

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ  
АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО  
ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ  
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
ИТМО»

ФАКУЛЬТЕТ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО  
ИНТЕЛЛЕКТА



---

Расчётно-графическая работа № 1  
по дисциплине «Теория вероятностей  
и продвинутая математическая статистика»

---

Выполнили студенты группы J3211:

Воробьев А.П.

ИСУ: 465440

Шакина А.С.

ИСУ: 396675

Преподаватель практики:

Кононов И.А.

Санкт-Петербург

2025

# Содержание

Задача 1	3
Задача 2	4
Задача 3	7
Задача 4	9
Задача 5	13
Задача 6	21
Задача 7	24

# Задача 1

## Постановка задачи

Требуется аналитически решить задачу. Помимо этого нужно написать программу, которая будет считать данные вероятности приблизительно с помощью метода Монте-Карло. Особо любознательные могут задаться поиском (аналитическим) оптимального количества итераций, которых будет достаточно для получения ответа с достаточно «разумной» точностью. При решении задач могут быть полезны принципы произведения вероятностей и включений-исключений.

## Условие варианта 2

В городе живут  $n + 1$  людей. Человек, условный «прародитель», пишет два письма случайно выбранным адресатам, которые образуют «первое» поколение. Те делают то же самое, в результате чего образуется «второе поколение». В общем, на каждое полученное письмо горожанин готовит два ответа и отправляет двум случайным жителям. Найти вероятность того, что прародитель не входит ни в одно из поколений с номерами  $1 \dots r$ .

## Аналитическое решение

### Обозначения и допущения

Обозначения:

- $n + 1$  - количество жителей города.
- $r$  - количество поколений, которые мы рассматриваем.
- $S_i$  - число входящих писем в поколении  $i$ .
- $A$  - событие, что прародитель не является адресатом письма

Допущение: Отправитель письма не может быть его адресатом.

### Решение

Рассмотрим  $i$ -е поколение. Каждое входящее письмо порождает два исходящих, которые отправляются двум случайным жителям города (кро-

ме самого отправителя). Так как в первом поколении -  $S_1 = 2$  письма, то во втором поколении будет  $S_2 = 2 \cdot S_1 = 4$  и так далее. Таким образом, количество писем в  $i$ -м поколении можно выразить формулой:

$$S_i = 2 \cdot S_{i-1} = 2^i$$

Заметим, что в первом поколении не может быть прародителя, поэтому количество писем  $M$  за  $r$  поколений, адресатом для которого может быть прародитель, равно:

$$M = \sum_{i=2}^r S_i = \sum_{i=2}^r 2^i = 2^{r+1} - 4$$

Теперь найдем вероятность  $P(A)$ , что прародитель не является адресатом одного письма. Так как отправитель не может быть адресатом, то для каждого письма есть  $n$  возможных адресатов, из которых  $n - 1$  не являются прародителем. Значит:

$$P(A) = \frac{n - 1}{n}$$

Так как все адресаты выбираются независимо, то вероятность  $P(A_M)$ , что прародитель не является адресатом ни одного из  $M$  писем:

$$P(A_M) = P(A)^M = \left( \frac{n - 1}{n} \right)^{2^{r+1} - 4}$$

### Ответ

Таким образом, искомая вероятность, что прародитель не входит ни в одно из поколений с номерами  $1 \dots r$ , равна:

$$P(A_M) = \left( \frac{n - 1}{n} \right)^{2^{r+1} - 4}$$

## Приближённое решение методом Монте-Карло

Решение методом Монте-Карло реализовано в блокноте 1.ipynb.

# Задача 2

## Постановка задачи

Тоже как и в первом упражнении нужно решить аналитически и написать скрипт, вычисляющий вероятность приближенно по методу Монте-Карло. Дополнительный вопрос о поиске числа итерация тоже имеет место быть.

## Условие варианта 2

На отрезке длины  $l$  независимо одна от другой поставлены две точки, положение каждой из которых равновозможно на этом отрезке. Определить вероятность того, что из трех получившихся частей одного отрезка можно построить треугольник.

## Аналитическое решение

### Упрощения и обозначения

Обозначения:

- $x, y$  - случайные величины, соответствующие координатам точек на отрезке.
- $u = \min(x, y)$ ,  $v = \max(x, y)$  - упорядоченные координаты точек.
- $a = u$ ,  $b = v - u$ ,  $c = 1 - v$  - длины отрезков, на которые разбиваются исходный отрезок.
- $A$  - событие, что из отрезков можно построить треугольник.

Упрощения:

- Без потери общности можно считать, что  $l = 1$ , так как вероятность не зависит от длины отрезка.
- Две точки на отрезке можно рассматривать, как координаты на единичном квадрате  $[0, 1] \times [0, 1]$ .
- Поскольку пары  $(x, y)$  и  $(y, x)$  приводят к одинаковым отрезкам, то можно рассматривать только половину квадрата, где  $x \leq y$ . Площадь этой половины равна  $\frac{1}{2}$ .

