соответствие между точками отрезка $[0, 1]$ и канторовым множеством (представив точки первого множества в двоичной системе, а второго — в троичной, и отождествив единицы и двойки), получим, что мощность канторова множества равна континууму.

Посчитаем меру Лебега канторова множества. Вычтем длины отрезков, которые мы выкинули, из длины целого отрезка $[0, 1]$:

$$\mu(K) = 1 - \sum_{k=1}^{+\infty} \frac{2^{k-1}}{3^k} = 1 - \frac{1}{3} \cdot \frac{1}{1 - \frac{2}{3}} = 1 - 1 = 0.$$

Тем самым, получаем, что любая функция распределения, точки роста которой совпадают с точками канторова множества, является сингулярной. Наиболее часто встречающимся примером такой функции является канторова лестница. Построим её как предел функциональной последовательности.

Итак, пусть

$$F_1(x) = \begin{cases} 1/2, & x \in (1/3, 2/3), \\ 0, & x = 0, \\ 1, & x = 1, \\ \text{линейная интерполяция,} & \text{иначе;} \end{cases}$$

$$F_2(x) = \begin{cases} 1/2, & x \in (1/3, 2/3), \\ 1/4, & x \in (1/9, 2/9), \\ 3/4, & x \in (7/9, 8/9), \\ 0, & x = 0, \\ 1, & x = 1, \\ \text{линейная интерполяция,} & \text{иначе.} \end{cases}$$

И так далее. Продолжая процесс, получим последовательность $F_n(x)$, $n = \overline{1, +\infty}$, которая сходится к непрерывной функции $F(x)$, называемой канторовой лестницей. Очевидно, точки роста данной функции образуют канторово множество K .

Чтобы построить датчик распределения, имеющий в качестве функции распределения канторову лестницу, нам необходимо, во-первых, задать точность вычислений $\varepsilon = 10^{-10}$, а во-вторых, воспользоваться уже написанным генератором схемы Бернулли. Как мы уже выяснили, точки канторова множества в троичной системе задаются только нулями и двойками, следовательно, представимы в следующем виде:

$$x = \sum_{k=1}^{+\infty} \frac{2\xi_k}{3^k},$$

где $\xi_k \sim \text{Bern}(1/2)$. Очевидно, что перебрать все точки канторова множества невозможно, поэтому переберём только те, которые получаются из указанного ряда при взятии первых N слагаемых. Как подобрать N для обеспечения нужной точности ε ? Посчитаем отбрасываемый остаток ряда:

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

то есть

$$N = \left\lceil \log_3 \frac{1}{\varepsilon} \right\rceil.$$

Тем самым, чтобы получить требуемый датчик, мы возьмём N бернуллиевских случайных величин ξ_1, \dots, ξ_N , реализуем их и получим конкретную точку канторова множества.

По сгенерированной выборке построим так называемую эмпирическую функцию распределения. Для этого в точках выборки будем повышать значение этой функции на $1/n$, где n — длина выборки, левее самой левой точки значение этой функции будет равно нулю, а правее самой правой точки — единице.

Чтобы получить точное значение канторовой лестницы в сгенерированной точке, представим вектор $(\xi_1 \dots \xi_N)_2 = l$ в виде двоичного числа, которое равно l , и значение канторовой лестницы

в рассматриваемой точке, очевидно, будет равно

$$F(x) = \frac{l}{2^N}.$$

Можно посчитать ещё проще:

$$F(x) = \sum_{k=1}^N \frac{\xi_k}{2^k}.$$

График эмпирической и теоретической функции распределения для данного случая представлен на рисунке 6 в пункте 2.3.

Для проверки корректности работы построенного датчика воспользуемся критерием Колмогорова, а именно: посчитаем максимальное по модулю отклонение эмпирической функции распределения от теоретической в точках выборки. Обозначим это число через D_n . Для проверки простой гипотезы H_0 о соответствии выборки известному закону распределения с заданным уровнем значимости α рассчитаем p -value по статистике $\sqrt{n}D_n$, то есть найдём значение функции распределения Колмогорова в заданной точке. Для этого используем следующее представление для функции распределения Колмогорова:

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

Гипотеза принимается, если p -value, рассчитываемое как $p = 1 - F_K(\sqrt{n}D_n)$ превышает заданный уровень значимости α .

Данный критерий основан на применении следующей теоремы Колмогорова.

Теорема 3. Пусть $X_1 \dots X_n, \dots$ — бесконечная выборка из распределения, задаваемого непрерывной функцией распределения $F(x)$. Пусть $F_n(x)$ — выборочная функция распределения, построенная на первых n элементах выборки. Тогда

$$\sqrt{n} \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \rightarrow K$$

по распределению при $n \rightarrow +\infty$, где K — случайная величина, имеющая распределение Колмогорова.

Доказательство этой теоремы можно найти, например, в [7].

Результаты расчётов критерия Колмогорова для различных выборок приведены в таблице 1 в пункте 2.3.

Рассмотрим теперь свойство симметричности относительно $1/2$ случайных величин X и $1 - X$. Рассмотрим случайную величину $1 - X$:

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

Так как $\eta_k \sim \text{Bern}(1/2)$, то X и $(1 - X)$ распределены одинаково.

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

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

Проверим последние упомянутые свойства эмпирически. Для этого воспользуемся критерием однородности Смирнова. Имеются две эмпирические функции распределения, необходимо найти максимум их абсолютной поточечной разности. Обозначив полученное значение через D_{nm} , где n, m — длины выборок, и посчитав p -value по статистике $\sqrt{\frac{nm}{n+m}}D_{nm}$, как это было проделано ранее, для заданного уровня значимости α проверим гипотезу H_0 о том, что обе выборки соответствуют однаковому закону распределения.

Иллюстрации обоих свойств приведены на рисунках 7-8 в пункте 2.3.

Критерий Смирнова основан на применении следующей теоремы Смирнова.

Теорема 4. Пусть $F_{1,n}(x), F_{2,m}(x)$ — эмпирические функции распределения, построенные по независимым выборкам объёмом n и m случайной величины ξ . Тогда, если $F(x) \in C^1(\mathbb{R})$, то

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

где $D_{n,m} = \sup_{x \in \mathbb{R}} |F_{1,n}(x) - F_{2,m}(x)|$.

То есть, аналогично теореме Колмогорова, статистика $\sqrt{\frac{nm}{n+m}}D_{n,m}$ будет по распределению стремиться к случайной величине, имеющей распределение Колмогорова. Доказательство данной теоремы можно найти в [7].

Результаты расчётов критерия Смирнова для данных свойств представлены в таблицах 2–3 в пункте 2.3.

Посчитаем теперь математическое ожидание и дисперсию рассматриваемой случайной величины. Из симметричности функции распределения относительно прямой $x = 1/2$ следует, что $\mathbb{E}\xi = 1/2$, однако можно вычислить математическое ожидание, исходя из других соображений, которые адаптируются и для моментов более высоких порядков.

Как было выяснено раньше, F самоподобна, то есть при $0 < x < 1/3$ выполнено соотношение $F(x) = F(3x)/2$, а при $2/3 < x < 1$ имеет место равенство $F(x) = 1/2 + F(3x - 2)/2$. Поэтому

$$\begin{aligned} \mathbb{E}\xi &= \int_{-\infty}^{+\infty} xdF(x) = \int_0^{1/3} xdF(x) + \int_{2/3}^1 xdF(x) = \frac{1}{2} \int_0^{1/3} xdF(3x) + \frac{1}{2} \int_{2/3}^1 xd(1/2 + F(3x - 2)) = \\ &= \{y = 3x \text{ в первом интеграле, } y = 3x - 2 \text{ во втором интеграле}\} = \\ &= \frac{1}{2} \int_0^1 \frac{y}{3} dF(y) + \frac{1}{2} \int_0^1 \frac{y+2}{3} dF(y) = \frac{1}{6} \int_0^1 ydF(y) + \frac{1}{6} \int_0^1 ydF(y) + \frac{1}{3} \int_0^1 dF(y) = \frac{1}{3}\mathbb{E}\xi + \frac{1}{3}. \end{aligned}$$

Решая уравнение, получаем, что $\mathbb{E}\xi = 1/2$.

Второй момент случайной величины ξ вычисляется аналогичным образом с использованием свойства самоподобия и уже вычисленного значения $\mathbb{E}\xi = 1/2$:

$$\begin{aligned} \mathbb{E}\xi^2 &= \int_0^{1/3} x^2 dF(x) + \int_{2/3}^1 x^2 dF(x) = \frac{1}{2} \int_0^1 \left(\frac{y}{3}\right)^2 dF(y) + \frac{1}{2} \int_0^1 \left(\frac{y+2}{3}\right)^2 dF(y) = \\ &= \frac{1}{9}\mathbb{E}\xi^2 + \frac{2}{9}\mathbb{E}\xi + \frac{2}{9} = \frac{1}{9}\mathbb{E}\xi^2 + \frac{1}{9} + \frac{2}{9}, \end{aligned}$$

откуда $\mathbb{E}\xi^2 = 3/8$ и дисперсия $\mathbb{D}\xi = 3/8 - (1/2)^2 = 1/8$.

Графики зависимости фактических математических ожиданий и дисперсий от размера выборки приведены на рисунках 9-10 в пункте 2.3.

2.3 Примеры работы программы

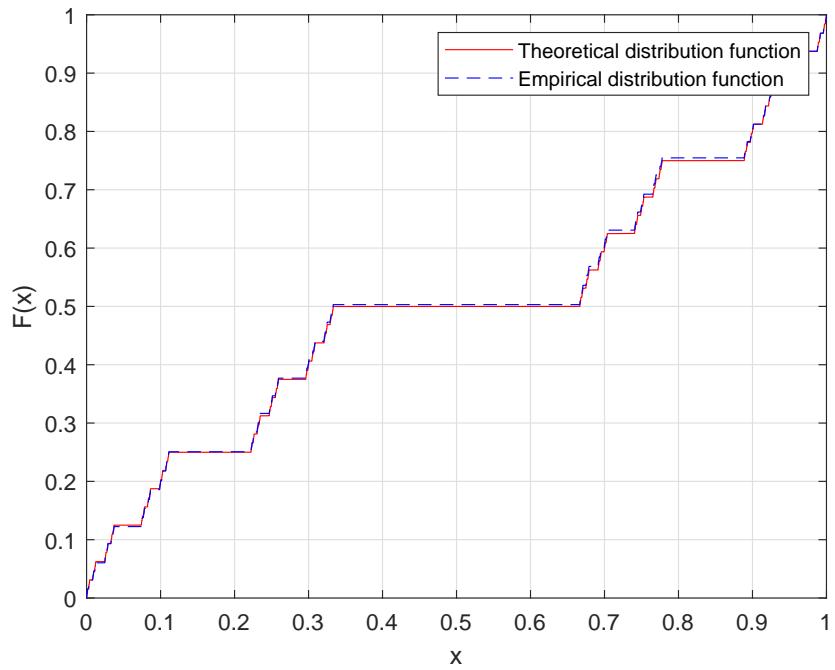


Рис. 6: Эмпирическая и теоретическая функции сингулярного распределения (канторова лестница), общее количество элементов $n = 20000$

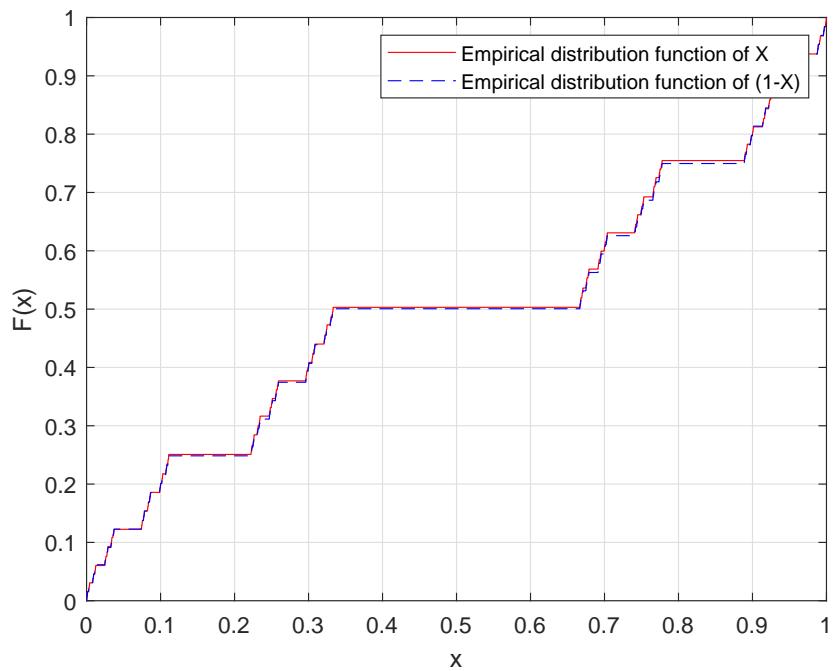


Рис. 7: Свойство симметричности сингулярного распределения относительно $1/2$, общее количество элементов: $n = 20000$

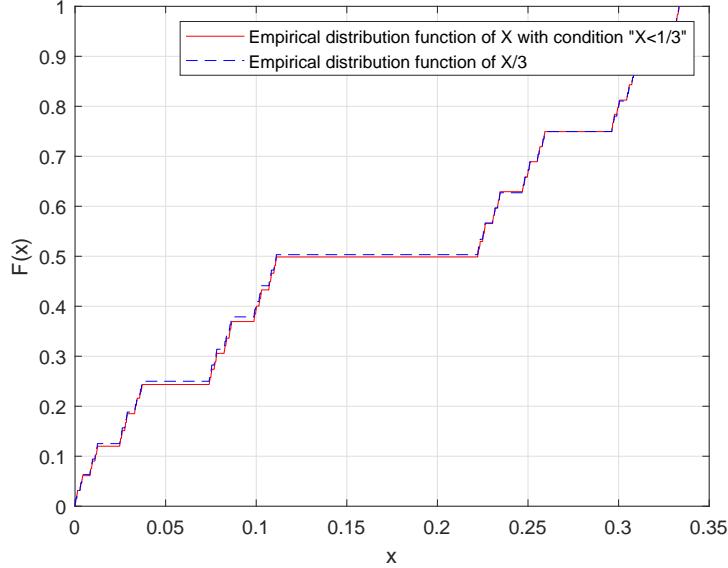


Рис. 8: Свойство самоподобия сингулярной случайной величины: $Y|(Y \leq 1/3) \sim Y/3$, общее количество элементов $n = 20000$

запуски\объём выборки	$n = 10^3$	$n = 10^4$
$m = 10^3$	94, 7%	94, 9%
$m = 10^4$	94, 8%	94, 87%

Таблица 1: Значения в таблице показывают, какая доля запусков привела к положительному с точки зрения критерия Колмогорова результату. Уровень значимости $\alpha = 5\%$. Проверялась гипотеза о принадлежности построенного распределения сингулярному закону, описанному в пункте 2.2.

запуски\объём выборки	$n = 10^3$	$n = 10^4$
$m = 10^3$	94, 9%	96, 2%
$m = 10^4$	95, 25%	95, 02%

Таблица 2: Обозначения и уровень значимости аналогичны таблице 1 (здесь рассматривается критерий Смирнова). Проверялось свойство симметричности сингулярного распределения относительно $1/2$.

запуски\объём выборки	$n = 10^3$	$n = 10^4$
$m = 10^3$	95, 3%	95%
$m = 10^4$	95, 48%	95, 13%

Таблица 3: Обозначения и уровень значимости аналогичны таблице 2. Проверялось свойство самоподобия сингулярного распределения (подробнее см. пункт 2.2).

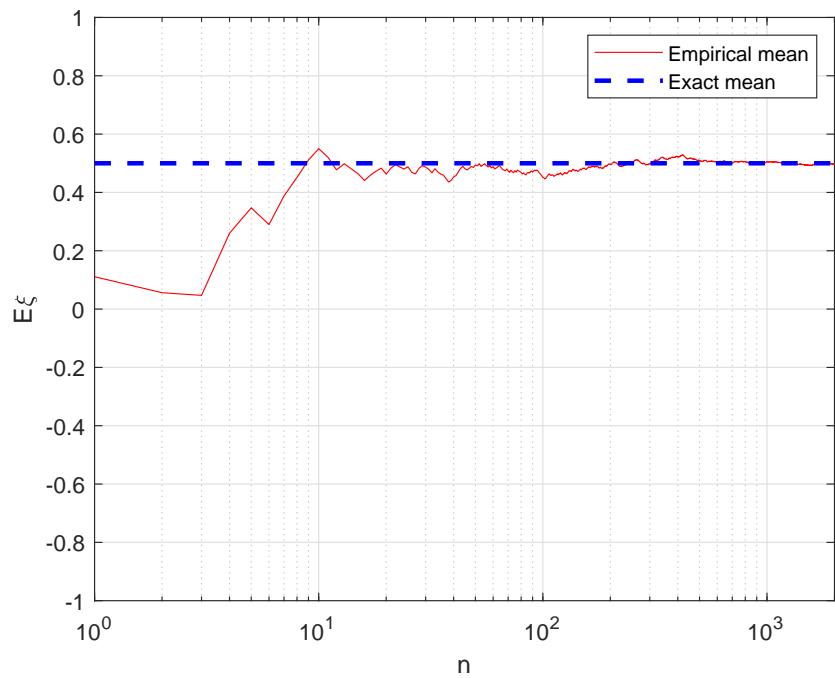


Рис. 9: Иллюстрация сходимости эмпирического математического ожидания к теоретическому при увеличении размера выборки

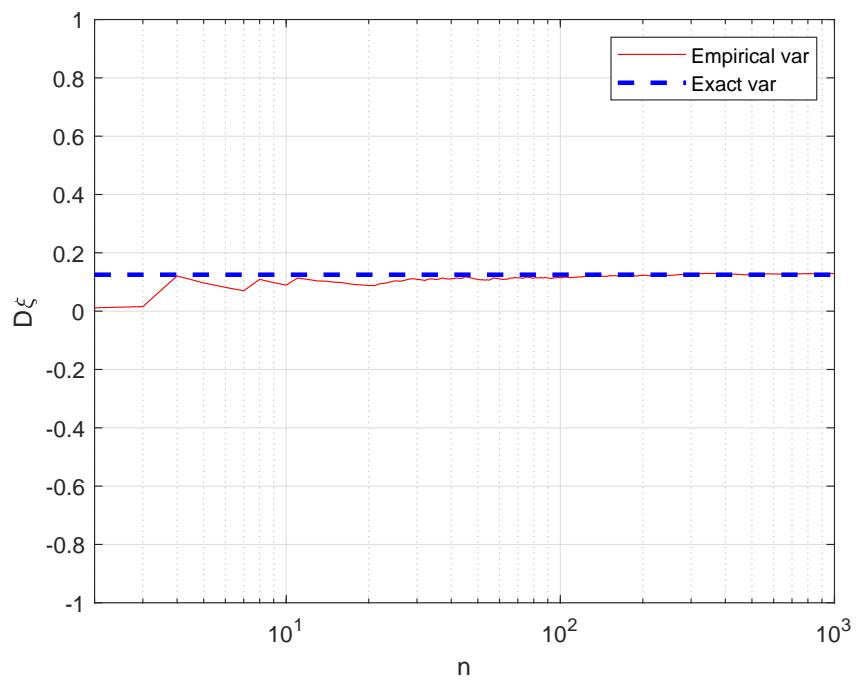


Рис. 10: Иллюстрация сходимости эмпирической дисперсии к теоретической при увеличении размера выборки

3 Задание №3

3.1 Постановка задачи

- Построить датчик экспоненциального распределения. Проверить для данного распределения свойство отсутствия памяти. Пусть X_1, X_2, \dots, X_n — независимо экспоненциально распределённые случайные величины с параметрами $\lambda_1, \lambda_2, \dots, \lambda_n$ соответственно. Найти распределение случайной величины $Y = \min(X_1, X_2, \dots, X_n)$.
- На основе датчика экспоненциального распределения построить датчик пуассоновского распределения.
- Построить датчик пуассоновского распределения как предел биномиального распределения. С помощью критерия хи-квадрат Пирсона убедиться, что получен датчик распределения Пуассона.
- Построить датчик стандартного нормального распределения методом моделирования случайных величин парами с переходом в полярные координаты. Проверить при помощи t -критерия Стьюдента равенство математических ожиданий, а при помощи критерия Фишера равенство дисперсий.

3.2 Теория и необходимая аналитика

Определение 8. Случайная величина ξ имеет экспоненциальное распределение с параметром $\lambda > 0$, если её функция распределения имеет вид:

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

Обозначение: $\xi \sim \text{Exp}(\lambda)$.

Из [3] известна следующая

Теорема 5. Пусть функция $F(x)$ непрерывна и монотонно возрастает на $x \in \mathbb{R}$, причём $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow +\infty} F(x) = 1$, случайная величина $Y \sim U[0, 1]$. Тогда случайная величина $X = F^{-1}(Y)$ имеет функцию распределения $F_X(x) = F(x)$.

Применим эту теорему для построения датчика экспоненциального распределения. Имеем:

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

Применяя теорему 5, получим, что если $Y \sim U[0, 1]$, то

$$X = -\frac{\log(1-Y)}{\lambda} \sim \text{Exp}(\lambda).$$

Заметим, что если $Y \sim U[0, 1]$, то $1 - Y \sim U[0, 1]$, и выражение для экспоненциальной случайной величины можно записать проще:

$$X = -\frac{\log(Y)}{\lambda} \sim \text{Exp}(\lambda).$$

Иллюстрация правильности работы датчика приведена на рисунке 11 в пункте 3.3.

Как известно из [1], экспоненциальное распределение является пределом геометрического распределения и так же обладает свойством отсутствия памяти, то есть получаем, что

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

для любых неотрицательных x, t .

Доказательство.

$$\begin{aligned} \mathbb{P}(X > t + x | X \geq t) &= \frac{\mathbb{P}(X > t + x, X \geq t)}{\mathbb{P}(X \geq t)} = \frac{\mathbb{P}(X > t + x)}{\mathbb{P}(X \geq t)} = \{F_X(x) \text{ непрерывна}\} = \\ &= \frac{1 - F_X(t + x)}{1 - F_X(t)} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} = 1 - F_X(x) = \mathbb{P}(X > x). \end{aligned}$$

□

Иллюстрация свойства отсутствия памяти у экспоненциального распределения представлена на рисунке 12 в пункте 3.3.

