

Лекция 1. Необходимые сведения из теории вероятностей.

Вероятностные пространства. Основные определения.

Тройка объектов (Ω, \mathcal{F}, P) будет обозначать вероятностное пространство. Здесь:

- Ω - *пространство состояний* (или *пространство элементарных событий*). Например, при моделировании физической системы Ω – это ансамбль состояний рассматриваемой системы. Часто это гигантское пространство, недоступное для эксперимента. Каждую точку пространства Ω часто называют *элементарным событием*.
- \mathcal{F} - совокупность подмножеств множества Ω , каждое такое подмножество называется *событием*. Вообще говоря, \mathcal{F} содержит не все подмножества множества Ω . В конкретной ситуации это множество всех результатов экспериментов, которые можно осуществить над системой.
- P – *вероятность*, для всякого события $A \in \mathcal{F}$ вещественное число $P(A)$ можно интерпретировать как степень правдоподобия события A , $0 \leq P(A) \leq 1$. Математически, вероятность – это отображение $P: \mathcal{F} \rightarrow [0,1]$ такое, что $P(\Omega) = 1$, и для каждой системы попарно непересекающихся событий $A_i \in \mathcal{F}$

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i).$$

Пример. Бросание трёх игральных костей. В качестве Ω можно взять

$$\Omega = \{(i, j, k) | i, j, k \in \{1, 2, 3, 4, 5, 6\}\}.$$

В качестве \mathcal{F} рассмотрим совокупность всех подмножеств множества Ω . Предположим, что все элементарные события

$\{i, j, k\}$ равновероятны. То есть $P(\{i, j, k\}) = \frac{1}{6^3}$. Заметим, что нельзя писать $P(\{i, j, k\}) = \frac{1}{6^3}$ для $\{i, j, k\} \notin \mathcal{F}$.

Пример события: «две шестёрки»:

$$\{(i, 6, 6), (6, j, 6), (6, 6, k) | i, j, k \neq 6\}.$$

Вопрос: сколько элементов в этом событии?

Заметим, что Ω в примере содержит 216 элементов. Если мы не учитываем порядок расположения выпавших очков, то есть то, на какой кости что выпало, то, например, результат (i, j, k) - это тот же самый результат, что и (j, k, i) . Так будет в случае, если мы подбросили три кости в воздух и не запоминаем, где какая кость (все кости одного цвета). В этом случае пространство Ω будет содержать меньше элементов.

Задача 1. Сколько элементов будет содержать Ω в этом случае? Мы будем требовать, чтобы совокупность \mathcal{F} была **σ -алгеброй**. Это означает, что

- i) $\Omega \in \mathcal{F}$,
- ii) $A \in \mathcal{F} \rightarrow \Omega \setminus A \in \mathcal{F}$,
- iii) $A_1, A_2, \dots \in \mathcal{F} \rightarrow \bigcup_n A_n \in \mathcal{F}$.

Аксиомы для P :

- i) $P(A) \in [0, 1], A \in \mathcal{F}$,
- ii) $P(\Omega) = 1$,
- iii) A_1, \dots, A_n, \dots **непересекающиеся** элементы \mathcal{F} , тогда

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n).$$

Из этих аксиом следует, в частности, что $P(\Omega \setminus A) = 1 - P(A)$.

Случайные величины.

Отображение $X: \Omega \rightarrow R$ является случайной величиной, если оно \mathcal{F} - измеримо. Это означает, что $\{\omega \in \Omega: X(\omega) \leq \alpha\} \in \mathcal{F}$ для всех $\alpha \in R$. В этом случае функция распределения $F(x)$ случайной величины X определяется следующим образом

$$F(x) = P(X \leq x) = P(\omega \in \Omega: X(\omega) \leq x).$$

Простые случайные величины.

Случайная величина называется простой, если она принимает конечное число значений. Это означает, что мы можем написать $X = \sum_{i=1}^n c_i I_{A_i}$, где

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

$A_1 \cup A_2 \cup \dots \cup A_n = \Omega$ – объединение непересекающихся измеримых множеств, в сумме дающих всё пространство Ω . Набор множеств $\{A_1, A_2, \dots, A_n\}$ иногда называют разбиением пространства Ω , соответствующим простой с.в. X .

Аппроксимация случайных величин простыми случайными величинами.