## Решение

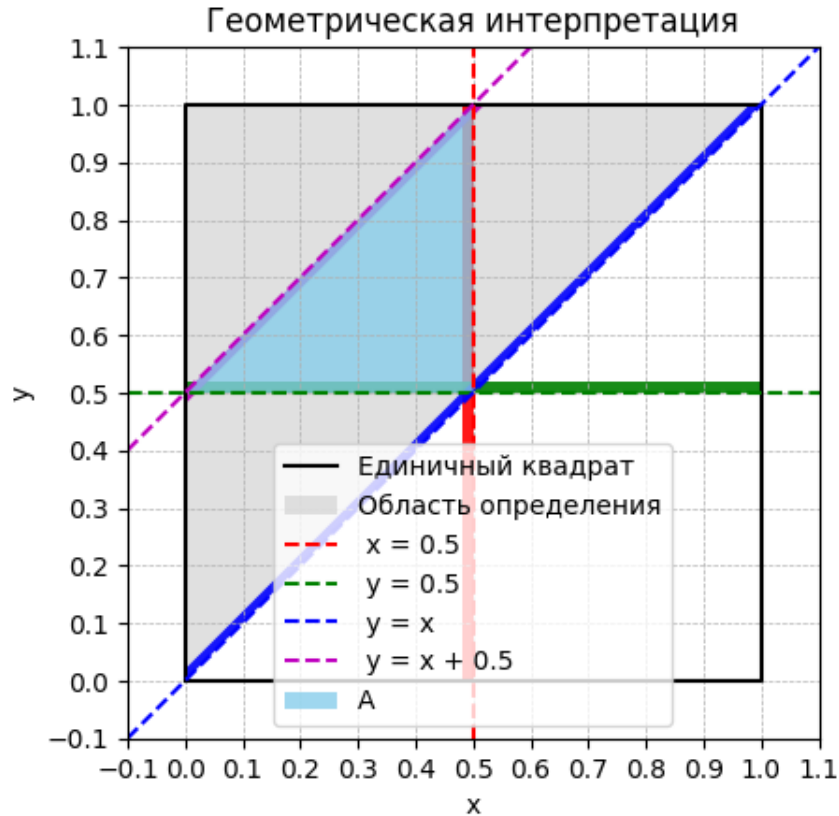
Из всего вышесказанного, мы имеем, что  $x = u$ ,  $y = v$ ,  $0 \leq x \leq y \leq 1$ ,  $a = x$ ,  $b = y - x$ ,  $c = 1 - y$ . Событие  $A$  - соблюдение неравенств треугольника можно записать как:

$$\begin{cases} a + b > c \leftrightarrow x + (y - x) > 1 - y \leftrightarrow y > \frac{1}{2} \\ a + c > b \leftrightarrow x + (1 - y) > y - x \leftrightarrow 2x + 1 > 2y \\ b + c > a \leftrightarrow (y - x) + (1 - y) > x \leftrightarrow 1 - x > x \leftrightarrow x < \frac{1}{2} \end{cases}$$

С учётом ограничений  $0 \leq x \leq y \leq 1$ , получаем систему неравенств:

$$\begin{cases} 0 \leq x < \frac{1}{2} \\ \frac{1}{2} < y \leq 1 \\ x \leq y \\ y < x + \frac{1}{2} \end{cases}$$

Построим область, удовлетворяющую системе неравенств:



Найдём область  $S_A$ , удовлетворяющую системе неравенств.

$$S_A = \int_{x=0}^{x=\frac{1}{2}} \int_{y=\frac{1}{2}}^{y=x+\frac{1}{2}} dydx = \int_{x=0}^{x=\frac{1}{2}} ((x + \frac{1}{2}) - \frac{1}{2})dx = \int_{x=0}^{x=\frac{1}{2}} xdx = \frac{1}{8} - 0 = \frac{1}{8}$$

Таким образом, вероятность  $P(A)$  равна отношению площади  $S_A$  к площади треугольной области, которая равна  $\frac{1}{2}$ :

$$P(A) = \frac{S_A}{\frac{1}{2}} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}$$

### Ответ

Таким образом, искомая вероятность, что из трёх частей отрезка можно построить треугольник, равна:

$$\boxed{P(A) = \frac{1}{4}}$$

## Приближённое решение методом Монте-Карло

Решение методом Монте-Карло реализовано в блокноте 2.ipynb.

## Задача 3

### Постановка задачи

Требуется аналитически решить задачу.

### Условие варианта 2

Введём события  $A_i = \{X = i\}$ ,  $B_i = \{Y = i\}$ ,  $i \geq 0$ . Известно, что для любых  $i \geq 0, j \geq 0$  события  $A_i$  и  $B_j$  – независимы, при этом

$$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}, \lambda > 0, i \geq 0,$$

$$P(Y = j) = e^{-\mu} \frac{\mu^j}{j!}, \mu > 0, j \geq 0.$$

Требуется найти  $P(X = i | X + Y = j)$ .

### Решение

Для нахождения вероятности  $P(X = i | X + Y = j)$  используем формулу условной вероятности:

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

причём  $A = \{X = i\}$ ,  $B = \{X + Y = j\}$ .

Найдём  $P(A \cap B)$ . Пересечение событий означает, что  $X$  должно быть равно  $i$ , и одновременно  $X + Y$  должно быть равно  $j$ . Это возможно только если  $Y = j - i$ .

Таким образом, событие  $\{X = i \cap X + Y = j\}$  эквивалентно событию  $\{X = i \cap Y = j - i\}$ . Подставим в формулу:

$$P(A \cap B) = e^{-\lambda} \frac{\lambda^i}{i!} \cdot e^{-\mu} \frac{\mu^{j-i}}{(j-i)!} = e^{-(\lambda+\mu)} \cdot \frac{\lambda^i \mu^{j-i}}{i!(j-i)!}$$

Найдём  $P(B) = P(X + Y = j)$ . Запишем событие  $\{X + Y = j\}$  как  $\{X = k \cap Y = j - k\} \quad \forall k \in 1, \dots, j$ . Так как события несовместны и



независимы:

$$P(X + Y = j) = \sum_{k=0}^j P(X = k \cap Y = j - k) = \sum_{k=0}^j P(X = k)P(Y = j - k)$$

По условию:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad P(Y = j - k) = e^{-\mu} \frac{\mu^{j-k}}{(j-k)!},$$