Докажем следующее

Утверждение 1. Пусть X_1, X_2, \dots, X_n — независимые экспоненциально распределённые случайные величины с параметрами $\lambda_1, \lambda_2, \dots, \lambda_n$ соответственно. Тогда случайная величина $Y = \min(X_1, X_2, \dots, X_n) \sim \text{Exp}\left(\sum_{i=1}^n \lambda_i\right)$.

Доказательство.

$$\begin{aligned} F_Y(x) &= \mathbb{P}(Y \leq x) = 1 - \mathbb{P}(Y > x) = 1 - \mathbb{P}(\min(X_1, X_2, \dots, X_n) > x) = \\ &= 1 - \mathbb{P}(X_1 > x, X_2 > x, \dots, X_n > x) = \{X_1, X_2, \dots, X_n \text{ независимы}\} = \\ &= 1 - \mathbb{P}(X_1 > x) \cdot \mathbb{P}(X_2 > x) \cdot \dots \cdot \mathbb{P}(X_n > x) = \\ &= 1 - (1 - F_{X_1}(x)) \cdot (1 - F_{X_2}(x)) \cdot \dots \cdot (1 - F_{X_n}(x)) = 1 - e^{-\lambda_1 x} \cdot e^{-\lambda_2 x} \cdot \dots \cdot e^{-\lambda_n x} = 1 - e^{(\sum_{i=1}^n \lambda_i)x}. \end{aligned}$$

□

Иллюстрация доказанного свойства представлено на рисунке 17 в пункте 3.3.

Определение 9. Случайная величина ξ имеет распределение Пуассона с положительным параметром λ , если

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

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

Из [2] известна следующая

Теорема 6. Пусть $\eta_1, \eta_2, \dots, \eta_n, \dots$ — независимые случайные величины, и пусть

$$\eta_i \sim \text{Exp}(\lambda), \quad i = \overline{1, \infty}.$$

Если

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

то случайная величина $\xi \sim \text{Pois}(\lambda)$. В случае $\eta_1 > 1$ полагаем $\xi = 0$.

Исходя из теоремы 6, получаем следующий алгоритм построения датчика пуассоновского распределения. Будем генерировать экспоненциальные случайные величины до тех пор, пока их сумма не превзойдёт единицы. Последний номер случайной величины, при которой сумма всё еще меньше единицы, и будет нужным числом.

Иллюстрация работы датчика пуассоновского распределения на основе датчика экспоненциального распределения представлена на рисунке 13 в пункте 3.3.

Теперь реализуем датчик пуассоновского распределения как предел биномиального распределения. Из [2] известна следующая

Теорема 7. Пусть случайная величина X имеет биномиальное распределение с параметрами n и p . Обозначим $np = \lambda = \text{const}$. Тогда при $n \rightarrow +\infty$

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

Датчик пуассоновского распределения как предела биномиального можно построить следующим образом. Пусть $N = 10000$ — фиксированное достаточно большое число, тогда из условия $\lambda = Np$ однозначно определяется число $p = \lambda/N$. Далее генерируем случайную величину $\xi \sim \text{B}(N, p)$ и приближённо принимаем её за пуассоновскую.

Иллюстрация работы датчика пуассоновского распределения на основе датчика биномиального распределения представлена на рисунке 14 в пункте 3.3.

Для проверки правильности построенного датчика воспользуемся критерием хи-квадрат Пирсона. Его суть заключается в следующем. Пусть X_1, \dots, X_n — некоторая выборка. Необходимо проверить, принадлежит ли эта выборка заданному распределению. Проверяемая гипотеза H_0 гласит, что выборка принадлежит пуассоновскому распределению. В критерии хи-квадрат Пирсона предполагается разбиение области значений случайной величины на непересекающиеся интервалы. В нашем случае достаточно взять целые точки, так как пуассоновская случайная величина принимает только целые неотрицательные значения. Обозначим через n_i количество элементов выборки, равных i . Теоретическое значение вероятности выпадения значения i в распределении Пуассона равно

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

Пусть k — максимальное значение в выборке. Тогда статистика критерия согласия хи-квадрат Пирсона в нашем случае имеет вид:

$$X_n^2 = n \sum_{i=0}^k \frac{(n_i/n - p_i)^2}{p_i}.$$

Гипотеза H_0 верна, если вычисленное значение X_n^2 не превосходит критического значения $\chi_{r,\alpha}^2$, где $r = k - 1$ — количество степеней свободы, α — заданный уровень значимости.

Результат расчётов критерия хи-квадрат Пирсона представлен в таблице 4 в пункте 3.3.

Более подробно о распределении хи-квадрат (в частности, его определение) будет рассказано в пункте 5.2.

Построим теперь датчик стандартного нормального распределения методом моделирования случайных величин парами с переходом в полярные координаты. Для этого необходимо доказать следующее

Утверждение 2. Пусть $\omega \sim \text{Exp}(1/2)$, $\phi \sim U[0, 2\pi]$. Тогда случайные величины $\xi = \sqrt{\omega} \cos \phi$ и $\eta = \sqrt{\omega} \sin \phi$ имеют стандартное нормальное распределение.

Доказательство. Достаточно доказать, что совместное распределение ξ и η совпадает с совместным распределением $\sqrt{\omega} \cos \phi$ и $\sqrt{\omega} \sin \phi$:

$$\begin{aligned}\mathbb{P}(\xi < x, \eta < y) &= \int_{-\infty}^x \int_{-\infty}^y \frac{1}{2\pi} e^{-\frac{x_1^2+x_2^2}{2}} dx_1 dx_2 = \{\text{полярные координаты}\} = \\ &= \iint_{\substack{r \cos \phi < x \\ r \sin \phi < y}} \frac{r}{2\pi} e^{-\frac{r^2}{2}} dr d\phi = \{\omega = r^2\} = \iint_{\substack{\sqrt{\omega} \cos \phi < x \\ \sqrt{\omega} \sin \phi < y}} \frac{1}{2} e^{-\frac{\omega}{2}} \frac{1}{2\pi} d\omega d\phi.\end{aligned}$$

□

Иллюстрация работы датчика стандартного нормального распределения, построенного указанным способом, представлена на рисунках 15-16 в пункте 3.3.

Проверим равенство математических ожиданий построенных случайных величин ξ и η , используя t -критерий Стьюдента. Его суть заключается в следующем.

Обозначим через M_1 и M_2 математические ожидания первой и второй выборки соответственно. Рассмотрим разность выборочных средних $\Delta = \bar{\xi} - \bar{\eta}$. Очевидно, если нулевая гипотеза о равенстве математических ожиданий выполнена, то $\mathbb{E}(\Delta) = M_1 - M_2 = 0$. Так как известно, что $\mathbb{D}(\xi) = \mathbb{D}(\eta) = 1$, то так как

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

и аналогично для $\bar{\eta}$, то

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

Используя несмешённую 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 несмешённую оценку дисперсии разности выборочных средних:

$$s_\Delta^2 = \frac{s_1^2 + s_2^2}{n}.$$

Следовательно (см., например, [2]), t -статистика для проверки нулевой гипотезы равна

$$t = \frac{\bar{\xi} - \bar{\eta}}{\sqrt{s_1^2 + s_2^2}} \sqrt{n}.$$

Число степеней свободы $r = 2n - 2$. Далее, если полученное значение статистики t превосходит критическое значение $t_{\alpha, r}$ для заданного уровня значимости α , то нулевая гипотеза отвергается.

Результаты расчётов t -критерия Стьюдента представлены в таблице 5 в пункте 3.3.

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

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

Критерий Фишера утверждает, что если $F < F_{\alpha/2}(n - 1, m - 1)$ или $F > F_{1-\alpha/2}(n - 1, m - 1)$, то нулевая гипотеза о равенстве дисперсий отвергается, где $F_{\alpha}(n - 1, m - 1)$ есть α -квантиль распределения Фишера с $(n - 1)$ и $(m - 1)$ степенями свободы.

Результаты расчётов критерия Фишера представлены в таблице 6 в пункте 3.3.

Более подробно о распределении Стьюдента, выборочном среднем и выборочной дисперсии будет рассказано в пункте 5.2.

3.3 Примеры работы программы

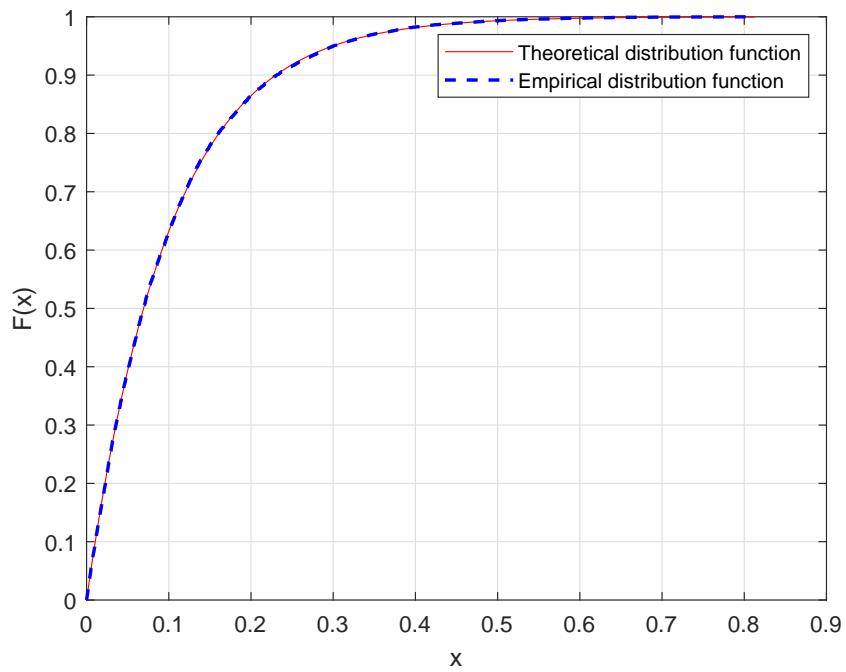


Рис. 11: Эмпирическая и теоретическая функции экспоненциального распределения, параметр $\lambda = 10$, общее количество элементов $n = 10000$

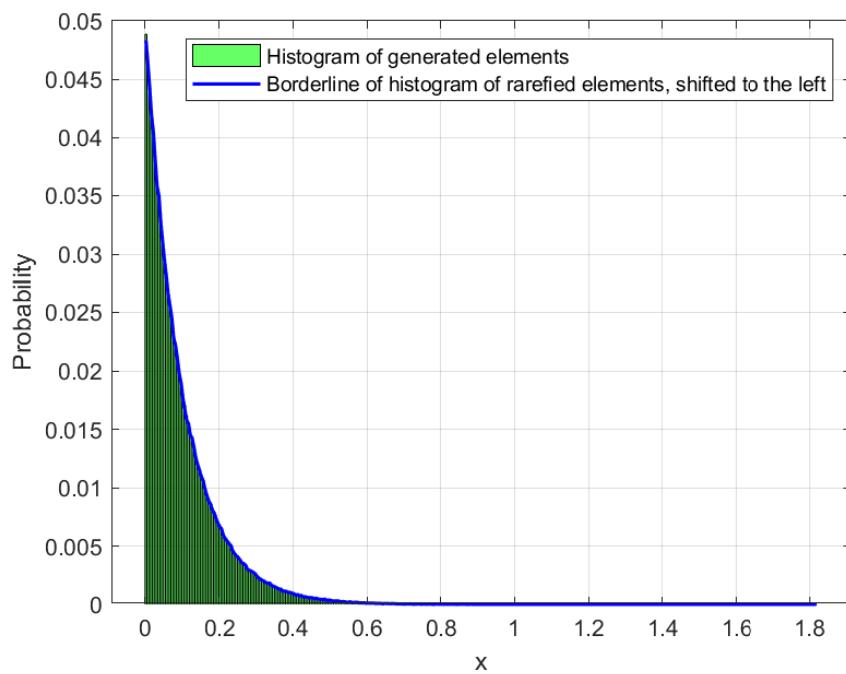


Рис. 12: Иллюстрация свойства отсутствия памяти у экспоненциального распределения, параметр $\lambda = 10$, разделятель $t = 0.3$, общее количество элементов $n = 10000000$

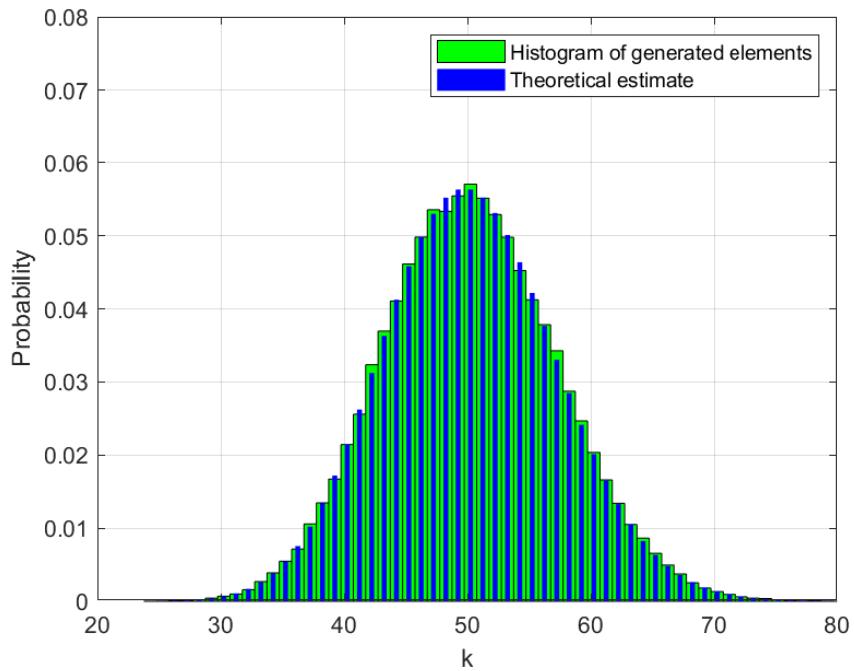


Рис. 13: Иллюстрация работы датчика пуассоновского распределения, использующего датчик экспоненциального распределения, параметр $\lambda = 50$, общее количество элементов $n = 100000$

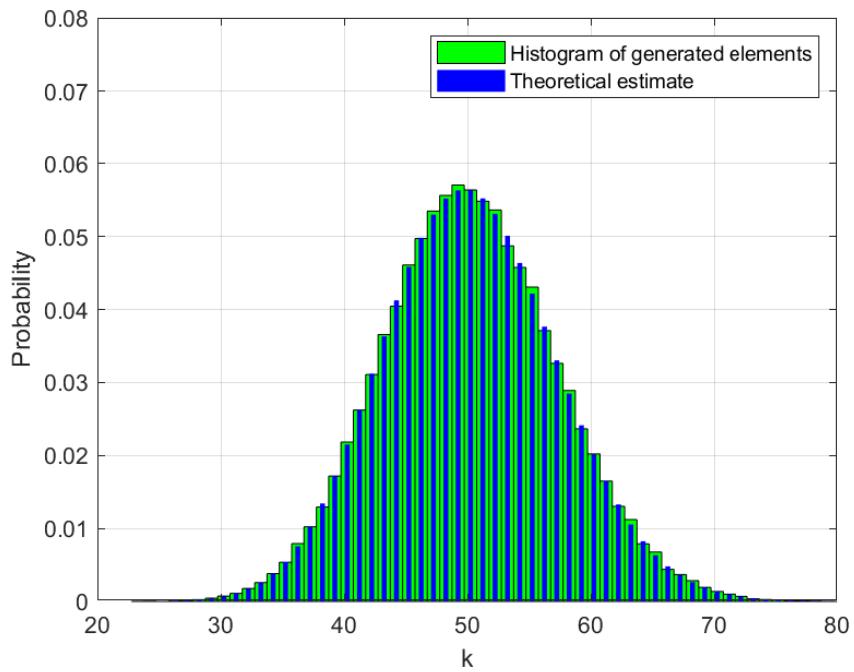


Рис. 14: Иллюстрация работы датчика пуассоновского распределения, использующего датчик биномиального распределения, параметр $\lambda = 50$, номер используемого элемента последовательности $N = 10000$, общее количество элементов $n = 100000$

запуски \ объём выборки	$n = 10^3$	$n = 10^4$
$m = 10^3$	95, 8%	95, 5%
$m = 10^4$	95, 13%	95, 03%

Таблица 4: Для генерации пуассоновского вектора брался элемент биномиального распределения под номером $N = 1000$. Значения в таблице показывают, какая доля запусков привела к положительному с точки зрения критерия хи-квадрат Пирсона результату. Уровень значимости $\alpha = 5\%$. Проверялась гипотеза о принадлежности построенного распределения пуассоновскому закону, описанному в пункте 3.2.

запуски \ объём выборки	$n = 10^3$	$n = 10^4$
$m = 10^3$	95, 1%	94, 6%
$m = 10^4$	94, 8%	95, 23%

Таблица 5: Обозначения и уровень значимости аналогичны таблице 4 (здесь рассматривается t -критерий Стьюдента). Проверялась гипотеза о равенстве математических ожиданий двух случайных величин, сгенерированных методом моделирования пар.

запуски \ объём выборки	$n = 10^3$	$n = 10^4$
$m = 10^3$	94, 5%	94, 9%
$m = 10^4$	94, 86%	94, 89%

Таблица 6: Обозначения и уровень значимости аналогичны таблице 4 (здесь рассматривается критерий Фишера). Проверялась гипотеза о равенстве дисперсий двух случайных величин, сгенерированных методом моделирования пар.

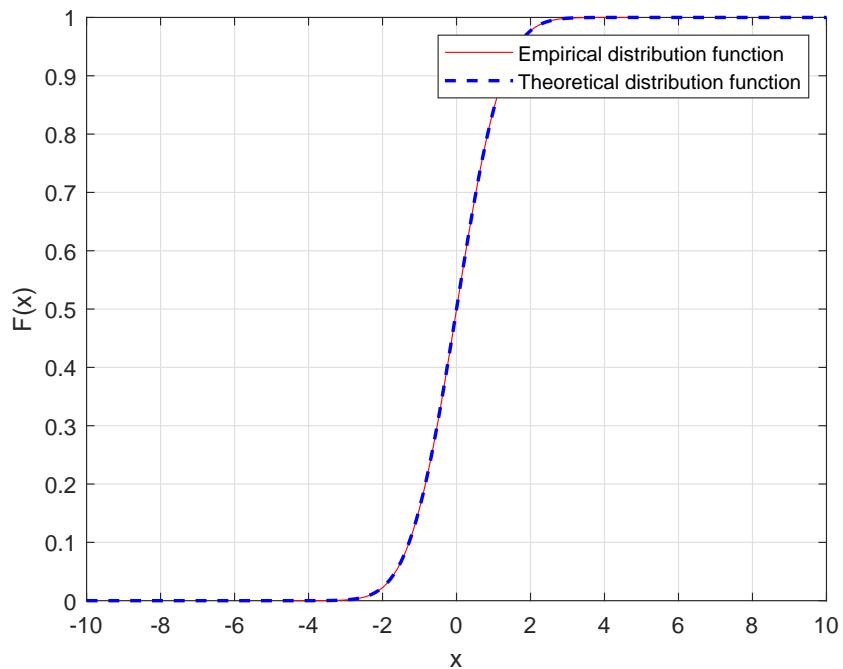


Рис. 15: Иллюстрация работы датчика стандартного нормального распределения (первая случайная величина), построенного с помощью метода перехода в полярные координаты, общее количество элементов $n = 1000000$

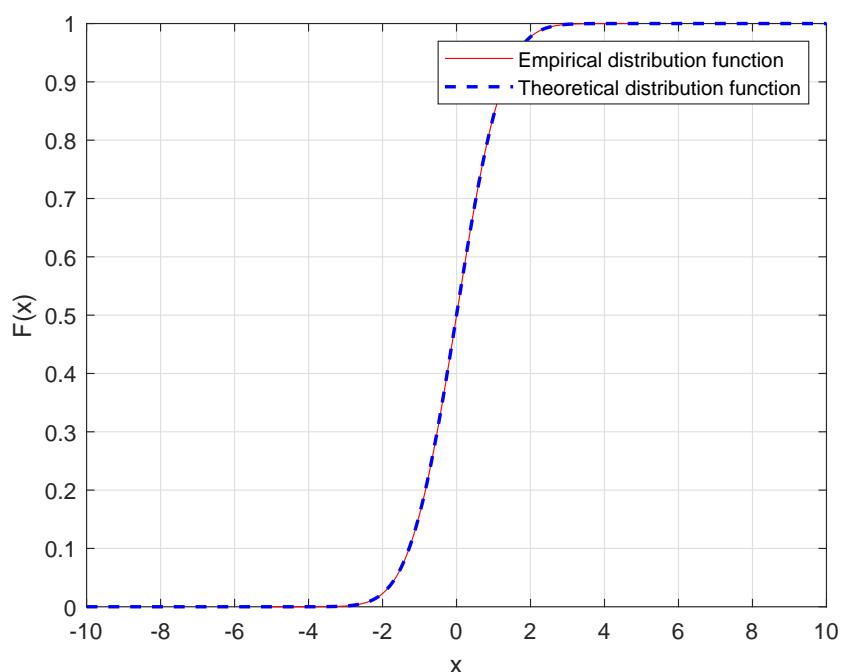


Рис. 16: Иллюстрация работы датчика стандартного нормального распределения (вторая случайная величина), построенного с помощью метода перехода в полярные координаты, общее количество элементов $n = 1000000$