Заметим, что каждую случайную величину X можно записать в виде разности двух неотрицательных величин

$$X = X^+ - X^-,$$

где $X^+ = \max(X, 0)$, $X^- = \max(-X, 0)$ - неотрицательные случайные величины.

Для каждой неотрицательной с.в. $X \geq 0$ существует возрастающая последовательность с.в. $\{X_n\}$ такая, что $X_n \uparrow X$.

Имеет место сходимость $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega), \forall \omega \in \Omega$.

Такая последовательность с.в. строится следующим образом:

$$X_n(\omega) = \begin{cases} n & \text{для } X(\omega) \geq n, \\ \frac{k-1}{2^n} & \text{для } \frac{k-1}{2^n} \leq X(\omega) < \frac{k}{2^n}, k = 1, \dots, 2^n n. \end{cases}$$

Математические ожидания случайных величин.

Основная идея: если

$$X = \begin{cases} a & \text{на } A, P(A) = p, \\ b & \text{на } \Omega \setminus A, P(\Omega \setminus A) = 1 - p, \end{cases}$$

то есть

$$X = a \cdot I_A + b \cdot I_{\Omega \setminus A},$$

тогда положим

$$E[X] = a \cdot E[I_A] + b \cdot E[I_{\Omega \setminus A}] = a \cdot P(A) + b \cdot P(\Omega \setminus A).$$

Это определение обобщается на простые с.в.:

$$X = \sum_i c_i \cdot I_{A_i} \Rightarrow E[X] = \sum_i c_i \cdot P(A_i).$$

Если X и Y простые с.в. и $X \geq Y$, то $E[X] \geq E[Y]$. В самом деле, без ограничения общности, можем предполагать, что разбиения, соответствующие X и Y , совпадают (почему?). В этом случае

$$X = \sum_i c_i \cdot I_{A_i}, \quad Y = \sum_i d_i \cdot I_{A_i}, \quad c_i \geq d_i,$$

откуда сразу следует неравенство $E[X] \geq E[Y]$.

Пример.

$$\begin{aligned} X &= 0 \cdot I_{(-\infty, 1]} + 1 \cdot I_{(1, \infty)} = 0 \cdot I_{(-\infty, 1]} + 1 \cdot I_{(1, 2]} + 1 \cdot I_{(2, \infty)} \\ Y &= 0 \cdot I_{(-\infty, 2]} + \frac{1}{2} \cdot I_{(2, \infty)} = 0 \cdot I_{(-\infty, 1]} + 0 \cdot I_{(1, 2]} + \frac{1}{2} \cdot I_{(2, \infty)} \end{aligned}$$

Очевидно, $X \geq Y$ и $EX > EY$.

Неотрицательные случайные величины.

Пусть $X \geq 0$ и $X_n \uparrow X$, как это было выше. Поскольку $\{E[X_n]\}$ возрастающая последовательность неотрицательных чисел, она сходится (возможно, к $+\infty$). Мы затем можем положить

$$E[X] = \lim_{n \rightarrow \infty} E[X_n].$$

Аналогично, для произвольной с.в. $X = X^+ - X^-$ можно определить

$$E[X^+] = \lim_{n \rightarrow \infty} E[X_n^+],$$

$$E[X^-] = \lim_{n \rightarrow \infty} E[X_n^-].$$

Теперь положим

$$E[X] = E[X^+] - E[X^-]. \quad (1)$$

Правая часть последнего равенства не определена, если $E[X^+] = E[X^-] = +\infty$. Однако, всегда $|X| = X^+ + X^-$. Потребуем, чтобы $E\{|X|\} < \infty$. Тогда $E[X^+] \leq E[|X|] < \infty$ и $E[X^-] \leq E[|X|] < \infty$. Поэтому если X абсолютно интегрируемая с.в., то $E[X]$ в (1) конечно.

Один полезный результат.

Если $X \geq 0$ и $E[X] = 0$, то $X = 0$ с вероятностью 1. Это означает, что $P(\{\omega \in \Omega : X(\omega) \neq 0\}) = 0$.