тогда

$$P(X + Y = j) = e^{-(\lambda+\mu)} \sum_{k=0}^j \frac{\lambda^k \mu^{j-k}}{k!(j-k)!}$$

Воспользуемся биномом Ньютона:

$$(\lambda + \mu)^j = \sum_{k=0}^j \frac{j!}{k!(j-k)!} \lambda^k \mu^{j-k}$$

Следовательно:

$$P(X + Y = j) = e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^j}{j!}.$$

Подставим полученные выражения в формулу условной вероятности:

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{e^{-(\lambda+\mu)} \frac{\lambda^i \mu^{j-i}}{i!(j-i)!}}{e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^j}{j!}} = \frac{j!}{i!(j-i)!} \cdot \frac{\lambda^i \mu^{j-i}}{(\lambda + \mu)^j} = \\ &= C_j^i \cdot \left(\frac{\lambda}{\lambda + \mu}\right)^i \cdot \left(\frac{\mu}{\lambda + \mu}\right)^{j-i} \end{aligned}$$

**Ответ**

Таким образом, искомая вероятность равна:

$$\boxed{C_j^i \cdot \left(\frac{\lambda}{\lambda + \mu}\right)^i \cdot \left(\frac{\mu}{\lambda + \mu}\right)^{j-i}}$$

# Задача 4

## Постановка задачи

Представьте, что вам нужно быстро создать некий программный продукт, в котором будут использоваться вероятностные модели. Возможно, вы будете использовать Python и модуль `SciPy`, в котором реализовано достаточно много математических алгоритмов и моделей, в том числе и вероятностно-статистических (подмодуль `stats`). Допустим, в вашей модели есть “нестандартное” распределение (например, не реализованное в `SciPy`), для которого известна плотность, и вам нужно быстро реализовать распределение, то есть определить функцию распределения, квантили и прочие методы. Тогда идея унаследоваться от класса `rv_continuous` кажется разумной: ведь достаточно переопределить метод, соответствующий плотности. Казалось бы, отличная идея, плюс применили объектно-ориентированный подход. Но допустим, что вам понадобилось сгенерировать случайные числа (на самом деле псевдослучайные) из данного распределения, но вдруг что-то пойдёт не так, например, будет слишком долго работать... И вдруг нужно будет сообразить другие методы.

*Само задание:*

1. Для заданной плотности  $p$  наследуйтесь от класса `rv_continuous` и сгенерируйте с помощью реализованного подкласса для вашего распределения  $n$  случайных чисел из данного распределения. Посмотрите, как будет работать алгоритм при росте  $n$  (сразу большим  $n$  не делайте). Убедитесь эмпирически, что распределение сгенерированных чисел согласуется с данным вероятностным законом.
2. Реализуйте генерацию для вашего распределения с помощью обратной к функции распределения (при написании функции для  $F^{-1}$  вызов специальных функций, например, квантили стандартного нормального закона разрешается; не забыть аналитический вывод  $F^{-1}$ ). Проведите тот же эксперимент.
3. Выберите ещё один метод для генерации случайных чисел (можно, например, `rejecting sampling`, `ratio of uniforms` или другой метод). Опишите его математическое обоснование. Воспользуйтесь реализацией

или сами реализуйте. Проведите тот же эксперимент.

## Условие варианта 2

$$p(x) = 3x^2 e^{-x^3} \cdot 1(x \geq 0)$$

## Аналитические вычисления

### Плотность распределения

$$p(x) = 3x^2 e^{-x^3} \cdot 1(x \geq 0)$$

### Функция распределения

$$F(x) = \int_{-\infty}^x p(t) dt = \int_0^x 3t^2 e^{-t^3} dt = -e^{-t^3} \Big|_0^x = 1 - e^{-x^3}$$

То есть:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x^3}, & x \geq 0 \end{cases}$$

### Обратная функция распределения

$$F(x) = 1 - e^{-x^3} \Leftrightarrow 1 - F(x) = e^{-x^3} \Leftrightarrow \ln(1 - F(x)) = -x^3 \Leftrightarrow x = (-\ln(1 - F(x)))^{1/3}$$

Тогда обратная функция распределения:

$$F^{-1}(y) = (-\ln(1 - y))^{1/3}, \quad y \in [0, 1)$$

## Метод rejection sampling

**Идея метода** Генерация семплов по  $f(x)$  происходит с помощью более простого вспомогательного распределения  $g(x)$ , которое мы можем просем-

плировать, и которое удовлетворяет следующему условию:

$$\forall x \ f(x) < cg(x), \quad c > 1$$

### Алгоритм

1. Взять семпл  $x$  по распределению  $g(x)$ ;
2. Выбрать случайное число  $u$  равномерно из отрезка  $[0, cg(x)]$ ;
3. Вычислить  $f(x)$ ;
4. Если  $u \leq f(x)$ , то  $x$  добавляется к семплам;
5. Если  $u > f(x)$ , то  $x$  отклоняется (отсюда и название метода).

**Обоснование** Пусть  $Y$  — случайная величина с плотностью распределения  $g(y)$ , а  $U$  — случайная величина, равномерно распределённая на отрезке  $[0, cg(x)]$ .