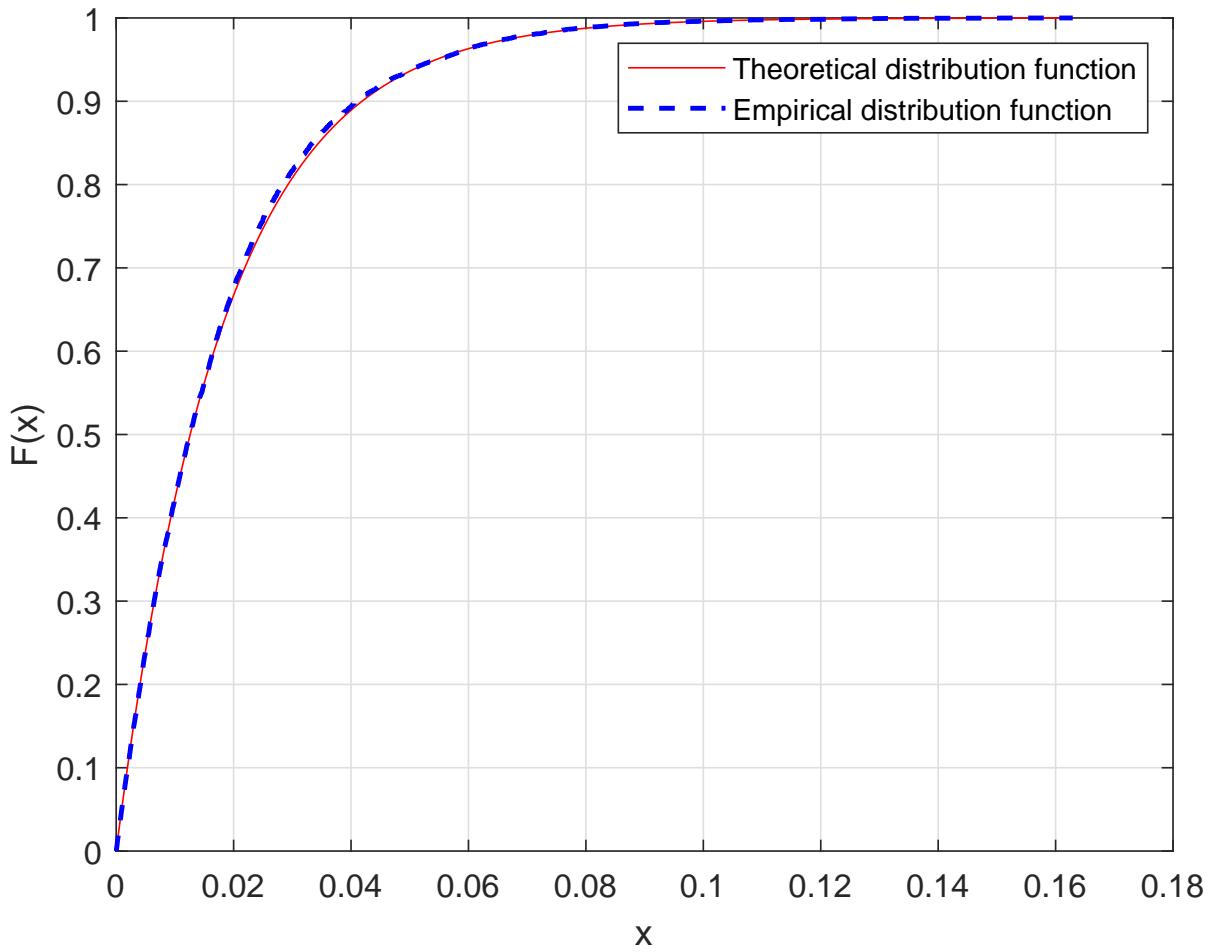


Рис. 17: Эмпирическая функция распределения случайной величины $\xi = \min(X_1, \dots, X_N)$, где $N = 10$ и $X_i \sim \text{Exp}(i)$, а также теоретическая функция экспоненциального распределения с параметром $\lambda = \sum_{i=1}^N i = 55$, общее количество элементов $n = 10000$

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 Теория и необходимая аналитика

Определение 10. Случайная величина ξ имеет распределение Коши с параметрами сдвига x_0 и масштаба $\gamma > 0$, если её функция распределения имеет вид:

$$F_\xi(x) = \frac{1}{\pi} \operatorname{arctg} \left(\frac{x - x_0}{\gamma} \right) + \frac{1}{2}.$$

Обозначение: $\xi \sim C(x_0, \gamma)$.

Заметим, что функция распределения возрастает и имеет обратную функцию:

$$F_\xi^{-1}(y) = x_0 + \gamma \operatorname{tg} \left(\pi \left(y - \frac{1}{2} \right) \right).$$

Воспользовавшись теоремой 5, получим, что если $Y \sim U[0, 1]$, то $X = F_\xi^{-1}(Y)$ имеет распределение Коши.

Результат работы датчика распределения Коши представлен на рисунке 18 в пункте 4.3.

Опишем теперь суть метода фон Неймана. Имеются плотности двух распределений: $p_1(x)$ и $p_2(x)$. Пусть также существует такая константа k , что $k p_1(t) \geq p_2(x)$ для всех x . Алгоритм выглядит следующим образом:

1. Генерируется случайный элемент x , соответствующий плотности $p_1(x)$.
2. Вычисляется $P = \frac{p_2(x)}{k p_1(x)}$.
3. Генерируется случайный элемент $y \sim \operatorname{Bern}(P)$, и если $y = 1$, то значение x возвращаем как результат генерации случайного элемента, имеющего плотность распределения $p_2(x)$, иначе повторяем процесс заново.

В нашем случае

$$p_1(x) = \frac{1}{\pi} \frac{\gamma}{(x - x_0)^2 + \gamma^2},$$
$$p_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Очевидно, что чем меньше k , тем быстрее работает метод фон Неймана. Будем считать, что $x_0 = 0$. Это логичное предположение, так как максимальное значение и $p_1(x)$, и $p_2(x)$

достигают в точке $x = 0$ (то есть плотность распределения Коши не нужно никуда «сдвигать», чтобы мажорировать плотность нормального распределения). Можно аналитически доказать, что при $x_0 \neq 0$ мы уже не сможем получить оптимальное значение k . Выберем параметр γ распределения Коши так, чтобы найти это минимальное значение k . Итак, требуется найти

$$k^* = \min_{\gamma} \max_x \frac{p_2(x)}{p_1(x)} = \min_{\gamma} \left(\frac{\sqrt{\pi}}{\gamma \sqrt{2}} \max_x \left((\gamma^2 + x^2) e^{\frac{x^2}{2}} \right) \right).$$

Исследуем функцию $f(x) = (\gamma^2 + x^2) e^{\frac{x^2}{2}}$ на экстремум.

$$f'(x) = -e^{-\frac{x^2}{2}} x (x^2 + \gamma^2 - 2) = 0.$$

Тем самым, максимум возникает при $x = 0$, если $|\gamma| > \sqrt{2}$, и при $x = \pm\sqrt{2 - \gamma^2}$ иначе. Тогда

$$k^* = \min \left\{ \min_{|\gamma| > \sqrt{2}} \left(\gamma \sqrt{\frac{\pi}{2}} \right), \min_{|\gamma| \leq \sqrt{2}} \left(\frac{\sqrt{2\pi}}{\gamma} e^{\frac{\gamma^2-2}{2}} \right) \right\} = \min \left\{ \sqrt{\pi}, \min_{|\gamma| \leq \sqrt{2}} \left(\frac{\sqrt{2\pi}}{\gamma} e^{\frac{\gamma^2-2}{2}} \right) \right\}.$$

Исследуем теперь функцию $g(\gamma) = \frac{\sqrt{2\pi}}{\gamma} e^{\frac{\gamma^2-2}{2}}$ на минимум.

$$g'(\gamma) = \sqrt{2\pi} e^{\frac{\gamma^2-2}{2}} \left(1 - \frac{1}{\gamma^2} \right),$$

то есть $g(\gamma)$ достигает своего минимума $\sqrt{\frac{2\pi}{e}}$ при $\gamma = 1 \in [-\sqrt{2}, \sqrt{2}]$. Тогда

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

Итак, оптимальными для скорости работы метода фон Неймана являются параметры $x_0 = 0$, $\gamma = 1$ с мажорантой $k = \sqrt{\frac{2\pi}{e}}$.

Обоснованием метода моделирования фон Неймана служит следующее

Утверждение 3. Пусть P и Q — вероятностные меры, заданные на измеримом пространстве $(\mathcal{X}, \mathcal{A})$, причём $P(A) < kQ(A)$ для всех $A \in \mathcal{A}$ для некоторого вещественного числа k . Пусть ν — бернульевская случайная величина с параметром $p = \frac{dP}{kdQ}$. Тогда

$$\mathbb{P}(X \in A | \nu = 1) = P(A).$$

Доказательство.

$$\mathbb{P}(X \in A | \nu = 1) = \frac{\mathbb{P}(X \in A, \nu = 1)}{\mathbb{P}(\nu = 1)} = \frac{\int_A \mathbb{P}(\nu = 1, X = x) dQ(x)}{\int_X \mathbb{P}(\nu = 1, X = x) dQ(x)} = \frac{\int_A \frac{1}{k} \frac{dP}{dQ} dQ(x)}{\int_X \frac{1}{k} \frac{dP}{dQ} dQ(x)} = P(A).$$

□

Иллюстрация работы построенного датчика, использующая функцию `normal probability plot`, представлена на рисунке 19 в пункте 4.3. Построение происходит в таком масштабе, чтобы график функции распределения стандартной нормальной случайной величины выглядел как прямая. По оси OX откладываются точки выборки, по оси OY откладываются квантили стандартного нормального распределения. На самом деле, в таком же масштабе функции любого другого нормального распределения тоже будут выглядеть как прямые.

Доказательство. Пусть $\xi \sim N(\mu, \sigma^2)$. Тогда функция распределения

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

где $F(x)$ — функция стандартного нормального распределения. Исходя из полученного равенства получаем, что и квантили различных распределений связаны между собой линейно. Из этого делаем вывод, что, во-первых, любую нормальную случайную величину $\xi \sim N(\mu, \sigma^2)$ можно представить в виде $\xi = \sigma\eta + \mu$, где $\eta \sim N(0, 1)$, а во-вторых, в масштабе, отрисовываемом функцией `normal probability plot`, функция любого нормального распределения $N(\mu, \sigma^2)$ будет прямой со сдвигом, равным μ и тангенсом угла наклона, равным σ , то есть корню из дисперсии. \square

Очевидно, что любое функции любого другого распределения в данном масштабе уже не будут иметь вид прямой, так как квантили любого другого распределения уже не будут линейно связаны с квантилями нормального распределения.

График сравнения времени работы метода пар и метода фон Неймана моделирования стандартной нормальной случайной величины изображён на рисунке 20 в пункте 4.3.

4.3 Примеры работы программы

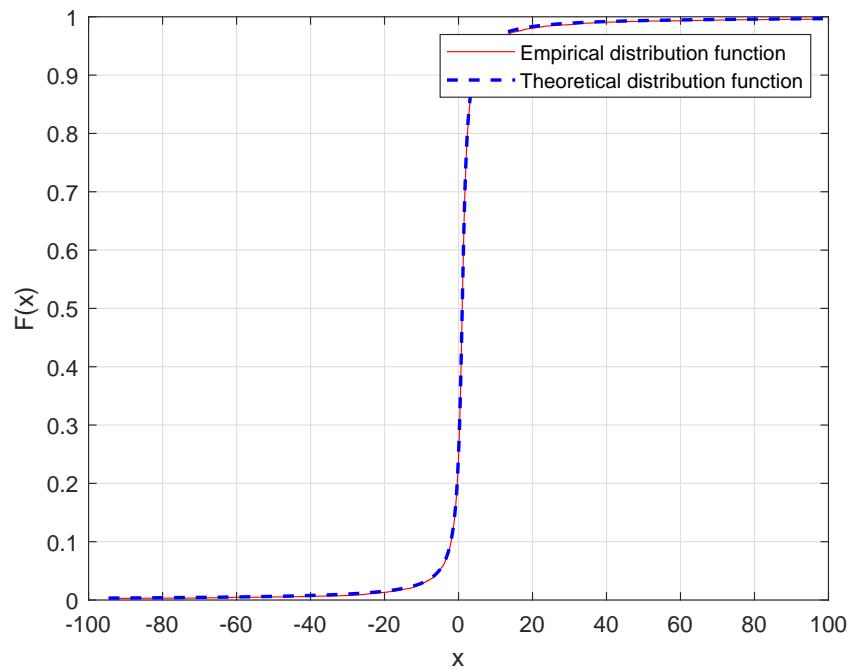


Рис. 18: Эмпирическая и теоретическая функции распределения Коши, параметры $x_0 = 1$, $\gamma = 1$, общее количество элементов $n = 10000$

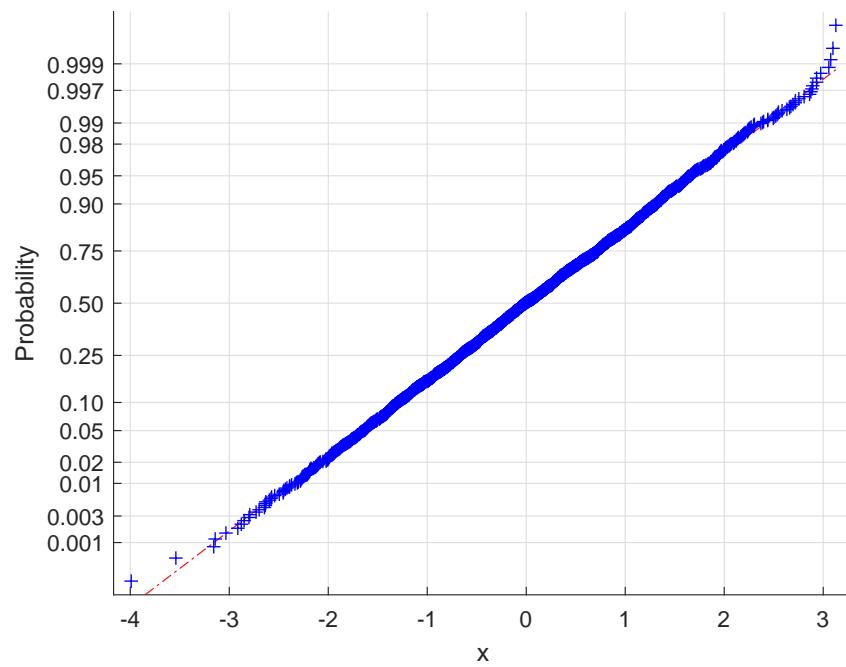


Рис. 19: Эмпирические и теоретические квантили стандартного нормального распределения, общее количество элементов $n = 3000$

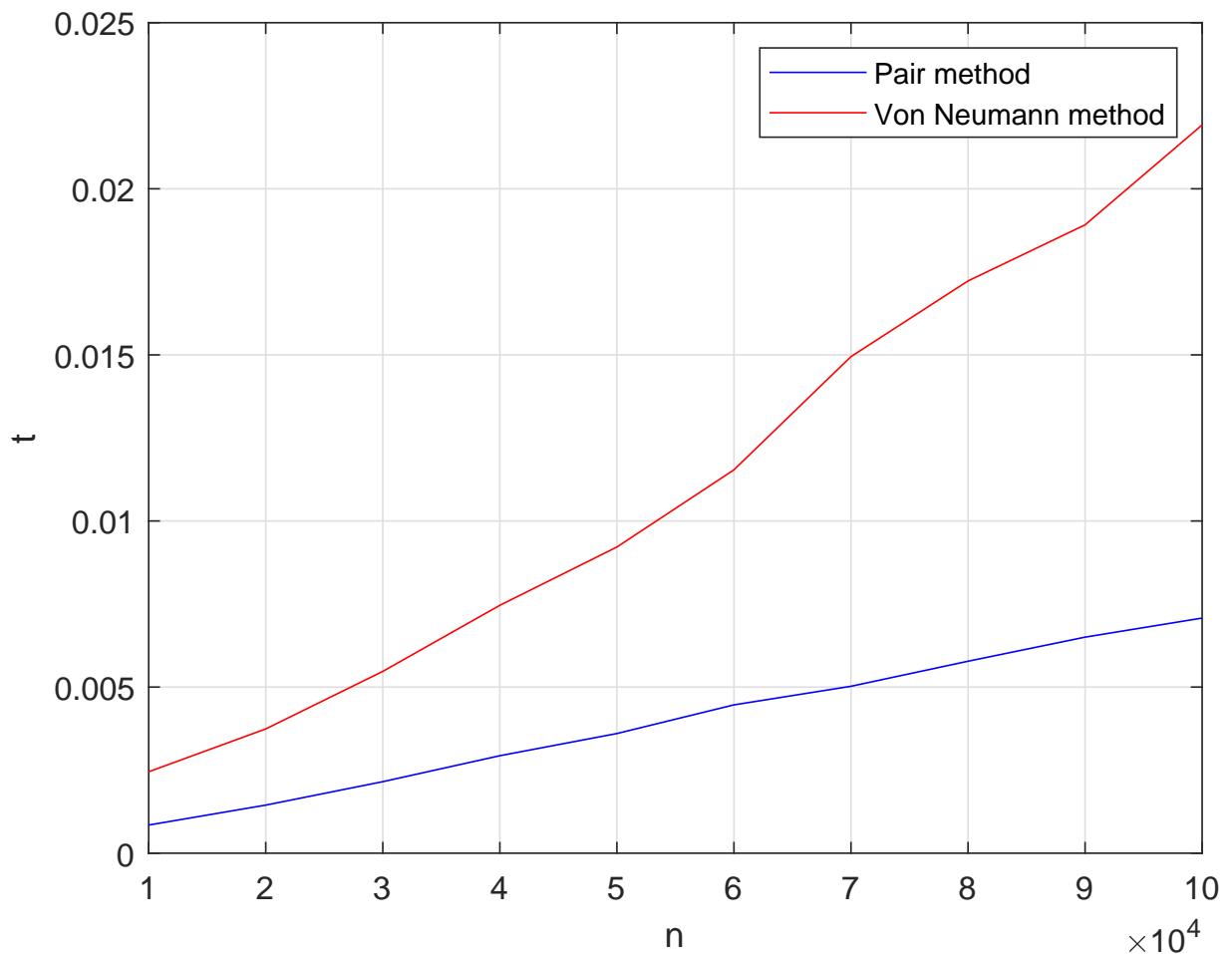


Рис. 20: Усреднённое (по $d = 200$ запускам) время работы метода пар и метода фон Неймана моделирования стандартной нормальной случайной величины в зависимости от длины выборки

5 Задание №5

5.1 Постановка задачи

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

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

- Считая μ и σ^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 объяснить, а также найти закон распределения данных сумм.

5.2 Теория и необходимая аналитика

Сформулируем закон больших чисел.

Теорема 8. Пусть X_1, \dots, X_n, \dots — последовательность независимых одинаково распределённых случайных величин, определённых на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ с конечным математическим ожиданием, равным a . Обозначим

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

Тогда для любого $\varepsilon > 0$

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|S_n - a| < \varepsilon) = 1.$$

Иными словами, последовательность S_n по вероятности сходится к a .

Пусть $X_1, \dots, X_n, \dots \sim N(\mu, \sigma^2)$. Иллюстрация сходимости последовательности S_n к математическому ожиданию μ представлена на рисунке 21 в пункте 5.3.

Вспомним теорему 2 (центральная предельная теорема). Иллюстрация её правильности изображена на рисунке 22 в пункте 5.3.

Построим теперь доверительные интервалы для среднего и дисперсии.

Определение 11. Доверительным интервалом параметра θ распределения случайной величины X с уровнем доверия p , порождённым выборкой (x_1, \dots, x_n) , называется интервал с границами $l(x_1, \dots, x_n)$ и $r(x_1, \dots, x_n)$, которые являются реализациями таких случайных величин $L(X_1, \dots, X_n)$ и $R(X_1, \dots, X_n)$, что

$$\mathbb{P}(L \leq \theta \leq R) = p.$$

Границы точке доверительного интервала l и r называются доверительными пределами.

Определение 12. Пусть X_1, \dots, X_n — выборка из независимых одинаково распределённых случайных величин. Тогда её выборочным средний называется случайная величина

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i;$$

выборочной дисперсией — случайная величина

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2;$$

несмешённой (исправленной) выборочной дисперсией — случайная величина

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

Определение 13. Пусть Y_0, \dots, Y_n — независимые одинаково распределённые случайные величины, имеющие стандартное нормальное распределение $N(0, 1)$. Тогда распределение случайной величины

$$t = \frac{Y_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n Y_i^2}}$$

называется распределением Стьюдента с n степенями свободы. Обозначение: $t \sim t(n)$.

Для данного распределения справедливо свойство симметричности:

$$t \sim t(n) \Rightarrow -t \sim t(n).$$

Определение 14. Случайная величина

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S},$$

где S — несмешённая выборочная дисперсия, имеет распределение Стьюдента с $n-1$ степенями свободы $t(n-1)$.

Определение 15. Пусть Z_1, \dots, Z_n — независимые одинаково распределённые случайные величины, имеющие стандартное нормальное распределение $N(0, 1)$. Тогда распределение случайной величины

$$X = \sum_{i=1}^n Z_i^2$$

называется распределением хи-квадрат с n степенями свободы. Обозначение: $X \sim \chi^2(n)$.

Из [2] известна следующая

Теорема 9. Случайная величина

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

где S^2 — несмешённая выборочная дисперсия, имеет распределение хи-квадрат с $(n-1)$ степенями свободы $\chi^2(n-1)$.

Пусть $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ — независимая выборка из нормального распределения, где μ, σ^2 — неизвестные параметры. Построим доверительный интервал для неизвестного среднего μ . Для этого введём случайную величину $T \sim t(n-1)$ и в силу симметрии распределения Стьюдента получим

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

где $t_{\alpha, n-1}$ — α -квантили распределения Стьюдента. После подстановки и преобразований получаем:

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

Пусть $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ — независимая выборка из нормального распределения, где μ, σ^2 — неизвестные параметры. Построим доверительный интервал для неизвестной дисперсии σ^2 . Для этого введём случайную величину $H \sim \chi^2(n-1)$. Тогда имеем:

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

где $\chi_{\alpha, n-1}^2$ — α -квантиль распределения хи-квадрат. После подстановки и преобразований получаем:

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

Графики посчитанных доверительных интервалов для среднего и дисперсии представлены на рисунках 23—24 в пункте 5.3.

Пусть теперь X_1, \dots, X_n — независимые случайные величины, имеющие стандартное распределение Коши со сдвигом a и масштабом b . Иначе говоря, $X_i \sim C(a, b)$. Проверим, выполняется ли закон больших чисел для данного распределения (напомним, что для его выполнения необходимо существование конечного математического ожидания, которого, как известно, нет у распределения Коши).