Мы хотим показать, что условная плотность  $Y$  при условии события  $U \leq f(Y)$  равна искомой плотности  $f(x)$ .

$$P(Y = y \mid U \leq f(Y)) = \frac{P(Y = y \cap U \leq f(Y))}{P(U \leq f(Y))}$$

Рассмотрим знаменатель. Заметим, что  $U$  и  $Y$  независимы, так как  $U$  выбирается независимо от  $Y$ .

$$\begin{aligned} P(U \leq f(Y)) &= \int_{-\infty}^{\infty} P(U \leq f(y))g(y)dy = \\ &= \int_{-\infty}^{\infty} \frac{f(y)}{cg(y)}g(y)dy = \frac{1}{c} \int_{-\infty}^{\infty} f(y)dy = \frac{1}{c} \end{aligned}$$

Рассмотрим числитель.

$$\begin{aligned} P(Y = y \cap U \leq f(Y)) &= P(U \leq f(Y) \mid Y = y) \cdot P(Y = y) = \\ &= P(U \leq f(y)) \cdot g(y) = \frac{f(y)}{cg(y)}g(y) = \frac{f(y)}{c} \end{aligned}$$

Теперь объединим и получим искомый результат:

$$P(Y = y \mid U \leq f(Y)) = \frac{f(y)}{c} \cdot c = f(y)$$

## Выбор вспомогательного распределения

В качестве вспомогательного распределения выберем экспоненциальное распределение  $g(x) = e^{-x}$ . Для метода необходимо найти константу  $c$  такую, что  $p(x) \leq c \cdot g(x)$  для всех  $x \geq 0$ . Это эквивалентно поиску максимума функции  $h(x) = \frac{p(x)}{g(x)}$ :

$$h(x) = \frac{3x^2 e^{-x^3}}{e^{-x}} = 3x^2 e^{x-x^3}$$

Найдём производную  $h(x)$ , чтобы найти её максимум:

$$\begin{aligned} h'(x) &= 6xe^{x-x^3} + 3x^2 e^{x-x^3} (1 - 3x^2) = \\ &= 3xe^{x-x^3} (2 + x(1 - 3x^2)) = 3xe^{x-x^3} (2 + x - 3x^3) \end{aligned}$$

Приравняем производную к нулю, чтобы найти критические точки,  $x \geq 0$ :

$$\begin{aligned} 3xe^{x-x^3} (2 + x - 3x^3) = 0 &\Leftrightarrow x = 0 \vee (x - 1)(3x^2 + 3x + 2) = 0 \Leftrightarrow \\ &\Leftrightarrow x = 0 \vee x = 1 \end{aligned}$$

- При  $x = 0$ ,  $h(0) = 0$ .
- При  $x = 1$ ,  $h(1) = 3(1)^2 e^{1-1} = 3$ .

Поскольку  $\lim_{x \rightarrow \infty} h(x) = 0$ , максимум функции достигается в точке  $x = 1$ . Таким образом, можно выбрать  $c = 3$ .

## Эмпирическое решение и выводы

Эксперименты, их результаты и выводы написаны блокноте 4.ipynb.

## Задача 5

### Постановка задачи

Для заданного распределения требуется найти:

- "нормирующие" константы;
- функцию распределения;
- математическое ожидание, дисперсию, среднее квадратическое отклонение, моду, медиану, асимметрию, эксцесс
- функцию распределения и плотность функции от исходной случайной величины

### Условие варианта 5

$$p_X(x) = c \cos x \cdot \text{ind}(x \in [-\frac{\pi}{2}, \frac{\pi}{2}]), Y = \sin X$$

### Решение

1. Найдём нормирующую константу  $c$ .

Так как  $p_X(x)$  – плотность, то  $\int_{-\pi/2}^{\pi/2} p_X(x) dx = 1$ . Подставим:

$$\int_{-\pi/2}^{\pi/2} c \cos x dx = 1$$

$$c \sin x \Big|_{-\pi/2}^{\pi/2} = 1$$

$$c(\sin(\pi/2) - \sin(-\pi/2)) = 1$$

$$c(1 - (-1)) = 1 \implies 2c = 1 \implies \boxed{c = \frac{1}{2}}$$

2. Найдём функцию распределения:

$$F_X(x) = \int_{-\pi/2}^x \frac{1}{2} \cos t dt = \frac{1}{2}(\sin x + 1).$$

Тогда:

$$F_X(x) = \begin{cases} 0, & x \leq -\frac{\pi}{2} \\ \frac{1}{2}(\sin x + 1), & -\frac{\pi}{2} < x < \frac{\pi}{2} \\ 1, & x \geq \frac{\pi}{2} \end{cases}$$

3. Математическое ожидание:

$$E[X] = \int_{-\infty}^{\infty} x \cdot p_X(x) dx$$

$$E[X] = \int_{-\pi/2}^{\pi/2} x \cdot \frac{1}{2} \cos x dx$$

Подынтегральная функция является нечётной, а интеграл по симметричному промежутку от нечётной функции равен нулю:

$$E[X] = 0$$

Дисперсия:

$$D[X] = E[X^2] - (E[X])^2$$

Так как  $E[X] = 0$ , то  $D[X] = E[X^2]$ .

$$E[X^2] = \int_{-\pi/2}^{\pi/2} x^2 \cdot \frac{1}{2} \cos x dx$$

Так как подынтегральная функция - чётная:

$$E[X^2] = 2 \cdot \frac{1}{2} \int_0^{\pi/2} x^2 \cos x dx = \int_0^{\pi/2} x^2 \cos x dx$$

Два раза интегрируем по частям:

$$\begin{aligned} u = x^2, dv = \cos x dx &\implies du = 2x dx, v = \sin x \\ \int_0^{\pi/2} x^2 \cos x dx &= x^2 \sin x \Big|_0^{\pi/2} - \int_0^{\pi/2} 2x \sin x dx = \\ &= \frac{\pi^2}{4} - 2 \int_0^{\pi/2} x \sin x dx \end{aligned}$$

$$\begin{aligned}
u = x, dv = \sin x dx &\implies du = dx, v = -\cos x \\
\int_0^{\pi/2} x \sin x dx &= -x \cos x \Big|_0^{\pi/2} - \int_0^{\pi/2} (-\cos x) dx = \\
&= \left(-\frac{\pi}{2} \cdot 0 - 0\right) + (\sin x \Big|_0^{\pi/2}) = 0 + (\sin \frac{\pi}{2} - \sin 0) = 1 - 0 = 1
\end{aligned}$$

Подставим обратно:

$$D[X] = E[X^2] = \frac{\pi}{4} - 2 \cdot 1 = \frac{\pi}{4} - 2$$

Среднее квадратическое отклонение:

$$\sigma_X = \sqrt{D[X]} = \sqrt{\frac{\pi}{4} - 2}$$

Мода:

$$mod_X = \arg \max_x p_X(x) = \arg \max_x \frac{1}{2} \cos x = 0$$

Медиана:

$$med = \arg \min_a E[|X - a|] = \arg \min_a \int_{-\infty}^{\infty} |x - a| p_X(x) dx$$

Для данной плотности  $p_X(x) = \frac{1}{2} \cos x \cdot ind\left(x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$ , функция для минимизации:

$$g(a) = \int_{-\pi/2}^{\pi/2} |x - a| \cdot \frac{1}{2} \cos x dx$$



Дифференцируем  $g(a)$  по  $a$ :

$$\begin{aligned} & \frac{\partial}{\partial a}|x-a| = \\ & = \begin{cases} \frac{\partial}{\partial a}(-(x-a)), & x < a \\ \frac{\partial}{\partial a}(x-a), & x > a \end{cases} = \\ & = \begin{cases} +1, & x < a \\ -1, & x > a \end{cases} \\ g'(a) &= \frac{1}{2} \left( \int_{-\pi/2}^a \cos x dx - \int_a^{\pi/2} \cos x dx \right) \end{aligned}$$

Минимум достигается при  $g'(a) = 0$ :

$$\begin{aligned} g'(a) = 0 &\implies \int_{-\pi/2}^a \cos x dx = \int_a^{\pi/2} \cos x dx \\ \sin x \Big|_{-\pi/2}^a &= \sin x \Big|_a^{\pi/2} \\ \sin a - \sin \left(-\frac{\pi}{2}\right) &= \sin \left(\frac{\pi}{2}\right) - \sin a \\ \sin a - (-1) &= 1 - \sin a \\ \sin a + 1 &= 1 - \sin a \\ \sin a &= 0 \\ \sin a &= 0 \end{aligned}$$

Единственное решение для  $a \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ :

$$a = 0$$

Следовательно:

$$\boxed{med = 0}$$

Коэффициент асимметрии:

$$As = \frac{E[(X - E[X])^3]}{\sigma_X^3}$$

Поскольку  $E[X] = 0$ :

$$E[(X - E[X])^3] = E[X^3] = \int_{-\pi/2}^{\pi/2} x^3 \cdot p_X(x) dx = \int_{-\pi/2}^{\pi/2} x^3 \cdot \frac{1}{2} \cos x dx$$

Подынтегральная функция  $f(x) = x^3 \cos x$  является нечетной. Интеграл от нечетной функции по симметричному интервалу  $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$  равен нулю:

$$E[(X - E[X])^3] = 0$$

Следовательно, коэффициент асимметрии равен нулю:

$$As = \frac{0}{\sigma_X^3} = 0$$

То есть, распределение является строго симметричным.

Коэффициент эксцесса:

$$Ex = \frac{E[(X - E[X])^4]}{\sigma_X^4} - 3$$

Поскольку  $E[X] = 0$ :

$$E[(X - E[X])^4] = E[X^4] = \int_{-\pi/2}^{\pi/2} x^4 \cdot \frac{1}{2} \cos x dx$$

Подынтегральная функция  $f(x) = x^4 \cos x$  является четной.

$$E[(X - E[X])^4] = 2 \cdot \frac{1}{2} \int_0^{\pi/2} x^4 \cos x dx = \int_0^{\pi/2} x^4 \cos x dx$$

Вычислим значение интеграла:

$$I_n = \int_0^{\pi/2} x^n \cos x dx = x^n \sin x \Big|_0^{\pi/2} - n \int_0^{\pi/2} x^{n-1} \sin x dx$$

$$J_m = \int_0^{\pi/2} x^m \sin x dx = -x^m \cos x \Big|_0^{\pi/2} + m \int_0^{\pi/2} x^{m-1} \cos x dx$$

Заметим  $\cos(\pi/2) = 0$  и для  $m \geq 1 : x^m = 0$ :

$$J_m = mI_{m-1}$$

$$I_n = \left(\frac{\pi}{2}\right)^n - nJ_{n-1} = \left(\frac{\pi}{2}\right)^n - n(n-1)I_{n-2}$$

Вычислим  $I_4$ :

$$I_0 = \int_0^{\pi/2} \cos x dx = \sin x|_0^{\pi/2} = 1 - 0 = 1$$

$$I_2 = \left(\frac{\pi}{2}\right)^2 - 2 \cdot 1 = \frac{\pi^2}{4} - 2$$

$$I_4 = \left(\frac{\pi}{2}\right)^4 - 4 \cdot 3 \left(\frac{\pi^2}{4} - 2\right) = \frac{\pi^4}{16} - 3\pi^2 + 24$$