График зависимости выборочного среднего от длины выборки, построенный на выборке случайных величин, имеющих распределение Коши, и изображённый на рисунке 25 в пункте 5.3, наглядно показывает, что закон больших чисел для выборки данного распределения не работает (что было ожидаемо).

Докажем, что у распределения Коши нет математического ожидания. А именно, покажем, что соответствующий математическому ожиданию интеграл не сходится абсолютно:

$$\mathbb{E}(|X|) = \int_{-\infty}^{+\infty} \frac{b|x|}{(x-a)^2 + b^2} dx = \frac{b}{\pi} \left(\int_0^{+\infty} \frac{x}{(x+a)^2 + b^2} dx + \int_0^{+\infty} \frac{x}{(x-a)^2 + b^2} dx \right) = +\infty.$$

Для продвижения дальше вспомним некоторые понятия из анализа и теории вероятностей.

Определение 16. Пусть имеется случайная величина ξ . Тогда её характеристической функцией называется функция

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

Определение 17. Интеграл вида

$$\int_0^{+\infty} \frac{\cos(ax)}{b^2 + x^2} dx = \frac{\pi}{2|b|} e^{-|ab|}, \quad b \neq 0,$$

называется интегралом Лапласа. Его подробное вычисление можно найти в, например, [6].

Утверждение 4. Пусть $\xi \sim C[a, b]$. Тогда

$$\varphi_\xi(t) = e^{ait - b|t|}.$$

Доказательство.

$$\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \int_{-\infty}^{+\infty} e^{ity} \cdot \frac{1}{\pi} \cdot \frac{b}{b^2 + (y-a)^2} dy = \left\{ x = \frac{y-a}{b} \right\} = e^{ait} \int_{-\infty}^{+\infty} \frac{1}{\pi} \cdot e^{itbx} \cdot \frac{1}{1+x^2} dx =$$

$$\begin{aligned}
&= \{e^{i\theta} = \cos(\theta) + i \sin(\theta), \sin(-\theta) = -\sin(\theta), \cos(-\theta) = \cos(\theta)\} = e^{ait} \cdot \frac{2}{\pi} \int_0^{+\infty} \frac{\cos(tbx)}{1+x^2} dx = \\
&= \{\text{интеграл Лапласа}\} = e^{ait} \cdot \frac{2}{\pi} \cdot \frac{\pi}{2} e^{-|bt|} = \{b > 0\} = e^{ait-b|t|}.
\end{aligned}$$

□

Из вида характеристической функции распределения Коши вытекает теорема сложения:

Теорема 10. *Если $\xi_1 \sim C(a_1, b_1)$, $\xi_2 \sim C(a_2, b_2)$, то $\xi_1 + \xi_2 \sim C(a_1 + a_2, b_1 + b_2)$. Или, что то же самое,*

$$C(a_1, b_1) * C(a_2, b_2) = C(a_1 + a_2, b_1 + b_2),$$

где $*$ — операция свёртки.

Из этой теоремы выводится следующее интересное свойство распределения Коши:

Утверждение 5. *Если случайные величины ξ_1, \dots, ξ_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 ξ_j .*

Иллюстрация данного свойства приведена на рисунке 26 в пункте 5.3.

5.3 Примеры работы программы

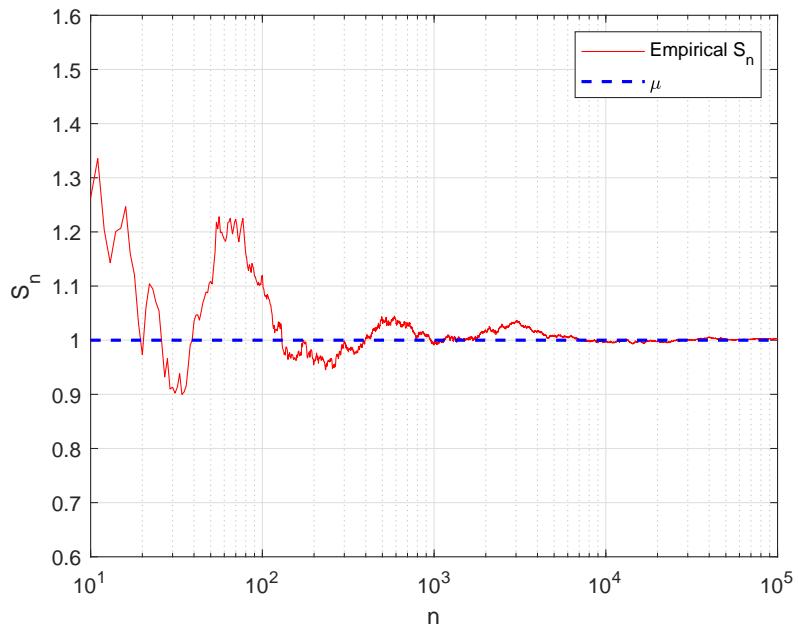


Рис. 21: Иллюстрация закона больших чисел на примере нормальных случайных величин, параметры $\mu = 1$, $\sigma^2 = 1$

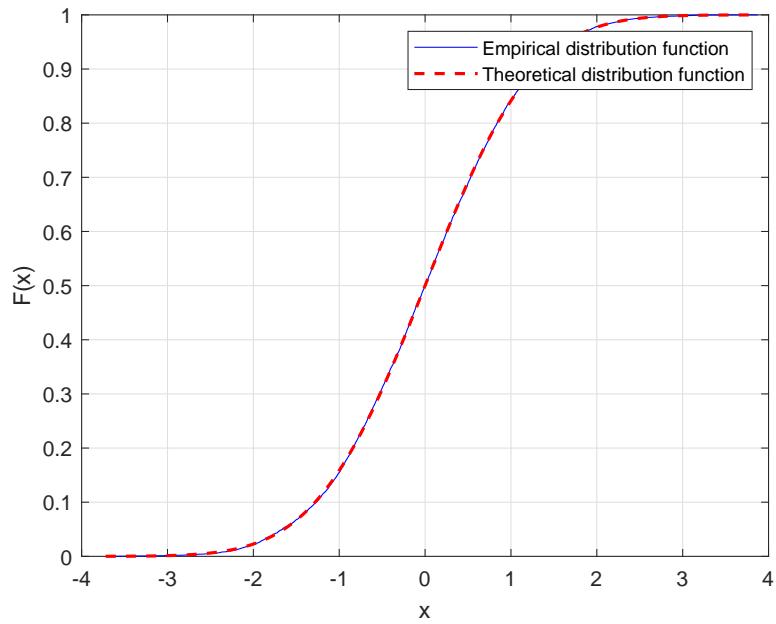


Рис. 22: Иллюстрация центральной предельной теоремы: взят достаточно большой номер $N = 10000$ подследовательности $Y(i)$, использовано $n = 10000$ выборок длины N , построена эмпирическая функция распределения получившейся случайной величины $Y(N)$, которая практически совпадает с функцией стандартного нормального распределения

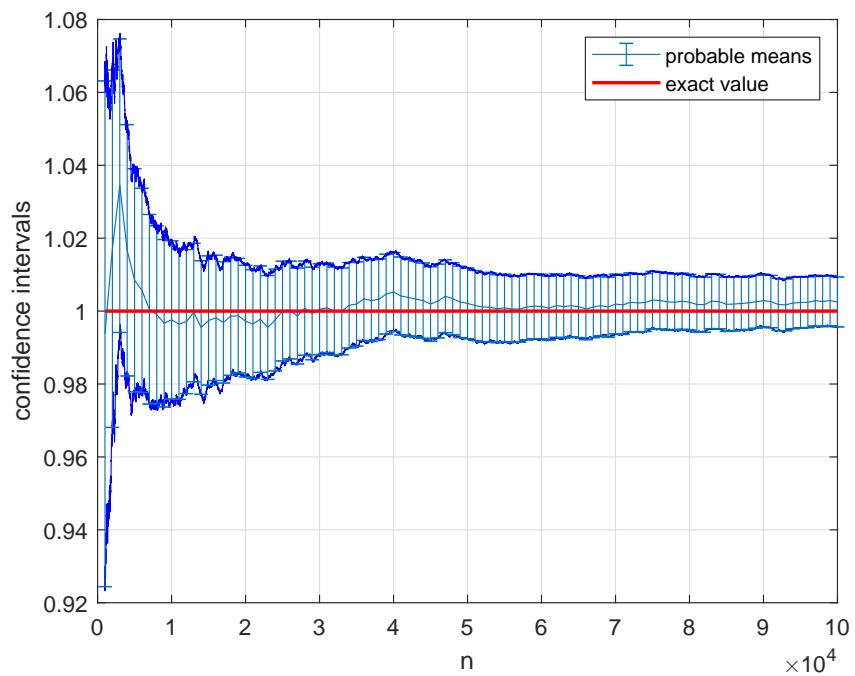


Рис. 23: Зависимость доверительного интервала для среднего от длины выборки (построение проводилось для одной выборки)

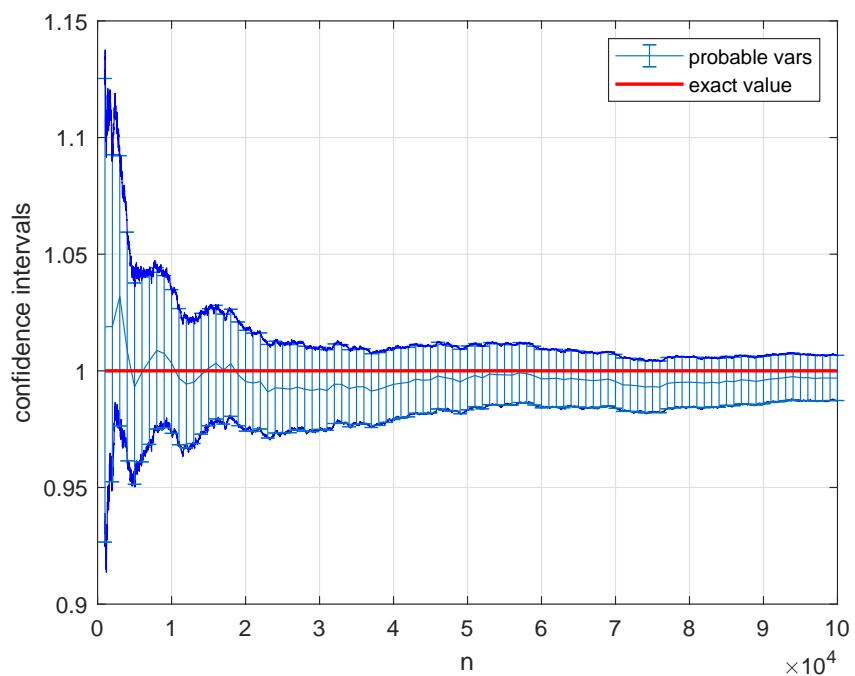


Рис. 24: Зависимость доверительного интервала для дисперсии от длины выборки (построение проводилось для одной выборки)

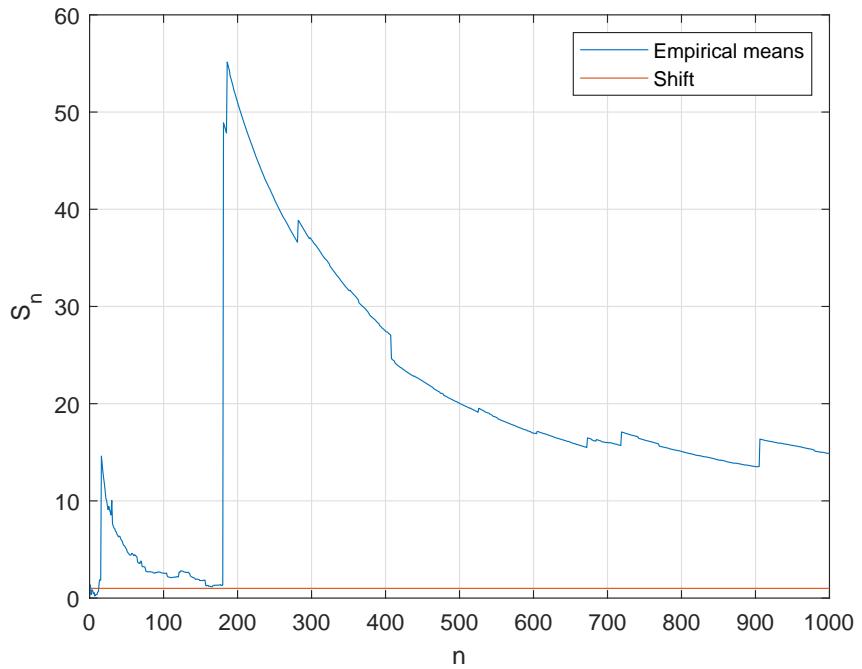


Рис. 25: Иллюстрация невыполнения закона больших чисел для распределения Коши, параметры $a = 1$, $b = 2$

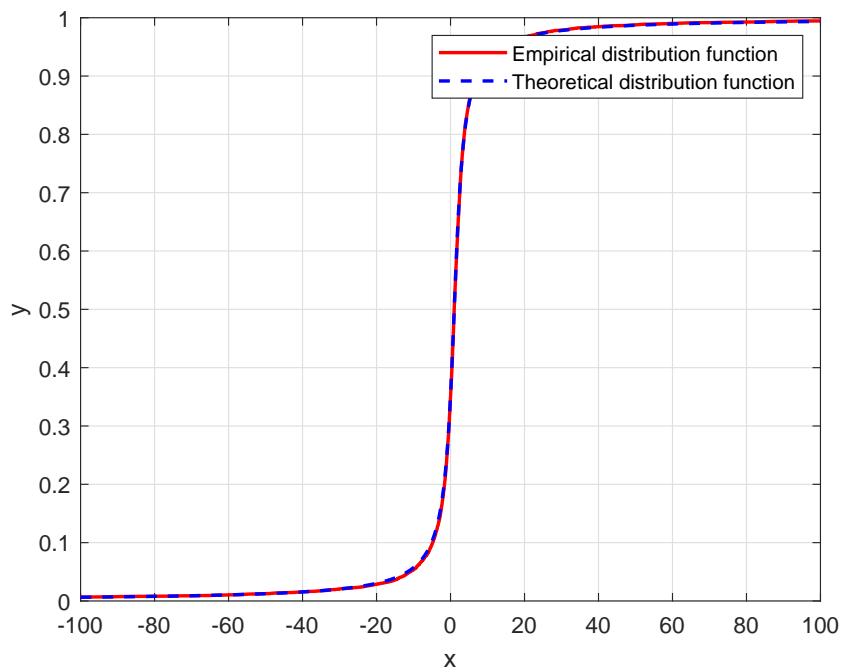


Рис. 26: Эмпирическое распределение среднего арифметического выборки, состоящей из случайных величин, распределённых по Коши с параметрами $a = 1$, $b = 2$, и теоретическое распределение Коши с теми же параметрами

6 Задание №6

6.1 Постановка задачи

1. Посчитать интеграл

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

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

2. Для каждого случая оценить точность вычислений.

6.2 Теория и необходимая аналитика

Представим интеграл

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

в виде

$$I = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, \dots, x_{10}) g(x_1, \dots, x_{10}) dx_1 \dots dx_{10},$$

где

$$f(x) = \pi^5 \cdot \frac{e^{-\frac{1}{2^7 \cdot x_1^2 \dots x_{10}^2}}}{x_1^2 \cdot \dots \cdot x_{10}^2},$$

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

Заметим, что $g(x)$ является совместной плотностью набора независимых случайных величин, имеющих нормальное распределение с параметрами 0 и $\frac{1}{2}$. Таким образом, можно переписать интеграл в виде:

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

В силу теоремы 21 (закона больших чисел) выборочное среднее будет стремиться к математическому ожиданию:

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n f(x_1^k, \dots, x_{10}^k) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(f(x_1, \dots, x_{10})), \quad x_i \sim N\left(0, \frac{1}{2}\right).$$

В описанной выше процедуре и заключается метод Монте-Карло. Оценим его погрешность на шаге n с помощью центральной предельной теоремы:

$$\mathbb{P}\left(\left|\frac{S_n}{n} - I\right| < \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - In}{n}\right| < \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n - In}{\sigma\sqrt{n}}\right| < \frac{\sqrt{n}}{\sigma}\varepsilon\right) =$$

$$\begin{aligned}
&= \mathbb{P} \left(\frac{S_n - In}{\sigma\sqrt{n}} < \frac{\sqrt{n}}{\sigma} \varepsilon \right) - \mathbb{P} \left(\frac{S_n - In}{\sigma\sqrt{n}} < -\frac{\sqrt{n}}{\sigma} \varepsilon \right) \approx \\
&\approx \Phi \left(\frac{\varepsilon\sqrt{n}}{\sigma} \right) - \Phi \left(-\frac{\varepsilon\sqrt{n}}{\sigma} \right) = 2\Phi \left(\frac{\varepsilon\sqrt{n}}{\sigma} \right) - 1 = \eta,
\end{aligned}$$

где

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

Тогда погрешность имеет вид

$$\varepsilon = \frac{k_{\eta+1}\sigma}{\sqrt{n}},$$

с вероятностью η , где k_α — квантиль стандартного нормального распределения.

Вместо σ будем пользоваться выборочной дисперсией:

$$\sigma_n^2 = \left(\frac{1}{n} \sum_{i=1}^n f^2(x^i) \right) - \left(\frac{1}{n} \sum_{i=1}^n f(x^i) \right)^2,$$

где под x^i подразумевается набор переменных x_1, \dots, x_{10} на i -й итерации.

Результаты расчётов с помощью метода Монте–Карло приведены в таблице 7 в пункте 6.3, а график сходимости расчётов к точному значению интеграла при увеличении числа итераций представлен на рисунке 27 в том же пункте.

Теперь посчитаем интеграл I , используя метод прямоугольников. Для этого необходимо свести задачу к вычислению собственного интеграла Римана. Сделаем это с помощью замены

$$x_i = \operatorname{tg} \left(\frac{\pi}{2} t_i \right).$$

Тогда

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

Проведём равномерное разбиение отрезка $[-1, 1]$ 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 нам необходимо выбрать середины нашего разбиения:

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

Тогда наш интеграл приближённо можно посчитать следующим образом:

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

Как известно из [5], погрешность метода средних прямоугольников для функции одной переменной $f(x)$ равна

$$\varepsilon = \frac{\max_{\xi \in [a,b]} |f''(\xi)|}{24} (b-a) h^2.$$

Распространяя эту оценку на функцию 10-ти переменных (достаточно грубо), получим итоговую оценку погрешности (учитывая $h = 2/N$):

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

Введём обозначение: $M = \max(\max |f''_{x_1, x_2}|, \max |f''_{x_1, x_1}|)$. В силу того, что функция $f(x)$ зависит от своих переменных симметрично, погрешность метода квадратур можно записать в виде:

$$\varepsilon = \frac{50M}{3N^2}.$$

Результаты расчетов методом квадратур приведены в таблице 8 в пункте 6.3.

6.3 Примеры работы программы

испытание\объём выборки	$n = 10^4$	$n = 10^5$	$n = 10^6$	$n = 10^7$
№1	$119,82 \pm 21,67$	$129,58 \pm 7,21$	$123,02 \pm 2,21$	$124,33 \pm 0,7$
№2	$126,91 \pm 22,58$	$127,3 \pm 7,13$	$124,77 \pm 2,23$	$125,05 \pm 0,7$
№3	$116,24 \pm 21,41$	$128,25 \pm 7,2$	$123,26 \pm 2,22$	$124,8 \pm 0,7$
№4	$120,26 \pm 21,51$	$128,15 \pm 7,2$	$124,56 \pm 2,23$	$124,44 \pm 0,7$
№5	$108,96 \pm 20,72$	$123,55 \pm 7,03$	$124,37 \pm 2,23$	$124,79 \pm 0,7$
среднее время подсчёта, сек	0,0061	0,0348	0.3136	3.0534
$n = 10^8$	$n = 10^9$	$n = 10^{10}$	$n = 10^{11}$	
$124,7098 \pm 0,221$	$124,8063 \pm 0,07$	$124,823 \pm 0,0221$	$124,8381 \pm 0,007$	
$124,8968 \pm 0,221$	$124,7887 \pm 0,07$	$124,834 \pm 0,0222$	$124,833 \pm 0,007$	
$124,7418 \pm 0,221$	$124,85 \pm 0,07$	$124,8461 \pm 0,0221$	$124,8354 \pm 0,007$	
$124,7061 \pm 0,221$	$124,8174 \pm 0,07$	$124,8352 \pm 0,0221$	$124,8324 \pm 0,007$	
$124,8099 \pm 0,221$	$124,8258 \pm 0,07$	$124,8211 \pm 0,0221$	$124,8366 \pm 0,007$	
15.9319	180.67	1423.66	31879.06	

Таблица 7: Расчёты интеграла методом Монте–Карло. Уровень значимости $\alpha = 5\%$.