Таким образом:

$$E[(X - E[X])^4] = \frac{\pi^4}{16} - 3\pi^2 + 24$$

Теперь подставляем  $E[(X - E[X])^4]$  и  $\sigma_X^4 = D[X]^2 = \left(\frac{\pi^2}{4} - 2\right)^2 = \frac{(\pi^2 - 8)^2}{16}$  в формулу:

$$\begin{aligned} Ex &= \frac{\frac{\pi^4}{16} - 3\pi^2 + 24}{\frac{(\pi^2 - 8)^2}{16}} - 3 \\ &= \frac{\pi^4 - 48\pi^2 + 384}{(\pi^2 - 8)^2} - 3 \\ &= \frac{(\pi^4 - 48\pi^2 + 384) - 3(\pi^4 - 16\pi^2 + 64)}{(\pi^2 - 8)^2} \\ &= \frac{\pi^4 - 48\pi^2 + 384 - 3\pi^4 + 48\pi^2 - 192}{(\pi^2 - 8)^2} \\ &= \frac{-2\pi^4 + 192}{(\pi^2 - 8)^2} \end{aligned}$$

Окончательный результат для коэффициента эксцесса:

$$Ex = \frac{192 - 2\pi^4}{(\pi^2 - 8)^2}$$

Численно:  $Ex \approx -0.806$ , что указывает на более "плоскую" вершину по сравнению с нормальным распределением.

4. Функция распределения  $F_Y(y)$  определяется как  $P(Y \leq y)$ .

1. Случай  $y < -1$ :  $\sin X$  всегда больше  $-1$ .

$$F_Y(y) = P(Y \leq y) = 0$$

2. Случай  $-1 \leq y \leq 1$ : Используем обратную функцию  $x = \arcsin y$ .

$$F_Y(y) = P(Y \leq y) = P(\sin X \leq y)$$

Так как  $\sin x$  возрастает на  $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ :

$$F_Y(y) = P(X \leq \arcsin y)$$

По определению функции распределения исходной величины  $X$ ,  $P(X \leq x) = F_X(x) = \int_{-\pi/2}^x p_X(t) dt$ :

$$\begin{aligned} F_Y(y) &= F_X(\arcsin y) = \int_{-\pi/2}^{\arcsin y} \frac{1}{2} \cos t dt \\ &= \frac{1}{2} \sin t \Big|_{-\pi/2}^{\arcsin y} \\ &= \frac{1}{2} \left( \sin(\arcsin y) - \sin\left(-\frac{\pi}{2}\right) \right) \\ &= \frac{1}{2}(y + 1) \end{aligned}$$

3. Случай  $y > 1$ :  $\sin X$  всегда меньше  $1$ .

$$F_Y(y) = P(Y \leq y) = 1$$

Полная функция распределения  $Y$ :

$$F_Y(y) = \begin{cases} 0, & y < -1 \\ \frac{1}{2}(y + 1), & -1 \leq y \leq 1 \\ 1, & y > 1 \end{cases}$$

Плотность распределения находится как производная функции распределения:  $p_Y(y) = \frac{d}{dy}F_Y(y)$ .

Для  $y \in (-1, 1)$ :

$$p_Y(y) = \frac{d}{dy} \left( \frac{1}{2}(y+1) \right) = \frac{1}{2}$$

Таким образом, плотность распределения  $Y$ :

$$p_Y(y) = \begin{cases} \frac{1}{2}, & y \in [-1, 1] \\ 0, & y \notin [-1, 1] \end{cases}$$

## Задача 6

### Постановка задачи

Пусть случайная величина  $R$  имеет распределение Рэлея, то есть распределение с плотностью

$$f(x) = \frac{x}{\sigma^2} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} \text{ind}(X \geq 0),$$

где  $\sigma > 0$  – масштабирующий параметр, и пусть  $\theta \sim U[0, 2\pi]$ . Показать, что величины  $X = R \cos \theta$  и  $Y = R \sin \theta$  независимы и распределены нормально с нулевым математическим ожиданием и дисперсией  $\sigma^2$ , если  $R$  и  $\Theta$  независимы.

### Решение

По условию  $R$  имеет распределение Рэлея и  $\Theta \sim U[0, 2\pi]$ , независимы, поэтому

$$p_{R,\Theta}(r, \theta) = f_R(r) \cdot \frac{1}{2\pi} = \frac{1}{2\pi} \cdot \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0, \theta \in [0, 2\pi).$$

Перейдём к  $(X, Y)$ . Пусть:

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

Заметим, что это является переходом между декартовыми и полярными координатами. Из известной информации можно сделать вывод, что отображение  $g : (r, \theta) \mapsto (x, y)$  является диффеоморфизмом.

Обратное преобразование задаётся формулами:

$$r = \sqrt{x^2 + y^2}, \quad \theta = \begin{cases} \arctan \frac{y}{x}, & x > 0 \\ \arctan \frac{y}{x} + \pi, & x < 0, y \geq 0 \\ \arctan \frac{y}{x} - \pi, & x < 0, y < 0 \\ \frac{\pi}{2}, & x = 0, y > 0 \\ -\frac{\pi}{2}, & x = 0, y < 0 \end{cases}$$

Матрица Якоби  $(r, \theta) \mapsto (x, y)$  и соответственно якобиан:

$$Dg(r, \theta) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

$$\det Dg(r, \theta) = r(\cos^2 \theta + \sin^2 \theta) = r$$

Для обратного преобразования  $(x, y) \mapsto (r, \theta)$  имеем

$$|\det Dg^{-1}(x, y)| = \frac{1}{r}.$$

Тогда по заданной из условия задачи формуле для плотности совместного распределения  $(X, Y)$ :

$$\begin{aligned} p_{X,Y}(x, y) &= p_{R,\Theta}(r(x, y), \theta(x, y)) \cdot |\det D(g^{-1})(x, y)| = \\ &= \frac{1}{2\pi} \cdot \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) \cdot \frac{1}{r} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \\ p_X(x) &= \int_{-\infty}^{+\infty} p_{X,Y}(x, y) dy = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) dy \\ &= \frac{1}{2\pi\sigma^2} e^{-x^2/(2\sigma^2)} \int_{-\infty}^{+\infty} e^{-y^2/(2\sigma^2)} dy \\ &= \frac{1}{2\pi\sigma^2} e^{-x^2/(2\sigma^2)} \cdot \sqrt{2\pi}\sigma = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)}. \end{aligned}$$

Аналогично,

$$p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2/(2\sigma^2)}.$$

Тогда

$$p_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)} = p_X(x) p_Y(y).$$

Следовательно,  $X$  и  $Y$  независимы.