N	значение интеграла	среднее время подсчёта, сек
3	272,6027	0,0317
4	183,4886	0,5714
5	116,3903	4,5384
6	120,6386	20,8854
7	128,1802	230,7313

Таблица 8: Расчёты интеграла методом квадратур

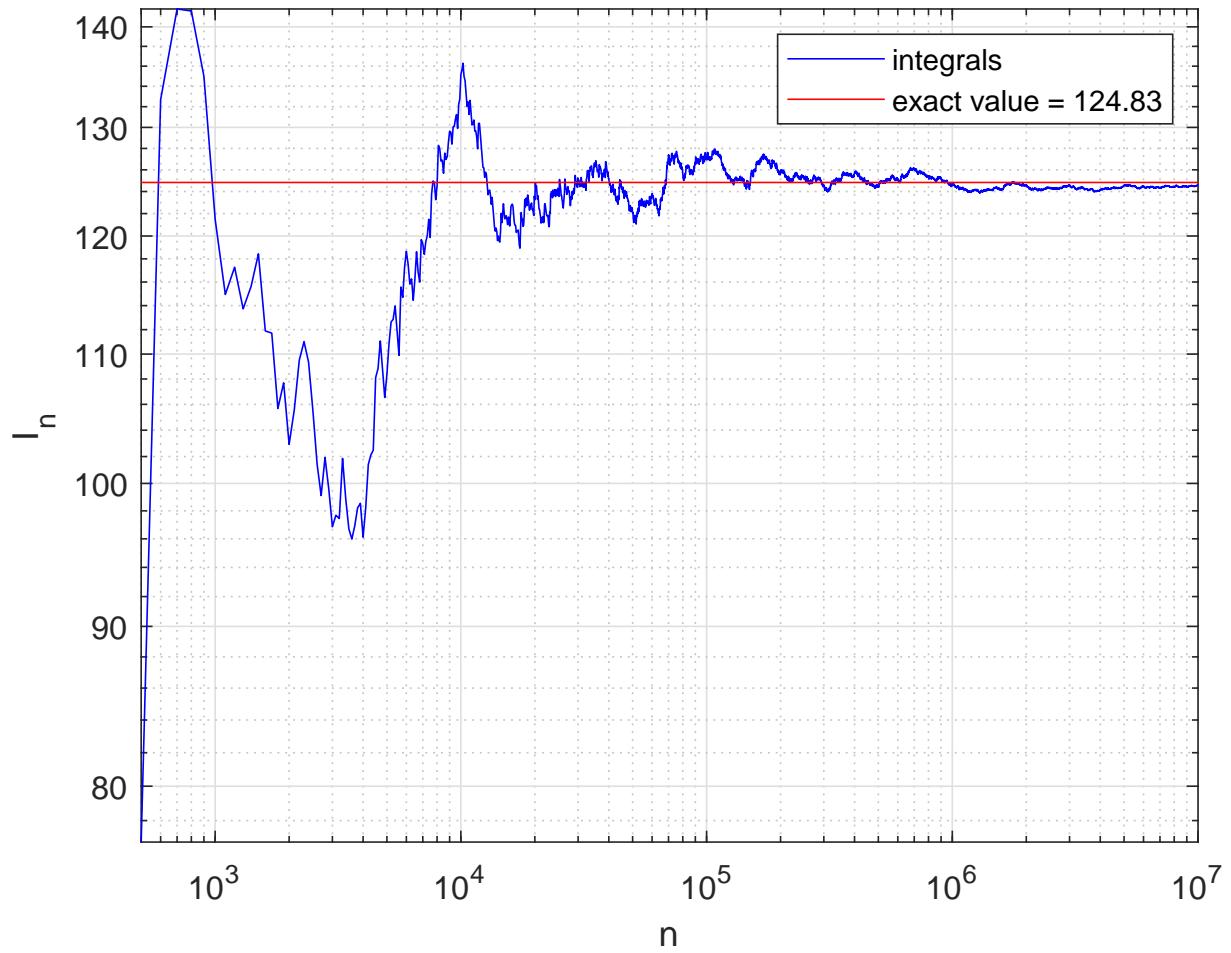


Рис. 27: Сходимость метода Монте–Карло к точному значению интеграла

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

$$f(x) = \begin{cases} x_1^3 \sin\left(\frac{1}{x_1}\right) + 10x_1x_2^4 \cos\left(\frac{1}{x_2}\right), & x_1 \neq 0, x_2 \neq 0, \\ x_1^3 \sin\left(\frac{1}{x_1}\right), & x_1 \neq 0, x_2 = 0, \\ 0, & x_1 = x_2 = 0. \end{cases}$$

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 Теория и необходимая аналитика

Для поиска минимума функции f будем N раз разыгрывать случайные величины x_1 и x_2 (для некоторого достаточно большого числа N), равномерно распределенные на единичном круге. После этого посчитаем значение $f(x_1, x_2)$ для каждой пары и выберем минимальное значение.

Оценим погрешность данного метода. Пусть точка (x_1^*, x_2^*) — точка минимума функции $f(x_1, x_2)$, а точка (x_1, x_2) — точка, найденная методом случайного поиска. Тогда по теореме Лагранжа (см., например, [9]):

$$|f(x_1^*, x_2^*) - f(x_1, x_2)| \leq \max_{(x_1, x_2) \in A} |\nabla f| \cdot |(x_1^*, x_2^*) - (x_1, x_2)|.$$

Оценим отдельные множители оценки.

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

Отсюда

$$|\nabla f| \leq \sqrt{\left(10 + \sqrt{10}\right)^2 + 17 \cdot 10^2}.$$

Оценим модуль разности аргументов. Пусть p — вероятность того, что в ε -окрестность точки минимума попала одна из N случайных точек. Заметим, что исходная функция имеет ровно два минимума: $(x_{1\min}, x_{2\min})$ и $(x_{1\min}, -x_{2\min})$. Возможны два варианта: точки минимума либо лежат внутри множества A , либо принадлежат его границе. В первом случае вероятность того, что хотя бы одна из точек попала в ε -окрестность точки минимума равна $p = 1 - (1 - 2\varepsilon^2)^n$. Во втором случае $p = 1 - (1 - \varepsilon^2)^n$. Откуда (рассмотрев худший случай) получаем, что

$$\varepsilon = \sqrt{1 - (1 - p)^{\frac{1}{n}}} \approx \sqrt{\frac{p}{n}}.$$

В итоге получаем, что погрешность метода случайного поиска составляет

$$|f(x_1^*, x_2^*) - f(x_1, x_2)| \leq \sqrt{\left(10 + \sqrt{10}\right)^2 + 17 \cdot 10^2} \sqrt{\frac{p}{n}}.$$

Зададим равномерное совместное распределение величин x_1, x_2 на единичном круге A , перейдя к полярной системе координат.

$$\frac{1}{\pi} \iint_{x_1^2 + x_2^2 \leq 1} dx_1 dx_2 = \frac{1}{\pi} \int_0^1 r dr \int_0^{2\pi} d\varphi = \int_0^{2\pi} \frac{1}{2\pi} d\varphi \int_0^1 dr^2.$$

Видно, что это распределение совпадает с совместным распределением случайной величины $\varphi \sim U[0, 2\pi]$ и случайной величины R , имеющей функцию распределения

$$F(t) = \begin{cases} 0, & x < 0, \\ t^2, & x \in [0, 1], \\ 1, & x > 1. \end{cases}$$

Заметим, что $F(t)$ удовлетворяет всем требованиям теоремы 5. Следовательно, мы можем моделировать случайную величину R как $F^{-1}(Y)$, где $Y \sim U[0, 1]$. А далее получаем, что

$$x_1 = R \cos \varphi, x_2 = R \sin \varphi.$$

Результаты подсчета минимума функции $f(x_1, x_2)$ методом случайного поиска представлены в таблице 9 в пункте 7.3, а график сходимости метода к минимуму представлен на рисунке 28 в том же пункте.

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

При помощи моделирования такого процесса ищется такая точка или множество точек, на котором достигается минимум 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 $\overline{x_1}$, «претендует» на то, чтобы лучше предыдущих приближать решение. Алгоритм принимает точку $\overline{x_0}$ 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 $\overline{x_{i+1}}$ по алгоритму получается на основе текущей точки $\overline{x_i}$ следующим образом. К точке $\overline{x_i}$ применяется оператор A , который случайным образом модифицирует соответствующую точку, в результате чего получается новая точка $\overline{x^*}$. Точка $\overline{x^*}$ становится точкой $\overline{x_{i+1}}$ с вероятностью $P(\overline{x^*}, \overline{x_{i+1}} | \overline{x_i})$, которая вычисляется в соответствии с распределением Гиббса:

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

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

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

В зависимости от выбора последовательности T_i и закона выбора точки-«претендента» на новое приближение существуют различные модификации метода имитации отжига. Мы остановимся на так называемом сверхбыстрым отжиге (Very Fast Annealing). В этом случае пространство X считается состоящим из D -мерных векторов (x_1, \dots, x_D) , где $x_i \in [A_i, B_i]$. Кроме этого, температура по каждой из координат может различаться, таким образом, T также является вектором размерности D . Семейство распределений строится следующим образом. Вводится функция

$$g_T(y) = \prod_{i=1}^D \frac{1}{2(|y_i| + T_i) \ln\left(1 + \frac{1}{T_i}\right)} \equiv \prod_{i=1}^D g_{(i;T)}(y_i), \quad y_i \in [-1, 1].$$

В качестве y для получения плотности распределений используется $\zeta = \frac{\Delta x}{B_i - A_i}$, то есть новое значение $x_i^{j+1} = x_i^j + z_i(B_i - A_i)$, где z_i — случайная величина с плотностью $g_{(i;T)}$ на $[-1, 1]$. При этом выходящие за границы интервала значения параметра генерируются заново (пока не произойдет попадание в интервал) или приравниваются соответствующим границам. Такую случайную величину легко промоделировать:

$$z_i = \operatorname{sgn}\left(\alpha_i - \frac{1}{2}\right) T_i \left[\left(1 + \frac{1}{T_i}\right)^{|2\alpha_i - 1|} - 1 \right].$$

Доказано, что закон изменения температуры

$$T_i(k) = T_{(i;0)} \exp\left(-c_i k^{\frac{1}{D}}\right), \quad c_i > 0.$$

дает статистическую гарантию нахождения глобального минимума. Для вероятности принятия также используется отдельная шкала температуры, изменяющаяся по такому же закону. Как правило, при реализации этого метода c_i управляется двумя параметрами:

$$c_i = m_i \exp\left(-\frac{n_i}{D}\right).$$

К сожалению, несмотря на широкое применение алгоритма, скорость его сходимости мало изучена (следовательно, получить конкретные оценки точности затруднительно, за исключением того, что точность данного метода не хуже точности метода градиентного спуска, обобщением которого он и является). На сегодняшний день существуют оценки только для дискретной оптимизации. Чтобы оценить скорость сходимости алгоритма имитации отжига, согласно [8], проделаем следующее. Назовём (x_1, x_2) состоянием оптимизационной задачи (далее ОЗ). Его соседями назовём те состояния, который отличаются от данного только лишь небольшим изменением одной из переменных x_1, x_2 . Построим ориентированный граф $G(V, E)$, связанный с ОЗ следующим образом: узлы в G — это просто состояния ОЗ, а ребра, идущие из любого узла, являются соседями этого состояния. Как легко заметить, метод имитации отжига реализует случайное блуждание по данному графу. Изменение состояние на каждом шаге зависит только лишь от текущего состояния. Также для конкретной температуры вероятности перехода из любого узла фиксированы. Эти факты подталкивают на моделирование метода имитации отжига цепью Маркова. Но вероятности перехода могут меняться с температурой. Поэтому полностью метод имитации отжига может быть смоделирован неоднородной по времени цепью Маркова. Мы предполагаем, что G является связным.

Для каждого узла i из G определим $N(i)$ как множество соседей состояния i , и пусть $g(i)$ — значение функции в состоянии i . Будем говорить, что метод имитации отжига сходится, если соответствующая ему цепь Маркова побывала в глобально оптимальном состоянии хотя бы один раз. Пусть T — минимальная температура, в которой когда либо находился алгоритм отжига. Пусть также $\Delta = \max_{i \in V, j \in N(i)} (g(i) - g(j))$. И, наконец, пусть d и D — это степень и диаметр графа $G(V, E)$ соответственно. Тогда в [8] доказана следующая

Теорема 11. *Метод имитации отжига сходится за время $\leq 2k[d \exp(\Delta/T)]^D$ с вероятностью $\geq (1 - 2^{-k})$ вне зависимости от начального состояния. Под временем в данном случае подразумевается количество итераций.*

Для функции Розенброка $g(x)$ известен глобальный минимум $(x_1^*, x_2^*) = (1, 1)$, минимальное значение $g(x^*) = 0$. Будем искать минимум методом имитации сверхбыстрого отжига на множестве $X = [-10, 10] \times [-10, 10]$. Будем останавливать алгоритм при достижении температуры $T_0(k) \leq T_{\lim} = 10^{-20}$ (это температура, которая влияет на вероятность выбора точки-кандидата в качестве нового приближения).

Сравним работу данного алгоритма с методом градиентного спуска, который является детерминированным. Его суть заключается в следующем. Очевидно, что если из какой-то точки пространства X двигаться в сторону антиградиента некоей функции $f(x)$, то по крайней мере в начальный промежуток этого движения значение функции будет уменьшаться (так как градиент функции является неким аналогом обычной производной функции одной переменной для функции многих переменных). Смысл метода градиентного спуска и состоит в движении в сторону антиградиента функции с некоторым шагом, который (в нашем случае) будет на каждой итерации уменьшаться в два раза (в при переходе в новую точку восстанавливаться в длине). Формально:

$$x^{j+1} = x^j - \lambda^j \nabla f(x^j).$$

Будем останавливать алгоритм, если расстояние между точками на соседних итерациях не больше, чем некоторое заранее заданное ε .

Результаты применения методов отжига и градиентного спуска к функции Розенброка представлены в таблице 10 в пункте 7.3.

Как мы видим, детерминированные методы имеют свойство сходиться к локальному минимуму, в то время как метод отжига благодаря элементам случайности позволяет во многих случаях эту проблему обойти.

7.3 Примеры работы программы

объём выборки	найденный минимум функции	погрешность
10^4	-1,2685	0,4306
10^5	-1,2827	0,1362
10^6	-1,2863	0,0431
10^7	-1,2882	0,0136
10^8	-1,2884	0,0043

Таблица 9: Поиск минимального значения функции методом случайного поиска. Здесь $p = 0.99$.

результаты	метод отжига	метод градиентного спуска
медиана	0,0818	—
минимальное значение	$2,3187 \cdot 10^{-6}$	0,36
среднее время, сек	0,005	0,002

Таблица 10: Поиск минимального значения функции Розенброка методом отжига (программа запускалась $d = 5000$ раз) и методом градиентного спуска. Начальное приближение $x^0 = (0, 0)$, точность $\varepsilon = 10^{-2}$. Остальные параметры метода сверхбыстрого отжига выбраны единичными.

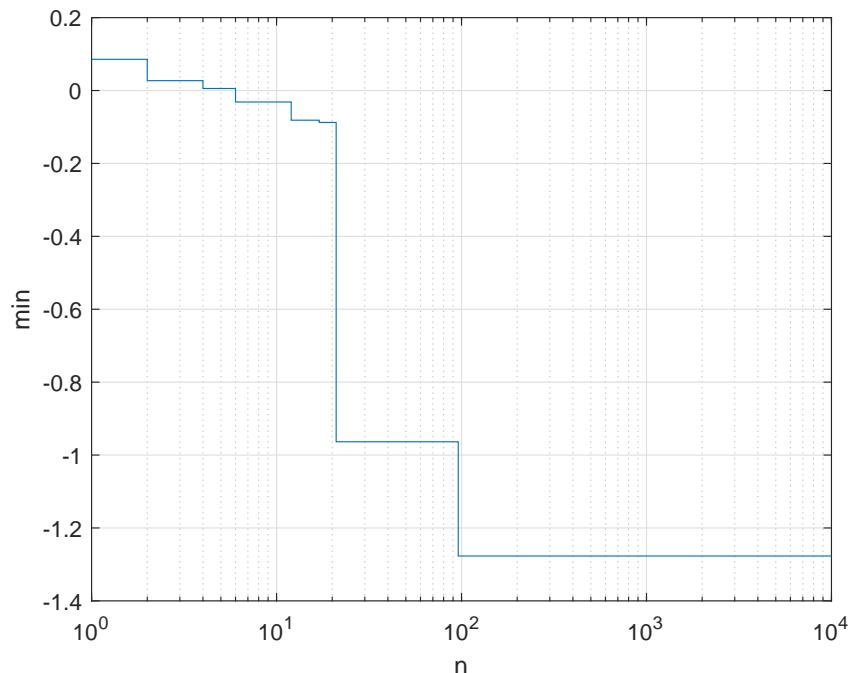


Рис. 28: Сходимость метода случайного поиска к минимуму

8 Задание №8

8.1 Постановка задачи

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

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

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

8.2 Теория и необходимая аналитика

Построим разностную схемы для задачи Дирихле (1). Для этого выберем достаточно мелкую квадратную сетку с шагом h . Координаты узлов $x_i = ih$, $y_j = jh$, а значения $u(x_i, y_j)$ и $f(x_i, y_j)$ будем для краткости обозначать $u_{i,j}$ и $f_{i,j}$.

Узел (x_i, y_j) называется внутренним, если он вместе со своими четырьмя соседями лежит во множестве D . В противном случае узел будет называть граничным. Обозначим через D_h множество внутренних узлов, а через δD_h — множество граничных узлов.

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

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

которое можно переписать в виде

$$u_{i,j} = \frac{1}{4}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}). \quad (2)$$

В граничном узле положим

$$u_{i,j} = f_{i,j}. \quad (3)$$

Заметим (см. [11]), что решение системы (2)–(3) существует, единствено и при $h \rightarrow 0$ стремится к решению задачи Дирихле (1).

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 одинаковой вероятностью $1/4$ может переместиться в один из четырёх соседних узлов, причём каждый такой единичный переход случаен и не зависит от положения частицы и истории её передвижений. Будем считать, что блуждание заканчивается, как только частица попадает в граничный узел. Можно доказать (см. [11]), что с вероятностью единица частица окажется в граничном узле через конечное число шагов.

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 блуждение точки неизбежно заканчивается на границе в первой же точке выхода её на границу, то

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

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

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

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 δD_h , то написанная выше сумма представляет собой математическое ожидание функции $f(x, y)$ на границе δD_h для траекторий, начинающихся в узле (x_i, y_j) . Тогда в силу закона больших чисел можно аппроксимировать математическое ожидание выборочным средним:

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

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

$$\begin{aligned} v_{i,j} &= \frac{1}{4} \sum_{(x_p, y_q) \in \delta D_h} (P(i-1, j, p, q) + P(i+1, j, p, q) + P(i, j-1, p, q) + P(i, j+2, p, q)) f_{pq} = \\ &= \frac{1}{4} (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}). \end{aligned}$$

То есть во внутреннем узле (x_i, y_j)

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

в граничном узле

$$v_{i,j} = f_{i,j}. \quad (5)$$

Заметим, что система (4)–(5) с точностью до переобозначения совпадает с системой (2)–(3).

По теореме о существовании решения внутренней задачи Дирихле (см. [10]) решение задачи (1) 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 условия на коэффициенты:

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

То есть мы получили, что функция $u(x, y) = x^2 - y^2$ является решением задачи (1), причём решение единственное (см. [10]).

Согласно приведённым выше выкладкам, численное решение может быть найдено по следующему алгоритму:

1. Построим квадратную сетку на $[-1, 1] \times [-1, 1]$ с шагом Δ .
2. Функцию во всех узлах, не принадлежащих кругу, положим равной NaN .

3. Все точки круга разделим на граничные и внутренние:

- В граничных точках положим $u(x, y) = f(x, y)$.
- Значение в каждой внутренней точке получим следующим образом. Попав во внутреннюю точку (x_i, y_j) , проведём серию из n случайных блужданий. Тогда

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

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

Результаты работы метода Монте–Карло на квадратной сетке при решении задачи Дирихле (1) при $f(x, y) = x^2 - y^2$ представлены на рисунке 29 в пункте 8.3, аналитическое решение представлено на рисунке 30, а функция ошибок — на рисунке 33 в том же пункте.

Решим теперь ту же самую задачу, используя полярную сетку в единичном круге. Как было показано в пункте 7.2, чтобы промоделировать равномерное распределение на единичном круге, достаточно взять $\varphi \sim U[0, 2\pi]$ и $r^2 \sim U[0, 1]$. В связи с этим возьмём равномерную сетку по φ и равномерную сетку по r^2 . Обозначим через h_φ шаг сетки по φ , и пусть $h_i = r_i - r_{i-1}$. Согласно [10], уравнение Лапласа в полярных координатах имеет вид

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \cdot \frac{\partial u}{\partial r} + \frac{1}{r^2} \cdot \frac{\partial^2 u}{\partial \varphi^2} = 0.$$

Для производной второго порядка на неравномерной сетке будем, согласно [5], использовать следующую разностную аппроксимацию:

$$y''(t_i) \approx \frac{2}{h_i + h_{i+1}} \cdot \left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right).$$

Обозначив $u_{i,j} = u(r_i \cos \varphi_j, r_i \sin \varphi_j)$, напишем разностную схему для выписанного выше уравнения Лапласа:

$$\begin{aligned} & \frac{2}{h_i + h_{i+1}} \cdot \frac{-u_{i,j}(h_i + h_{i+1}) + h_i u_{i+1,j} + h_{i+1} u_{i-1,j}}{h_{i+1} h_i} + \\ & + \frac{1}{r_i} \cdot \frac{(h_{i+1} - h_i) u_{i,j} + h_i u_{i+1,j} - h_{i+1} u_{i-1,j}}{2h_{i+1} h_i} + \frac{1}{r_i^2} \cdot \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_\varphi^2} = 0. \end{aligned}$$

Сгруппировав слагаемые, содержащие $u_{i,j}$, получим следующую формулу:

$$\begin{aligned} & \left(\frac{2}{h_i h_{i+1}} - \frac{1}{r_i} \cdot \frac{h_{i+1} - h_i}{2h_i h_{i+1}} + \frac{2}{r_i^2} \cdot \frac{1}{h_\varphi^2} \right) u_{i,j} = \\ & = u_{i+1,j} \left(\frac{2}{(h_i + h_{i+1}) \cdot h_{i+1}} + \frac{1}{r_i} \cdot \frac{1}{2h_{i+1}} \right) + u_{i-1,j} \left(\frac{2}{(h_i + h_{i+1}) \cdot h_i} - \frac{1}{r_i} \cdot \frac{1}{2h_i} \right) + \\ & + u_{i,j+1} \left(\frac{1}{r_i^2} \cdot \frac{1}{h_\varphi^2} \right) + u_{i,j-1} \left(\frac{1}{r_i^2} \cdot \frac{1}{h_\varphi^2} \right). \end{aligned}$$

Обозначим

$$A \equiv \frac{2}{h_i h_{i+1}} - \frac{1}{r_i} \cdot \frac{h_{i+1} - h_i}{2h_i h_{i+1}} + \frac{2}{r_i^2} \cdot \frac{1}{h_\varphi^2}, \quad \hat{p}_1 \equiv \frac{2}{(h_i + h_{i+1}) \cdot h_{i+1}} + \frac{1}{r_i} \cdot \frac{1}{2h_{i+1}},$$

$$\hat{p}_2 \equiv \frac{2}{(h_i + h_{i+1}) \cdot h_i} - \frac{1}{r_i} \cdot \frac{1}{2h_i}, \quad \hat{p}_3 = \hat{p}_4 \equiv \frac{1}{r_i^2} \cdot \frac{1}{h_\varphi^2}; \quad p_i = \frac{\hat{p}_i}{A}, \quad i = 1, 2, 3, 4.$$

Обратим внимание, что $p_1 + p_2 + p_3 + p_4 = 1$. Тогда, аналогично предыдущему способу, согласно закону больших чисел, можно аппроксимировать $u_{i,j}$ выборочным средним

$$u_{i,j} \approx \frac{1}{N} \sum_{k=1}^N f \left(r_p^{(k)} \cos(\varphi_q^{(k)}), r_p^{(k)} \sin(\varphi_q^{(k)}) \right).$$

Алгоритм с точностью до очевидных изменений совпадает с предыдущим алгоритмом. Единственное существенное отличие — центральную точку тоже считаем границей (заметим, что вероятности в ней не определены).

Результаты работы метода Монте–Карло на полярной сетке при решении задачи Дирихле (1) при $f(x, y) = x^2 - y^2$ представлены на рисунке 31 в пункте 8.3, аналитическое решение представлено на рисунке 32, а функция ошибок — на рисунке 34 в том же пункте.

8.3 Примеры работы программы

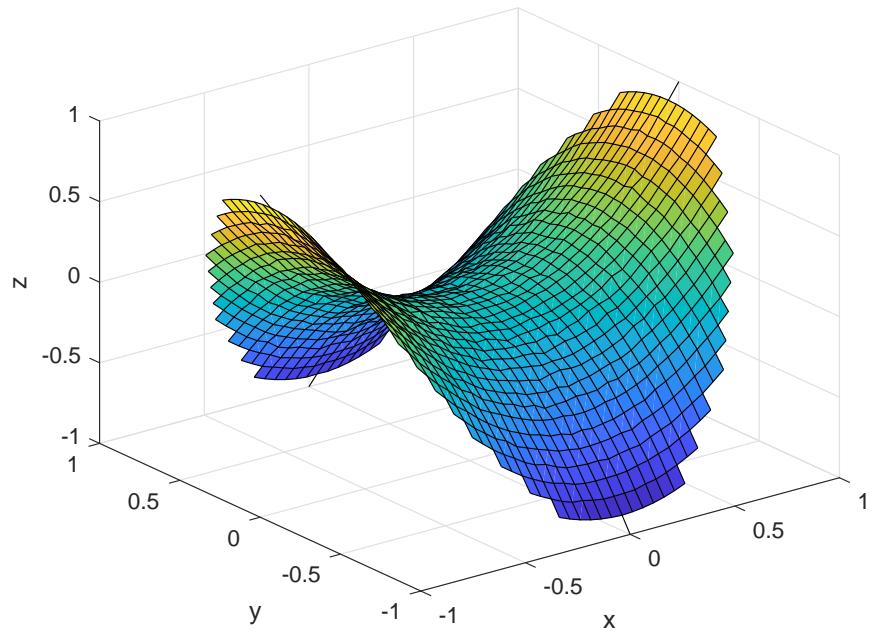


Рис. 29: Решение задачи Дирихле, полученное методом Монте–Карло на квадратной сетке, $\Delta = 10^{-1}$, $N = 10^6$

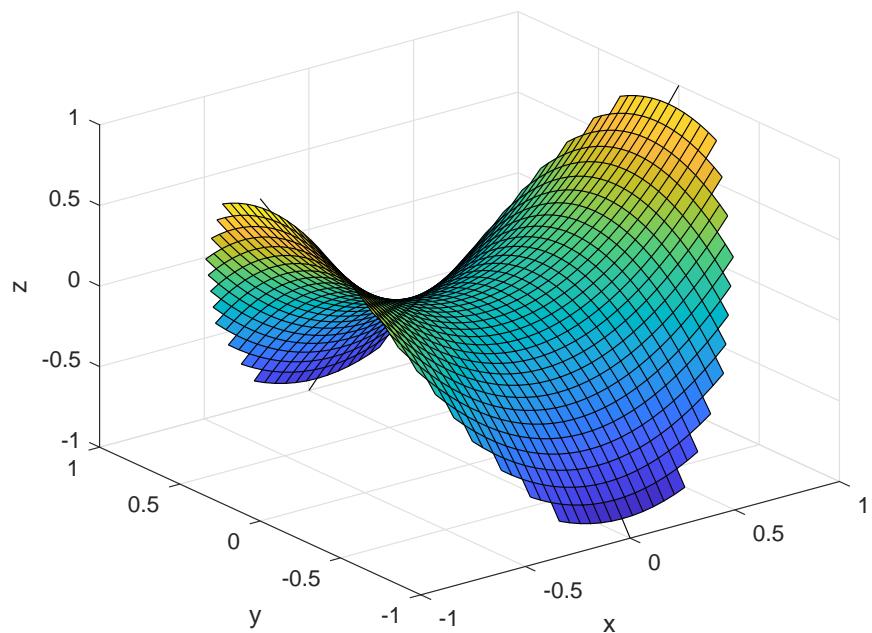


Рис. 30: Аналитическое решение задачи Дирихле на квадратной сетке

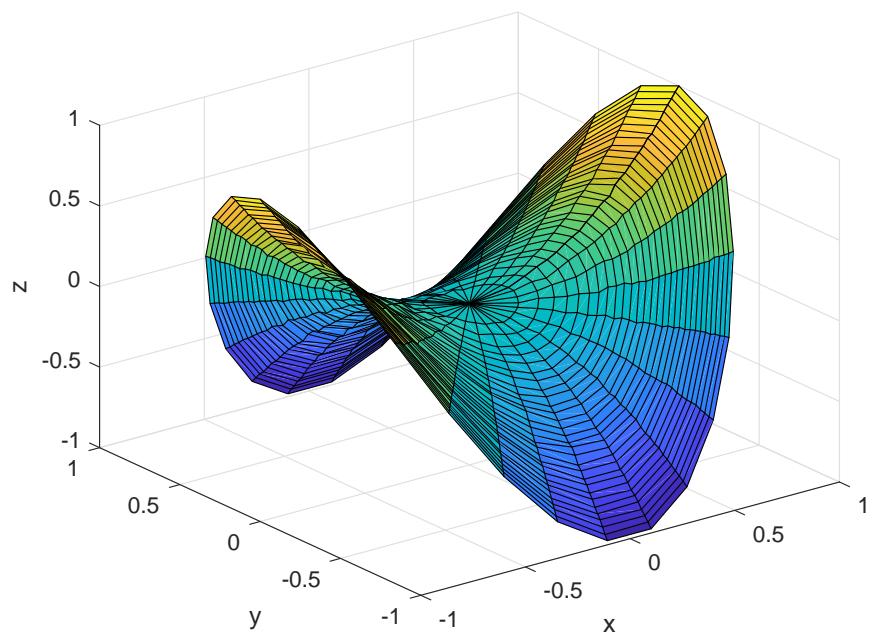


Рис. 31: Решение задачи Дирихле, полученное методом Монте–Карло на полярной сетке, $\Delta = 10^{-1}$, $N = 10^6$

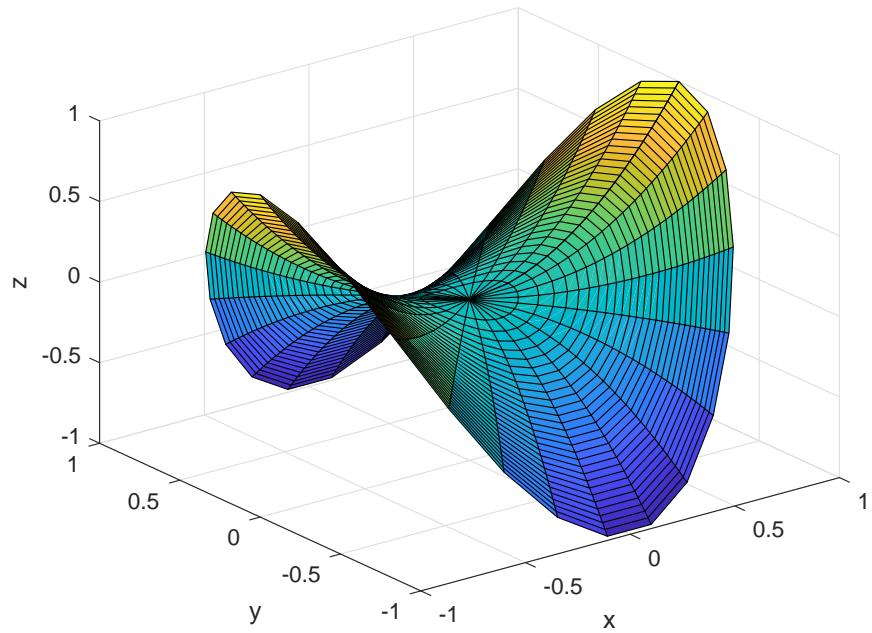


Рис. 32: Аналитическое решение задачи Дирихле на полярной сетке

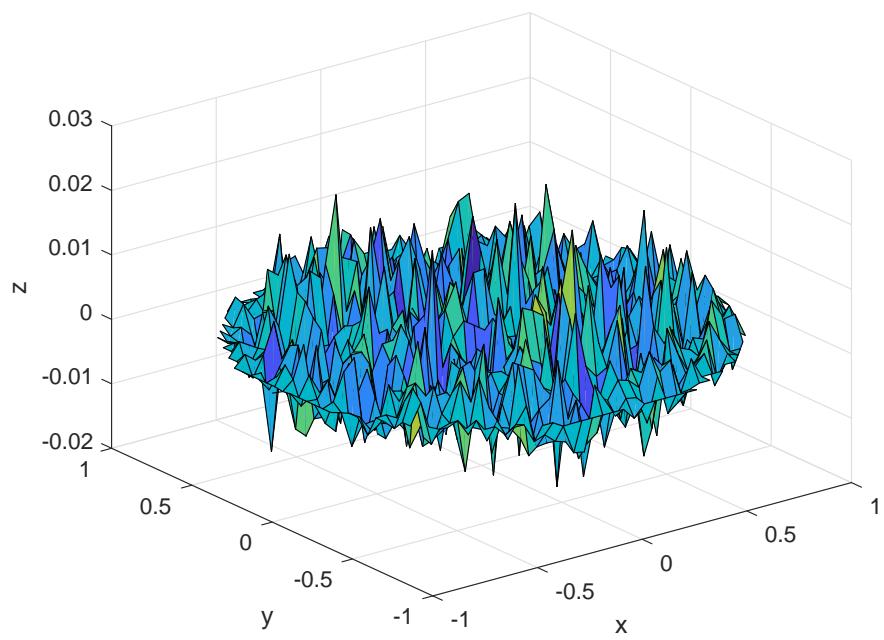


Рис. 33: Разность между теоретическим и найденным решениями на квадратной сетке

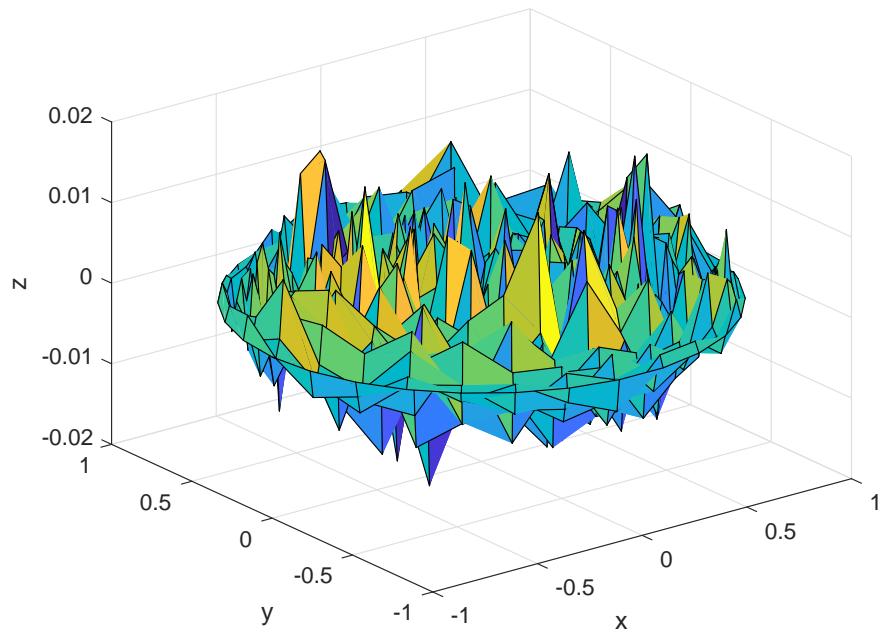


Рис. 34: Разность между теоретическим и найденным решениями на полярной сетке

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. Начальные значения X_0 генерируются 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 разбиения отрезка.
3. Построить график траектории, не соединяя точки ломаной, с целью получения визуально непрерывной линии.

9.2 Теория и необходимая аналитика

Определение 18. 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 19. 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 $\text{cov}(W(t_i), W(t_j)) = \min(t_i, t_j)$.

Так как процесс гауссовский, то случайные величины W_0 и W_1 — нормально распределены. Учитывая вид ковариационной функции, получим, что $W_0 = 0$ почти всюду, а $W_1 \sim N(0, 1)$.

Определение 20. Случайный процесс $\{W_t\}_{t \in T}$, где $T \subset [0, +\infty)$, называется процессом с независимыми приращениями, если для любых $t_0, t_1, \dots, t_n \in T$ таких, что

$$0 = t_0 < t_1 < \dots < t_{n-1} < t_n,$$

случайные величины $W_{t_0}, W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}}$ независимы.

Покажем, что винеровский процесс — процесс с независимыми приращениями. Так как для любых $t_0, t_1, \dots, t_n \in T$ случайный вектор $(W_{t_0}, W_{t_1}, \dots, W_{t_n})$ — гауссовский, то и вектор $(W_{t_0}, W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})$ имеет многомерное нормальное распределение (как линейная комбинация двух гауссовых векторов). Тогда независимость его компонент равносильна их некоррелированности (см. [2]). Пусть $t_0 \leq t_i \leq t_{i+1} \leq t_j \leq t_{j+1}$. Имеем

$$\text{cov}((W_{t_{i+1}} - W_{t_i}), (W_{t_{j+1}} - W_{t_j})) =$$

$$= \text{cov}(W_{t_{i+1}}, W_{t_{j+1}}) - \text{cov}(W_{t_{i+1}}, W_{t_j}) - \text{cov}(W_{t_i}, W_{t_{j+1}}) + \text{cov}(W_{t_i}, W_{t_j}) = t_{i+1} - t_{i+1} - t_i + t_i = 0.$$

А также

$$\text{cov}((W_{t_{i+1}} - W_{t_i}), W_{t_0}) = \text{cov}(W_{t_{i+1}}, W_{t_0}) - \text{cov}(W_{t_i}, W_{t_0}) = t_0 - t_0 = 0.$$

Отсюда и следует, что винеровский процесс является процессом с независимыми приращениями. Тогда, зная случайные величины W_{t_1} и W_{t_2} , мы можем сгенерировать случайную величину W_t , где $t_1 < t < t_2$.

Будем строить траекторию винеровского процесса рекурсией в глубину, т.е. будем генерировать W_t , где

$$t = t_1 + \alpha(t_2 - t_1), \quad \alpha \in (0, 1),$$

при условии, что $W_{t_1} = x_1$, $W_{t_2} = x_2$. В качестве начальных условий возьмём

$$W_0 = 0, \quad W_1 \sim N(0, 1).$$

Известно, что W_t — нормально распределена. Определим, как именно. Найдём выражение для условной плотности

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

Так как процесс — гауссовский, то

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

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

Используя выражение ковариационной функции, получаем, что

$$R_1 = \begin{pmatrix} t_1 & t_1 & t_1 \\ t_1 & t & t \\ t_1 & t & t_2 \end{pmatrix}, \quad R_2 = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix}.$$

После ряда преобразований получим следующее выражение для условной плотности:

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

Из этого получаем, что

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

Обозначив через n количество уровней рекурсии, мы получим значение W_t в $2^n - 1$ точках.

Траектория винеровского процесса, построенная методом добавления разбиения отрезка, представлена на рисунке 35 в пункте 9.3. На рисунке 36 в том же пункте построены несколько траекторий винеровского процесса, а также их доверительные интервалы.

Перейдём теперь к процессу Орнштейна–Уленбека.

Определение 21. Пусть дан случайный процесс $\{W_t\}_{t \in T}$. Тогда он называется стационарным, если конечномерные распределения инвариантны относительно сдвига времени.

Определение 22. Гауссовский процесс $\{W_t\}_{t \in T}$ называется процессом Орнштейна–Уленбека, если он является стационарным и марковским.

Из стационарности процесса Орнштейна–Уленбека следует, что

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

Без ограничения 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 12. Для того чтобы последовательность W_1, \dots, W_n нормально распределённых 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 этой теоремы можно найти в [12].

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

$$\rho(s, t) = \rho(s, \tau)\rho(\tau, t). \tag{6}$$

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

$$x = s - \tau,$$

$$y = \tau - t,$$

преобразуем выражение (6) к выражению

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

Теорема 13. 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 λ — некоторая положительная константа.

Доказательство этой теоремы также можно найти в [12].

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 является гауссовским), поэтому поделирование процесса Орнштейна–Уленбека заключается в моделировании случайных величин, имеющих распределение $N(a, \sigma^2)$.

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

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

Найдём переходную плотность

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

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

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

Поэтому

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

то есть

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

Так как рассматриваемый процесс является марковским, то, зная 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 аналогично моделированию винеровского процесса. Для упрощения положим $\alpha = 1/2$. Найдём условную плотность

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

где $t = (t_1 + t_2)/2$. Поскольку процесс W_t является гауссовским, то

$$\begin{aligned} p_{W_{t_1} W_t W_{t_2}}(x_1, x, x_2) &= \frac{1}{(2\pi)^{\frac{3}{2}} |R_1|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x_1, x, x_2)^T R_1^{-1}(x_1, x, x_2)\right\}, \\ p_{W_{t_1} W_{t_2}}(x_1, x_2) &= \frac{1}{2\pi |R_2|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x_1, x_2)^T R_2^{-1}(x_1, x_2)\right\}, \end{aligned}$$

где

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

После ряда преобразований 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 W_0 и W_1 возьмём

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

Траектории процесса Орнштейна–Уленбека для различных значений параметров представлены на рисунках 37–39 в пункте 9.3. На рисунке 40 в том же пункте построены несколько траекторий процесса Орнштейна–Уленбека, а также их доверительные интервалы.

9.3 Примеры работы программы

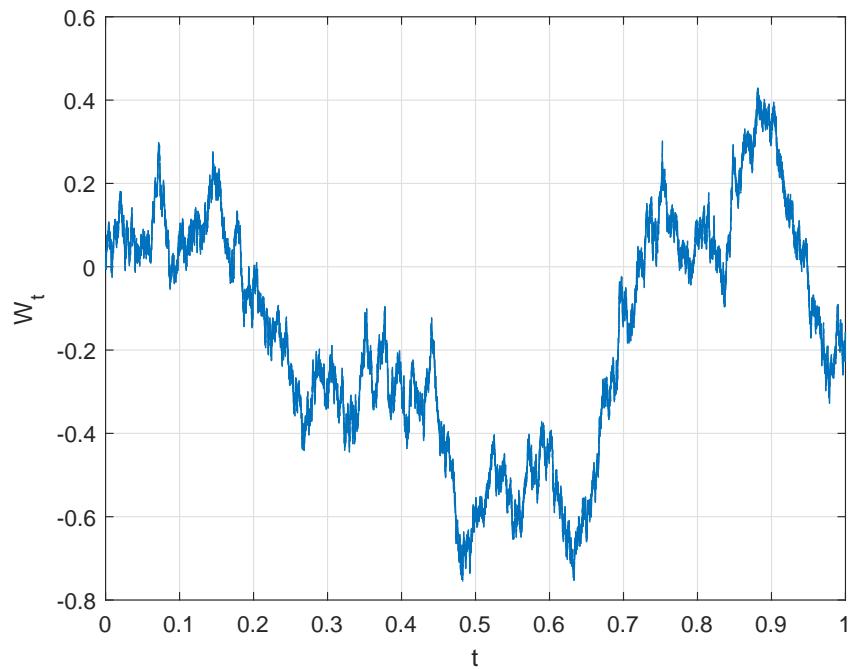


Рис. 35: Траектория винеровского процесса, $n = 20$, $\alpha = 0,5$

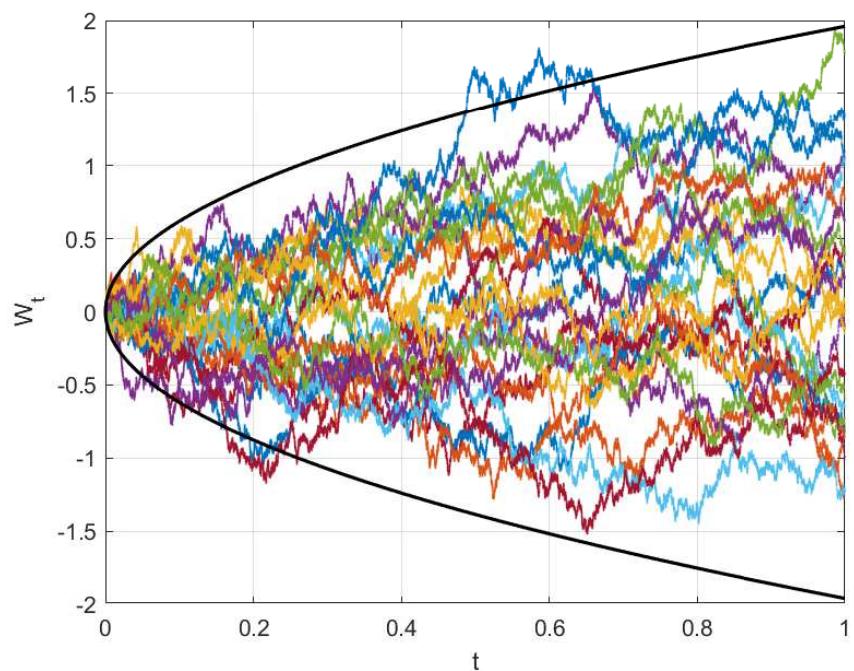


Рис. 36: Траектории винеровского процесса (количество $d = 50$), ограниченные доверительным интервалом $k_{1-\beta/2} \times [-\sqrt{t}, \sqrt{t}]$, где k_β — квантиль стандартного нормального распределения; $\beta = 5\%$, $n = 18$, $\alpha = 0,5$