Из вида маргинальных плотностей следует:

$$X \text{ имеет плотность } p_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)},$$

$$Y \text{ имеет плотность } p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2/(2\sigma^2)}.$$

Следовательно,

$$E[X] = E[Y] = 0, \quad \text{Var}(X) = \text{Var}(Y) = \sigma^2.$$



# Задача 7

## Постановка задачи

Дано совместное распределение дискретных случайных величин  $(X, Y)$ . Требуется найти:

- Пропущенное значение вероятности
- Маргинальные распределения для  $X$  и  $Y$
- Распределение случайной величины  $U = g(X, Y)$  для заданной функции  $g$
- Математическое ожидание, дисперсию, среднее квадратическое отклонение, моду, медиану, коэффициенты асимметрии и эксцесса случайных величин  $X$  и  $Y$
- Ковариацию, коэффициент корреляции между  $X$ ,  $Y$
- Исследовать на независимость

## Условие варианта 7

Таблица вероятностей  $p_{ij} = P(Y = y_i, X = x_j)$ :

$Y \setminus X$	-8	1	13	28
-9	0.04	0.04	0.11	0.01
-8	0.09	0.04	0.01	0.02
-5	0.07	0.06	0.01	0.03
-3	0.03	0.07	0.09	0.08
-2	0.11	0.05	0.03	NULL

## Решение

1. Нахождение пропущенной вероятности

Сумма известных вероятностей:

$$\sum_{\text{известные}} p_{ij} = 0.99.$$

Используя условие  $\sum_{i,j} p_{ij} = 1$ , получаем:

$$1 = 0.99 + p(Y = -2, X = 28),$$

$$\boxed{p(Y = -2, X = 28) = 0.01.}$$

## 2. Маргинальные распределения

$$P(X = -8) = 0.04 + 0.09 + 0.07 + 0.03 + 0.11 = 0.34,$$

$$P(X = 1) = 0.04 + 0.04 + 0.06 + 0.07 + 0.05 = 0.26,$$

$$P(X = 13) = 0.11 + 0.01 + 0.01 + 0.09 + 0.03 = 0.25,$$

$$P(X = 28) = 0.01 + 0.02 + 0.03 + 0.08 + 0.01 = 0.15.$$

Проверка:  $0.34 + 0.26 + 0.25 + 0.15 = 1$ .

$$P(Y = -9) = 0.04 + 0.04 + 0.11 + 0.01 = 0.20,$$

$$P(Y = -8) = 0.09 + 0.04 + 0.01 + 0.02 = 0.16,$$

$$P(Y = -5) = 0.07 + 0.06 + 0.01 + 0.03 = 0.17,$$

$$P(Y = -3) = 0.03 + 0.07 + 0.09 + 0.08 = 0.27,$$

$$P(Y = -2) = 0.11 + 0.05 + 0.03 + 0.01 = 0.20.$$

Проверка:  $0.20 + 0.16 + 0.17 + 0.27 + 0.20 = 1$ .

## 3. Распределение $U = (-1)^{X^2+Y^4}$

По определению:

$$U = \begin{cases} 1, & \text{если } X \text{ и } Y \text{ имеют одинаковую чётность,} \\ -1, & \text{если чётности разные.} \end{cases}$$

Чётности значений:

- $X$ : чётные  $\{-8, 28\}$ , нечётные  $\{1, 13\}$ ;
- $Y$ : чётные  $\{-8, -2\}$ , нечётные  $\{-9, -5, -3\}$ .

Вычислим  $P(U = 1)$  — сумму вероятностей ячеек, где чётности совпадают.

$$P(U = 1) = P_{\text{чёт/чёт}} + P_{\text{неч/неч}}.$$

Чёт/чёт:  $X \in \{-8, 28\}$ ,  $Y \in \{-8, -2\}$ .

$$\begin{aligned} P_{\text{чёт/чёт}} &= p(-8, -8) + p(-8, -2) + p(28, -8) + p(28, -2) \\ &= 0.09 + 0.11 + 0.02 + 0.01 = 0.23. \end{aligned}$$

Нечёт/нечёт:  $X \in \{1, 13\}$ ,  $Y \in \{-9, -5, -3\}$ .

$$\begin{aligned} P_{\text{неч/неч}} &= p(1, -9) + p(1, -5) + p(1, -3) + p(13, -9) + p(13, -5) + p(13, -3) \\ &= 0.04 + 0.06 + 0.07 + 0.11 + 0.01 + 0.09 = 0.38. \end{aligned}$$

$$P(U = 1) = 0.23 + 0.38 = 0.61,$$

$$P(U = -1) = 1 - 0.61 = 0.39.$$

$P(U = 1) = 0.61, \quad P(U = -1) = 0.39.$

#### 4. Характеристики $X$

$$E[X] = (-8) \cdot 0.34 + 1 \cdot 0.26 + 13 \cdot 0.25 + 28 \cdot 0.15 = 4.99,$$

$$E[X^2] = 64 \cdot 0.34 + 1 \cdot 0.26 + 169 \cdot 0.25 + 784 \cdot 0.15 = 181.87,$$

$$D(X) = E[X^2] - (E[X])^2 = 181.87 - 4.99^2 = 156.97,$$

$$\sigma_X = \sqrt{156.97} = 12.53.$$

Коэффициенты асимметрии и эксцесса:

$$As_X = \frac{E[(X - E[X])^3]}{\sigma_X^3}, \quad Ex_X = \frac{E[(X - E[X])^4]}{\sigma_X^4} - 3.$$