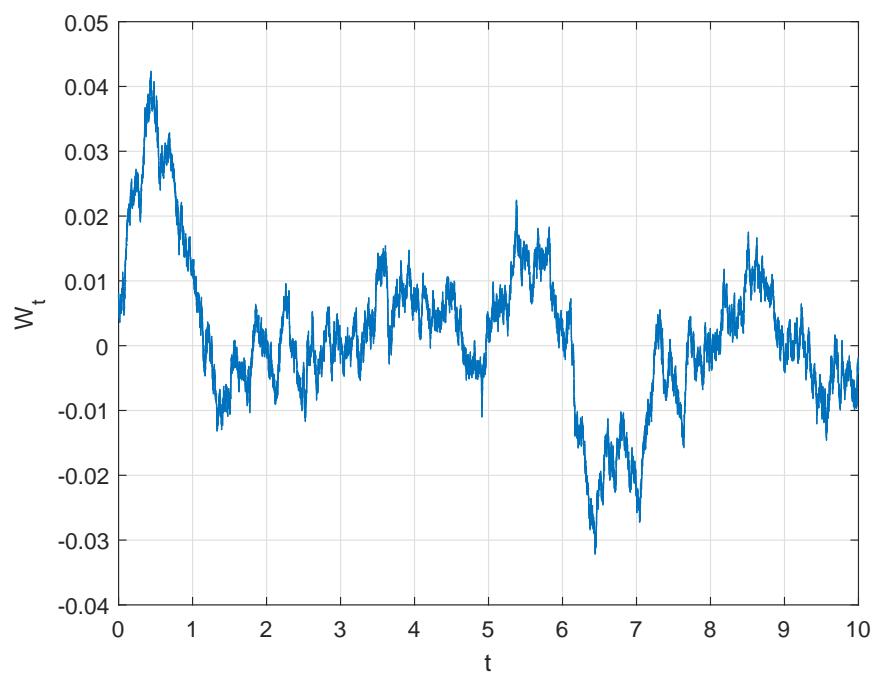


Рис. 37: Траектория процесса Орнштейна–Уленбека, $\lambda = 2$, $\sigma = 0,01$, $n = 20$

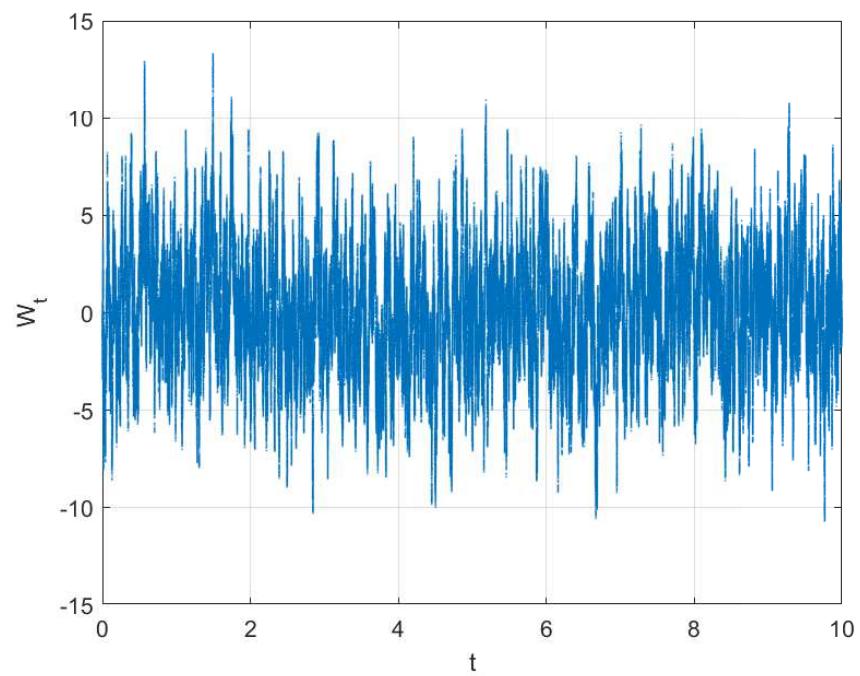


Рис. 38: Траектория процесса Орнштейна–Уленбека, $\lambda = 100$, $\sigma = 3$, $n = 20$

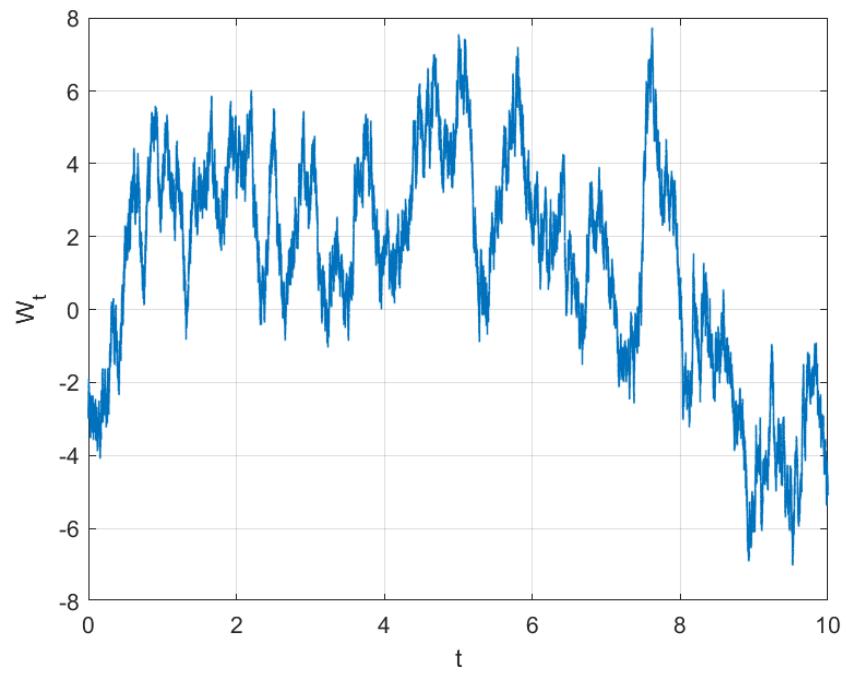


Рис. 39: Траектория процесса Орнштейна–Уленбека, $\lambda = 2$, $\sigma = 3$, $n = 20$

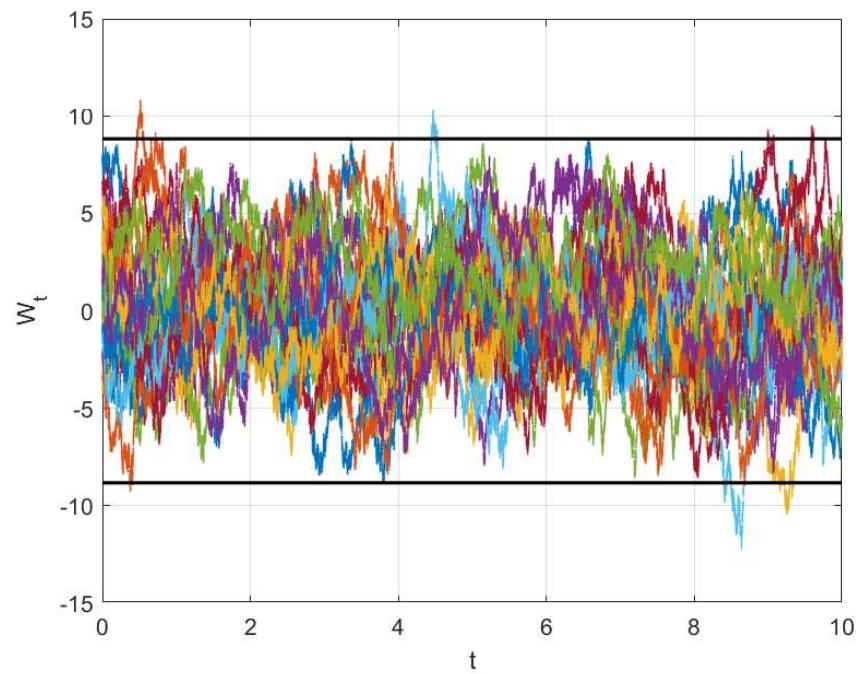


Рис. 40: Траектории процесса Орнштейна–Уленбека (количество $d = 25$), ограниченные доверительным интервалом $k_{1-\beta/2} \times \frac{\sigma^2}{\lambda} \times [-1, 1]$, где k_β — квантиль стандартного нормального распределения; $\beta = 5\%$, $\lambda = 2$, $\sigma = 3$, $n = 18$

10 Задание №10

10.1 Постановка задачи

Произвести 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 Теория и необходимая аналитика

Определение 23. Дискретным белым шумом называется последовательность $\varepsilon_1, \dots, \varepsilon_n, \dots$ независимых одинаково распределённых 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

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

Рассмотрим случай, когда $\omega(k)$ и x_k имеют гауссовское распределение. В этом случае условное математическое ожидание $\mathbb{E}(x_{k+1}|x_k)$ линейно по x_k . Тогда 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.

Итак, пусть имеется следующая система:

$$\begin{cases} x_{k+1} = A_k x_k + w_k, \\ y_{k+1} = C_{k+1} x_{k+1} + v_{k+1}, \end{cases} \quad (7)$$

причём $x_0, w_0, \dots, w_{N-1}, v_0, \dots, v_{N-1}$ 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, его надо найти. Для этого воспользуемся так называемым фильтром Калмана, а точнее, его схемой «шагаем–мерим», общий вид которой совпадает с системой (7).

Обозначим $\mathbb{E}x_0 = \bar{x}_0, \mathbb{D}x_0 = S, \mathbb{E}w_k = \mathbb{E}v_k = 0, \mathbb{D}w_k = M_k, \mathbb{D}v_k = N_k > 0$. Согласно [13], фильтр Калмана для схемы «шагаем–мерим» имеет вид:

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

В нашей задаче x_k — процесс Орнштейна–Уленбека с параметрами σ_W и λ , $y_{k+1} = x_{k+1} + v_{k+1}$, где v — белый шум. Пусть σ_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

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

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

Тогда фильтр Калмана для нашей задачи имеет вид:

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

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

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

Обратим внимание, что доверительным интервалом здесь будет отрезок

$$\hat{x}_k + k_{1-\alpha/2} \times [-\sqrt{R_{k|k}}, \sqrt{R_{k|k}}],$$

где k_α — квантиль стандартного нормального распределения.

Результаты работы фильтра Калмана для различных параметров представлены на рисунках 41–43 для белого гауссовского шума и на рисунках 44–46 для белого шума, имеющего распределение Коши (в этом случае фильтр Калмана не работает, так как распределение Коши не является гауссовским и не имеет конечных моментов, в качестве обобщённой дисперсии здесь брался квадрат коэффициента масштаба γ_n).

10.3 Примеры работы программы

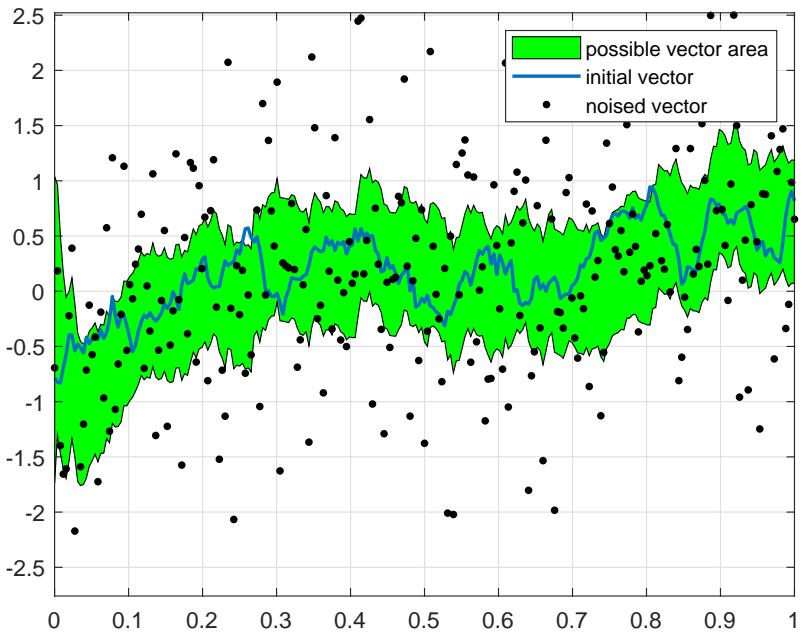


Рис. 41: Результат применения фильтра Калмана к процессу Орнштейна–Уленбека со случайной ошибкой в виде белого гауссовского шума, $\sigma_W = 1$, $\lambda = 1$, $\alpha = 5\%$, $\sigma_n = 1$

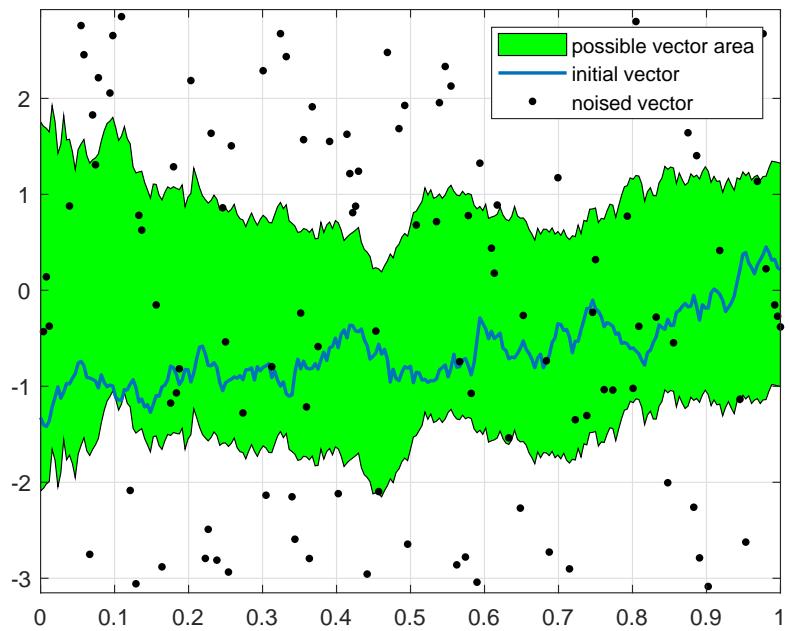


Рис. 42: Результат применения фильтра Калмана к процессу Орнштейна–Уленбека со случайной ошибкой в виде белого гауссовского шума, $\sigma_W = 1$, $\lambda = 1$, $\alpha = 5\%$, $\sigma_n = 5$

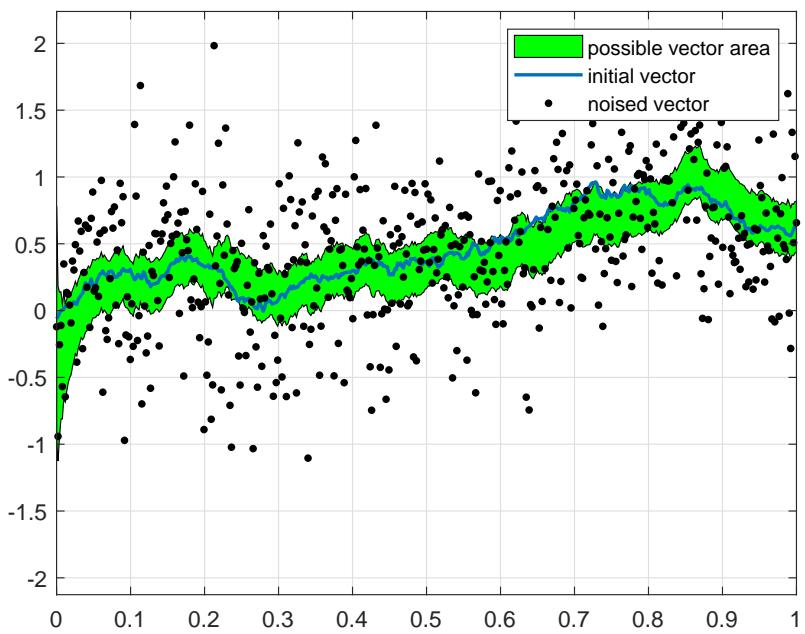


Рис. 43: Результат применения фильтра Калмана к процессу Орнштейна–Уленбека со случайной ошибкой в виде белого гауссовского шума, $\sigma_W = 1$, $\lambda = 0.1$, $\alpha = 5\%$, $\sigma_n = 0.1$

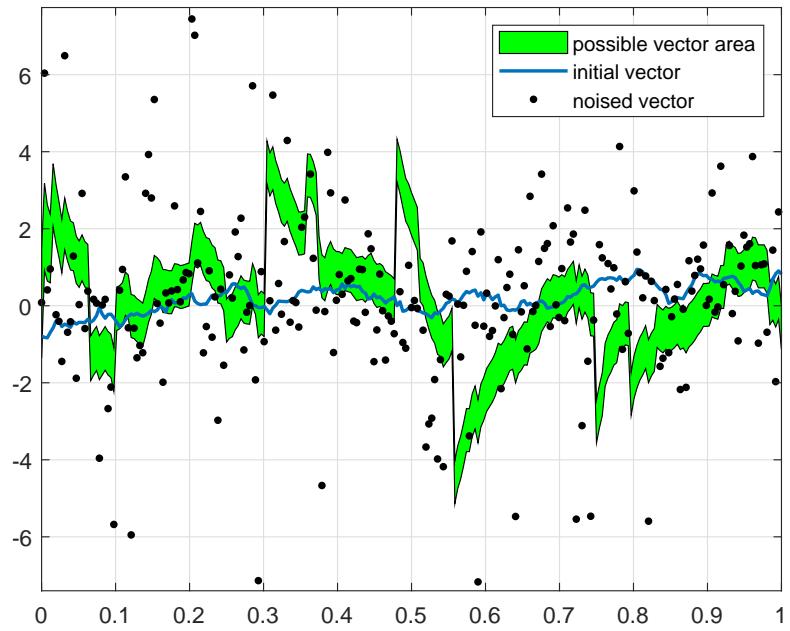


Рис. 44: Результат применения фильтра Калмана к процессу Орнштейна–Уленбека со случайной ошибкой в виде белого шума, имеющего распределение Коши, $\sigma_W = 1$, $\lambda = 1$, $\alpha = 5\%$, $\gamma_n = 1$

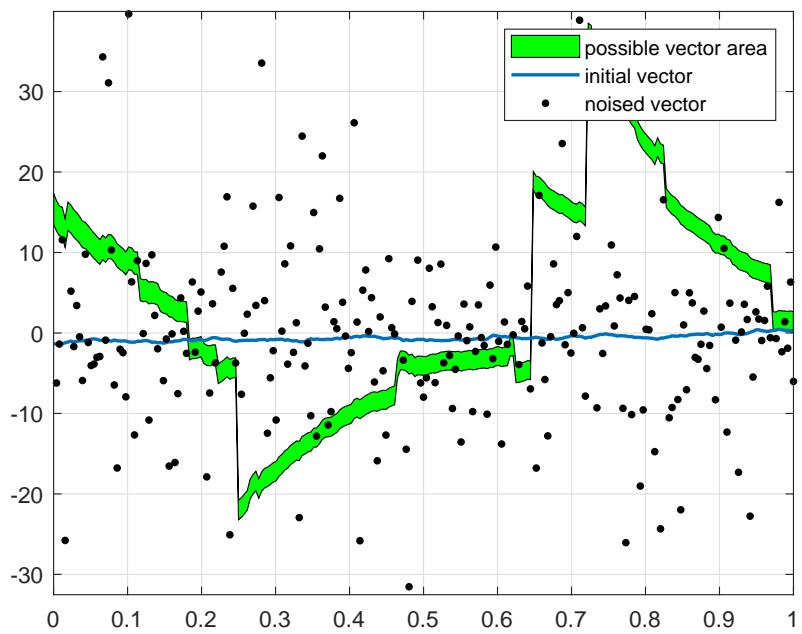


Рис. 45: Результат применения фильтра Калмана к процессу Орнштейна–Уленбека со случайной ошибкой в виде белого шума, имеющего распределение Коши, $\sigma_W = 1$, $\lambda = 1$, $\alpha = 5\%$, $\gamma_n = 5$

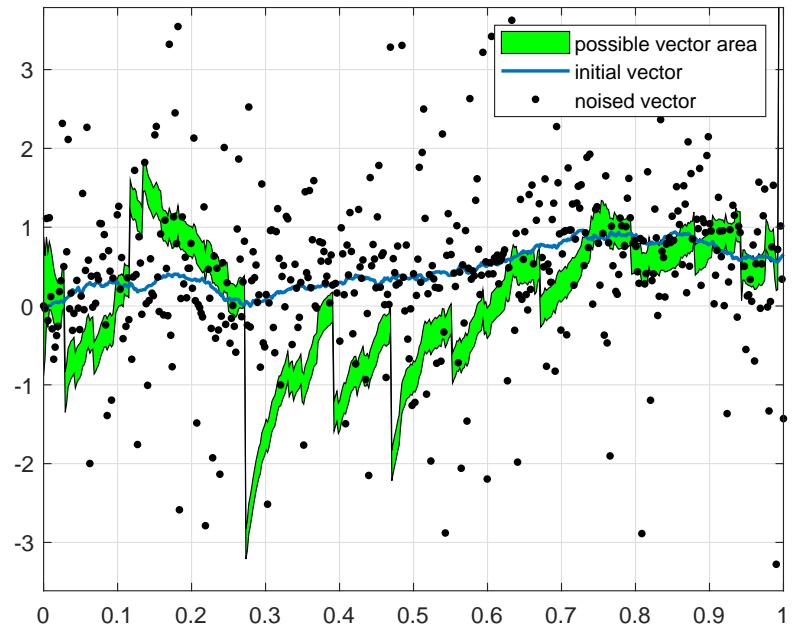


Рис. 46: Результат применения фильтра Калмана к процессу Орнштейна–Уленбека со случайной ошибкой в виде белого шума, имеющего распределение Коши, $\sigma_W = 1$, $\lambda = 0.1$, $\alpha = 5\%$, $\gamma_n = 0.1$

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 её обслуживания (распределение χ^2 с 10-ю степенями свободы).
2. Вторая интерпретация: система массового обслуживания с циклической интенсивностью $\lambda(t) = \lambda_0(1 + \cos(t))$ и единичными скачками. Свести данную задачу моделирования неоднородного пуассоновского процесса при помощи метода Льюиса и Шедлера к моделированию двумерного пуассоновского поля, где первая координата имеет равномерное распределение, а вторая — распределение Бернули.
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 Теория и необходимая аналитика

Смоделируем систему массового обслуживания по следующему алгоритму:

1. Сгенерируем времена поступления заявок в систему массового обслуживания на временном интервале $[0, T]$:

$$0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T,$$

причём $t_i - t_{i-1} \sim \text{Exp}(\lambda)$, $\lambda > 0$ — интенсивность потока заявок.