Вычислим  $E[(X - E[X])^3]$ :

$$\begin{aligned} E[(X - E[X])^3] &= \sum_i (x_i - E[X])^3 P(X = x_i) = \\ &= (-8 - 4.99)^3 \cdot 0.34 + (1 - 4.99)^3 \cdot 0.26 + \\ &\quad + (13 - 4.99)^3 \cdot 0.25 + (28 - 4.99)^3 \cdot 0.15 = \\ &= 1194.139098 \end{aligned}$$

Вычислим  $E[(X - E[X])^4]$ :

$$\begin{aligned} E[(X - E[X])^4] &= \\ &= (-12.99)^4 \cdot 0.34 + (-3.99)^4 \cdot 0.26 + (8.01)^4 \cdot 0.25 + (23.01)^4 \cdot 0.15 = \\ &= 52825.12138 \end{aligned}$$

Подставим в формулы:

$$\begin{aligned} \text{As}_X &= \frac{1194.139098}{(12.53)^3} = \frac{1194.139098}{1967.221277} = 0.607, \\ \text{Ex}_X &= \frac{52825.12138}{(12.53)^4} - 3 = \frac{52825.12138}{24649.2826} - 3 = -0.857. \end{aligned}$$

$\text{As}_X = 0.61, \quad \text{Ex}_X = -0.86.$

## 5. Характеристики $Y$

$$\begin{aligned} E[Y] &= (-9) \cdot 0.20 + (-8) \cdot 0.16 + (-5) \cdot 0.17 + (-3) \cdot 0.27 + (-2) \cdot 0.20 = -5.14, \\ E[Y^2] &= 81 \cdot 0.20 + 64 \cdot 0.16 + 25 \cdot 0.17 + 9 \cdot 0.27 + 4 \cdot 0.20 = 33.92, \\ D(Y) &= 33.92 - (-5.14)^2 = 7.50, \\ \sigma_Y &= \sqrt{7.50} = 2.74. \end{aligned}$$

$$\text{As}_Y = \frac{E[(Y - E[Y])^3]}{\sigma_Y^3}, \quad \text{Ex}_Y = \frac{E[(Y - E[Y])^4]}{\sigma_Y^4} - 3.$$

Вычислим  $E[(Y - E[Y])^3]$ :

$$\begin{aligned} E[(Y - E[Y])^3] &= \sum_i (y_i - E[Y])^3 P(Y = y_i) = \\ &= (-9 - (-5.14))^3 \cdot 0.20 + (-8 - (-5.14))^3 \cdot 0.16 + \\ &\quad + (-5 - (-5.14))^3 \cdot 0.17 + (-3 - (-5.14))^3 \cdot 0.27 + \\ &\quad + (-2 - (-5.14))^3 \cdot 0.20 = \\ &= -6.407088 \end{aligned}$$

Вычислим  $E[(Y - E[Y])^4]$ :

$$\begin{aligned} E[(Y - E[Y])^4] &= \\ (-3.86)^4 \cdot 0.20 + (-2.86)^4 \cdot 0.16 + (-0.14)^4 \cdot 0.17 + 2.14^4 \cdot 0.27 + 3.14^4 \cdot 0.20 &= \\ &= 80.2096 \end{aligned}$$

Подставим в формулы:

$$\begin{aligned} \text{As}_Y &= \frac{-6.407088}{(2.74)^3} = \frac{-6.407088}{20.570824} = -0.312, \\ \text{Ex}_Y &= \frac{80.2096}{(2.74)^4} - 3 = \frac{80.2096}{56.364} - 3 = -1.574. \end{aligned}$$

$$\boxed{\text{As}_Y = -0.312, \quad \text{Ex}_Y = -1.574.}$$

## 6. Ковариация и корреляция

$$E[XY] = \sum_{i,j} x_j y_i p_{ij} = -24.70.$$

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = -24.70 - (4.99 \cdot (-5.14)) = 0.95.$$

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{0.95}{12.53 \cdot 2.74} = 0.028.$$

$$\boxed{\text{Cov}(X, Y) = 0.95, \quad \rho_{XY} = 0.028.}$$

## 7. Исследование на независимость

Случайные величины  $X$  и  $Y$  независимы тогда и только тогда, когда для всех значений  $x_j$  и  $y_i$  выполняется

$$p_{ij} = P(Y = y_i, X = x_j) = P(Y = y_i) P(X = x_j).$$

Проверим условие независимости на нескольких конкретных ячейках таблицы.

1) Ячейка ( $Y = -9$ ,  $X = -8$ ):

$$p(Y = -9, X = -8) = 0.04,$$

а произведение

$$P(Y = -9) P(X = -8) = 0.20 \cdot 0.34 = 0.068.$$

Так как  $0.04 \neq 0.068$ , равенство  $p_{ij} = P(Y = y_i)P(X = x_j)$  нарушено уже для этой ячейки.

2) Ячейка  $(Y = -3, X = 28)$ :

$$p(Y = -3, X = 28) = 0.08,$$

а

$$P(Y = -3) P(X = 28) = 0.27 \cdot 0.15 = 0.0405,$$

и снова  $0.08 \neq 0.0405$ .

Вывод. Поскольку существует по крайней мере одна пара  $(y_i, x_j)$ , для которой

$$p(Y = y_i, X = x_j) \neq P(Y = y_i)P(X = x_j),$$

величины  $X$  и  $Y$  не являются независимыми.

$X$ и $Y$ зависимы.
---------------------