Время обработки s_i каждой заявки смоделируем с помощью распределения χ^2 с 10-ю степенями свободы.

2. Для каждой заявки будем считать время её исполнения Q_i . Здесь возможны два варианта:

- Если $Q_{i-1} < t_i$, то есть к моменту поступления i -й заявки очереди нет, то

$$Q_i = t_i + s_i.$$

- Если $Q_{i-1} \geq t_i$, то есть очередь есть, то

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

3. Для каждой заявки также посчитаем количество людей в очереди. Здесь тоже возможны два варианта:

- Если во время поступления i -й заявки очереди не было, то положим $n_i = 0$.

- Если же очередь была, то

$$n_i = \sum_{j=1}^n \mathbb{I}(k < i, Q_k > t_i),$$

то есть количество ещё не выполненных к моменту времени t_i заявок.

Заметим, что среднее время поступления новой заявки равно

$$\mathbb{E}(t_i - t_{i-1}) = \frac{1}{\lambda},$$

а среднее время обслуживания заявки равно

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

Тогда возможно три случая:

1. При значениях $\lambda > 0,1$ среднее время обслуживания заявки будет больше, чем среднее время поступления новой заявки, и очередь будет бесконечно расти.
2. При значениях $\lambda < 0,1$ система будет справляться с потоком заявок.
3. При $\lambda = 0,1$ система будет находиться в состоянии равновесия.

Результаты моделирования системы массового обслуживания представлены на рисунках 47–49 в пункте 11.3.

Перейдём теперь к СМО с циклической интенсивностью $\lambda(1 + \cos(t))$. Обозначим

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

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 $F(x) = 1 - e^{-(\Lambda(t+x) - \Lambda(t))}$, $x \geq 0$, если функция $\Lambda(t)$ неограниченно возрастает с увеличением t . Это следует из того факта, что

$$F(x) = \mathbb{P}(T_{n+1} - T_n > x | T_n = t) + \mathbb{P}(N(t, t+x) = 0 | T_n = t) = e^{-(\Lambda(t+x) - \Lambda(t))}.$$

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

К сожалению, аналитического выражения для обратной функции $\Lambda^{-1}(y)$ найти не удаётся, поэтому будем искать её численно. Заметим, что она существует, так как производная функции $\Lambda(t)$ (которая равна $\lambda(t) = \lambda_0(1 + \cos(t))$) почти всюду положительна (то есть функция возрастает).

Описанный выше метод моделирования неоднородного процесса Пуассона называется методом Льюиса–Шедлера.

Чтобы не искать обратную функцию, можно воспользоваться следующей модификацией метода Льюиса–Шедлера. Пусть имеется переменная t , в которой хранится текущее время (но не обязательно событие произошло строго в это время). На каждом шаге генерируем случайную

величину $\xi \sim \text{Exp}(2\lambda_0)$. Затем прибавляем к переменной t величину ξ и генерируем случайную величину $\eta = \text{Bern}((1 + \cos t)/2)$. Если она приняла значение 1, то полагаем $T_{i+1} = t$ и $i = i + 1$, иначе ничего не делаем и повторяем процесс заново.

Результаты моделирования системы массового обслуживания с циклической интенсивностью первым способом представлены на рисунке 50, а вторым способом — на рисунке 51. Как мы видим, эти способы позволяют добиться эквивалентного результата.

Наконец, рассмотрим работу страховой компании.

Определение 24. Случайная величина ξ называется случайной величиной, имеющей распределение Парето с параметрами x_m и k , если её функция распределения имеет вид

$$F_\xi(x) = \mathbb{P}(\xi < x) = 1 - \left(\frac{x_m}{x}\right)^k, \quad \forall x \geq x_m, \quad x_m, k > 0.$$

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

$$0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T,$$

причём $t_i - t_{i-1} \sim \text{Exp}(\lambda)$, $\lambda > 0$ — 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, распределённую по Парето, будем генерировать, воспользовавшись методом обратных функций (теорема 5):

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

Очевидно, 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 λ, x_m, k, W_0, c . Будем считать, что $k > 1$. Тогда

$$\frac{d\mathbb{E}W(t)}{dt} = c - \frac{d\mathbb{E}s(t)}{dt} = c - \frac{d}{dt} \left(\mathbb{E} \left[\sum_{t_i < t} s_i \right] \right) = c - \frac{d}{dt} \left(\frac{t}{\lambda} \mathbb{E}[s_i] \right) = c - \frac{\lambda k x_m}{k-1}.$$

В итоге получаем, что при $c(k-1) > \lambda k x_m$ капитал растёт, а при $c(k-1) < \lambda k x_m$ капитал уменьшается. При $c(k-1) = \lambda k x_m$ система находится в положении равновесия.

11.3 Примеры работы программы

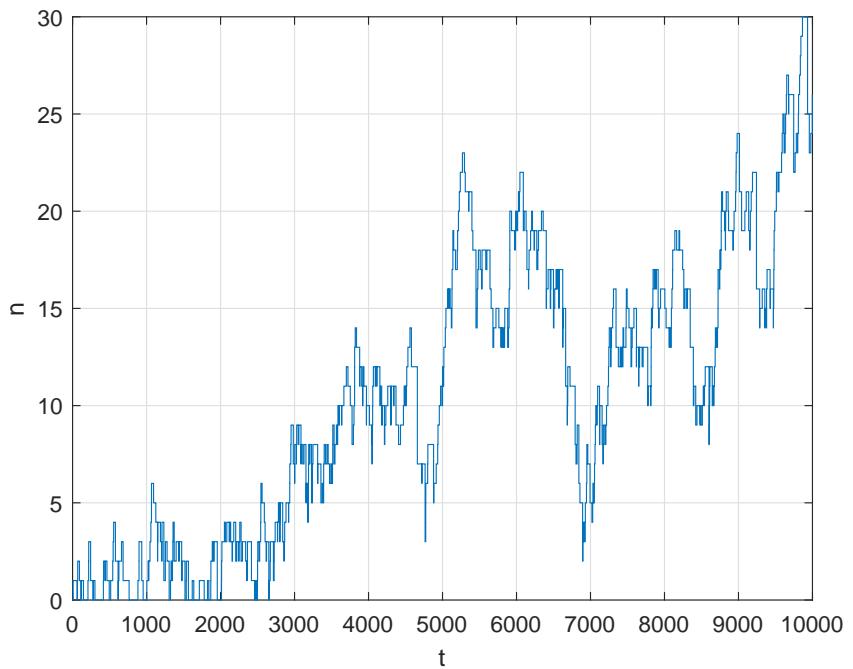


Рис. 47: Количество людей в системе массового обслуживания, $\lambda = 0,1$ (система в равновесии)

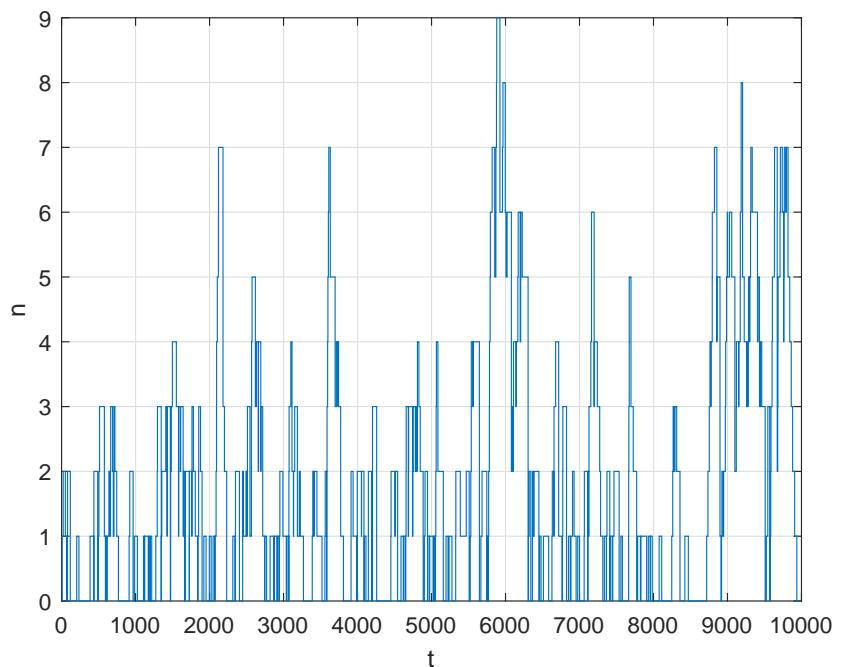


Рис. 48: Количество людей в системе массового обслуживания, $\lambda = 0,07$ (система справляется с потоком заявок)

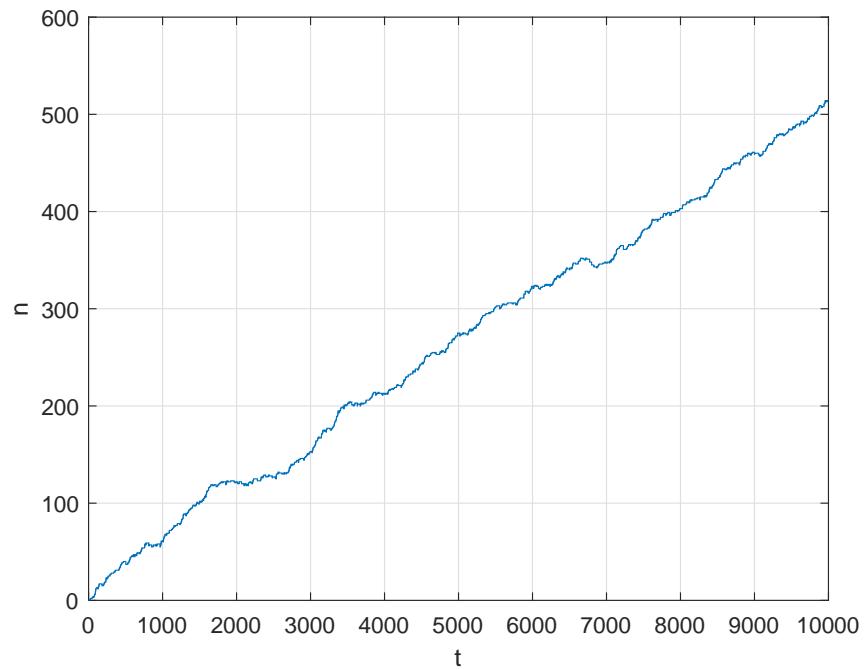


Рис. 49: Количество людей в системе массового обслуживания, $\lambda = 0,15$ (очередь бесконечно растёт)

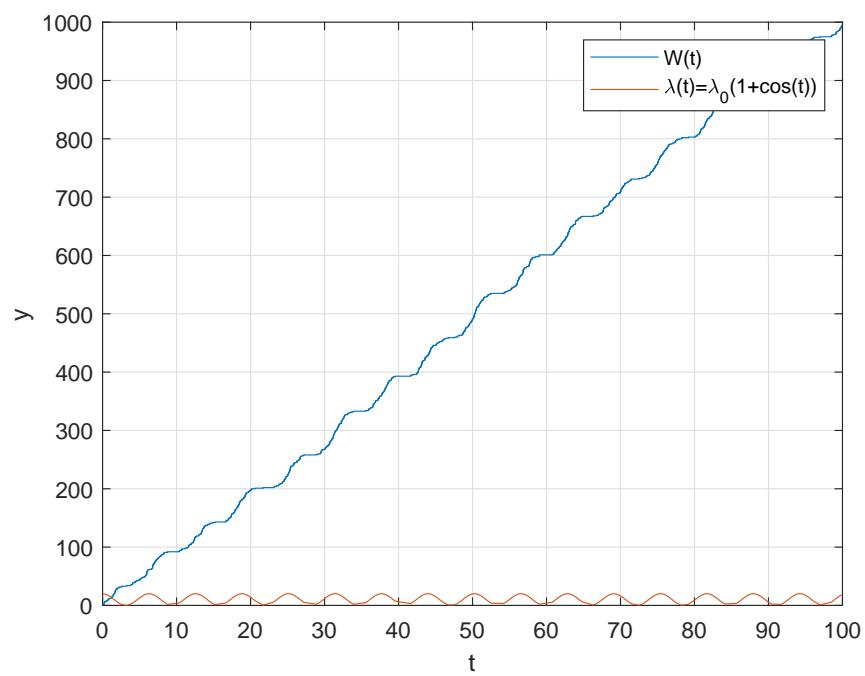


Рис. 50: Моделирование системы массового обслуживания с циклической интенсивностью, первый способ, $\lambda_0 = 10$

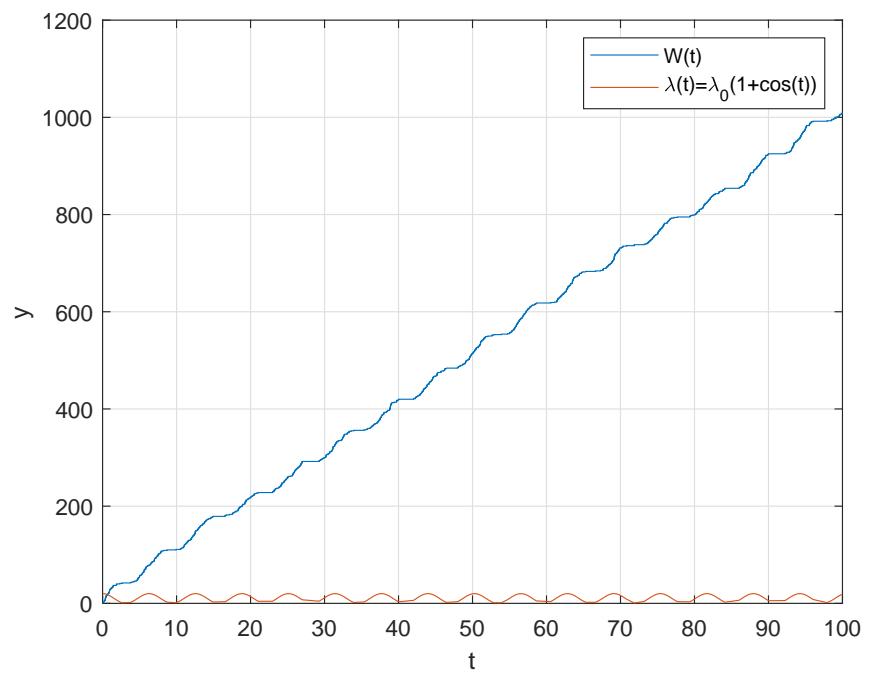


Рис. 51: Моделирование системы массового обслуживания с циклической интенсивностью, второй способ, $\lambda_0 = 10$

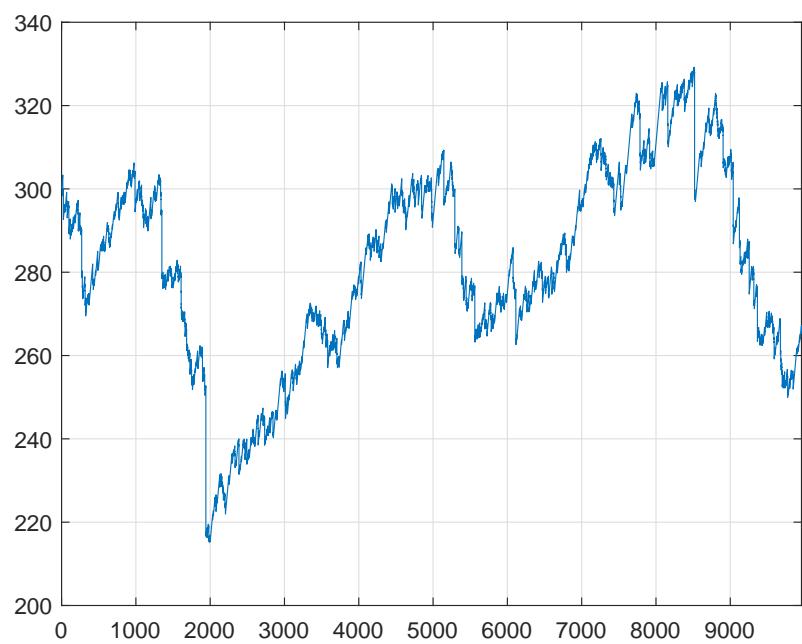


Рис. 52: Моделирование работы страховой компании, $c = 0,2$, $k = 2$, $x_m = 1$, $W_0 = 300$, $\lambda = 0,1$ (система в равновесии)

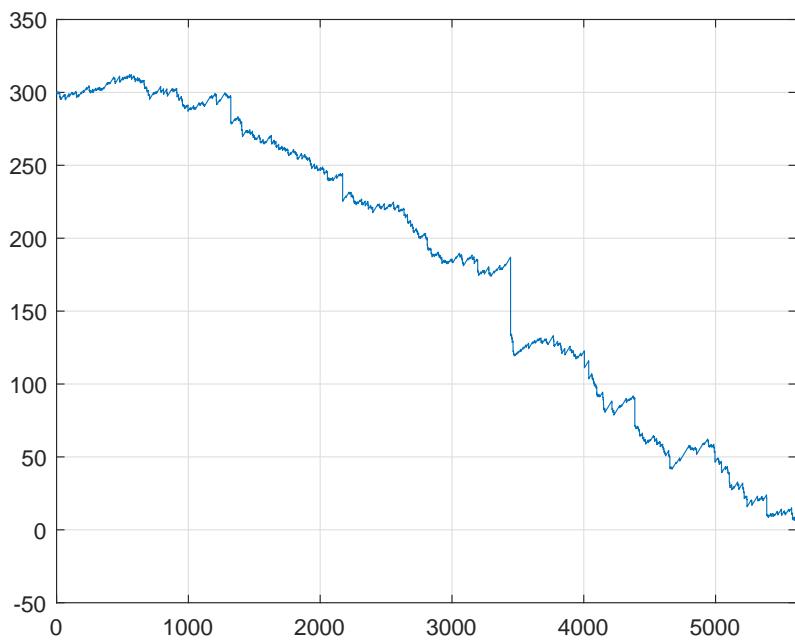


Рис. 53: Моделирование работы страховой компании, $c = 0,15$, $k = 2$, $x_m = 1$, $W_0 = 300$, $\lambda = 0,1$
(капитал уменьшается, останавливаемся, как только компания стала банкротом)

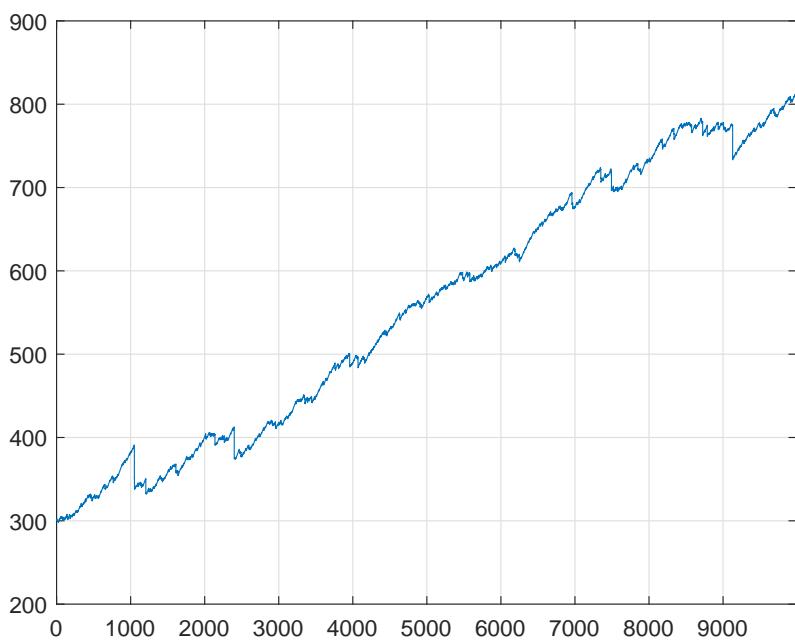


Рис. 54: Моделирование работы страховой компании, $c = 0,25$, $k = 2$, $x_m = 1$, $W_0 = 3000$, $\lambda = 0,1$
(капитал растёт)

Список литературы

- [1] Смирнов С. Н. *Лекции по курсу «Стохастический анализ и моделирование»*, 2017–2018.
- [2] Ширяев А. Н. *Вероятность*, Наука. М.: 1989.
- [3] Кропачёва Н. Ю., Тихомиров А. С. *Моделирование случайных величин: Метод. указания*, НовГУ им. Ярослава Мудрого, 2004.
- [4] Крамер Г. *Математические методы статистики*, М., Мир, 1975.
- [5] Самарский А. А., Гулин А. В. *Численные методы*, М., Наука, 1989.
- [6] Григорьев Е. А. *Введение в комплексный анализ*, М., факультет ВМК МГУ им. М. В. Ломоносова, 2015.
- [7] Колмогоров А. Н. *Избранные труды*, в 6 томах. Том 2. *Теория вероятностей и математическая статистика*. М., Математический институт им. В. А. Стеклова РАН.
- [8] Sanguthevar Rajasekaran *On the Convergence Time of Simulated Annealing*, University of Pennsylvania, 1990.
- [9] Зорич В. А. *Метаматематический анализ. Часть I*, М., МЦНМО, 2002.
- [10] Тихонов А. Н., Самарский А. А. *Уравнения математической физики*, М., МГУ, 2004.
- [11] Соболь И. М. *Численные методы Монте–Карло*, М., Наука, 1973.
- [12] Феллер В. *Введение в теорию вероятностей и её приложения*, в 2–х томах. Т.1, М., Мир, 1984.
- [13] Востриков И. В. *Лекции по курсу «Теория идентификации»*, 2018.