Доказательство. Пусть

$$A_n = \left\{ \omega \in \Omega : X(\omega) \geq \frac{1}{n} \right\} \in \mathcal{F}.$$

Далее,

$$0 = E[X] \geq E[X \cdot I_{A_n}] \geq \frac{1}{n} \cdot E[I_{A_n}] =$$

$$\frac{1}{n} \cdot P(A_n) \Rightarrow P(A_n) = 0.$$

$$\{\omega \in \Omega : X(\omega) \neq 0\} = \cup_n A_n,$$

$$P(\cup_n A_n) \leq \sum_n P(A_n) = 0.$$

Последнее неравенство доказывается следующим образом:

$$\cup_n A_n = \cup_n (A_n - \cup_{k < n} A_k) \Rightarrow P(\cup_n A_n) =$$

$$\sum_n P(A_n - \cup_{k < n} A_k) \leq \sum_n P(A_n).$$

Поэтому

$$P\{\omega \in \Omega : X(\omega) \neq 0\} = 0 \text{ и } X = 0 \text{ п. н.}$$

Функции распределения.

Функция распределения (ф.р.) с.в. X определяется следующим образом:

$$F(x) = P(X \leq x).$$

Эта функция непрерывна справа.

Замечание. Некоторые авторы определяют функцию распределения иначе

$$F(x) = P(X < x).$$

Так определённая ф.р. будет непрерывна слева.

Если $X \geq 0$, то для ступенчатых с.в. X_n , определённых выше, имеем:

$$E[X_n] = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \cdot P(A_{k,n}) + nP(B_n),$$

где

$$B_n = \{\omega \in \Omega : X(\omega) \geq n\}, A_{k,n} = \left\{ \omega \in \Omega : \frac{k-1}{2^n} \leq X(\omega) < \frac{k}{2^n} \right\}$$

Имеем

$$\begin{aligned} E[X_n] &= \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \cdot \left(F\left(\frac{k}{2^n}\right) - F\left(\frac{k-1}{2^n}\right) \right) + nP(B_n) + \\ &\quad \frac{1}{2^n} \sum_{k=1}^{n2^n} P\left(X = \frac{k}{2^n}\right) - nP(X = n) = \\ &\quad \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \cdot \Delta F\left(\frac{k-1}{2^n}\right) + nP(B_n) + \frac{1}{2^n} \sum_{k=1}^{n2^n} P\left(X = \frac{k}{2^n}\right) \\ &\quad - nP(X = n). \end{aligned}$$

Переходя к пределу при $n \rightarrow \infty$ и пользуясь определением интеграла Римана-Стилтьеса, получим

$$E(X) = \int_0^{\infty} x dF(x).$$

Если существует $F'(x) = f(x)$, то говорят, что ф.р. имеет плотность $f(x)$. В этом случае

$$E(X) = \int_0^{\infty} x f(x) dx.$$

Для произвольной интегрируемой с.в. X , $E\{|X|\} < \infty$, получим тот же самый результат

$$E(X) = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x) dx.$$

Заметим, что в вычислениях, приведённых выше: $Var[X^{\pm}] < \infty$ влечёт $n \cdot P(B_n^{\pm}) \rightarrow 0, n \rightarrow \infty$.

Задача 2. Доказать это утверждение. Указание: воспользоваться неравенством Чебышева (см. Лекцию 2).

Если с.в. X имеет плотность вероятности f , то мы будем писать $X \sim f$.

Как проверить, что $X \sim f$?

Предложение 1.1.

$$X \sim f \Leftrightarrow E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \quad (2)$$

для любой ограниченной и непрерывной функции g .

Для дальнейшего напомним некоторые определения: $\Omega = R$, пусть \mathcal{F} — σ - алгебра, порождённая *открытыми* множествами в R . Тогда \mathcal{F} называется σ - алгеброй **борелевских** множеств пространства Ω . Это понятие естественным образом обобщается на случай $\Omega = R^n$.

Будем обозначать $\sigma_{\mathcal{A}}$ — σ - алгебру, порождённую совокупностью подмножеств \mathcal{A} . Это наименьшая σ - алгебра, содержащая \mathcal{A} :

$$\sigma_{\mathcal{A}} = \bigcap_{\mathcal{G}} \{ \mathcal{G} \text{ является } \sigma - \text{ алгеброй, } \mathcal{G} \supset \mathcal{A} \}.$$

Будем говорить, что $g: R \rightarrow R$ измерима по Борелю (или g борелевская) $\Leftrightarrow \{x \in R: g(x) \leq \alpha\}$ — борелевское множество для каждого $\alpha \in R$.

Эквивалентная формулировка Предложения 1.1 состоит в том, что (2) выполнено для любой ограниченной борелевской функции g , поскольку такие функции являются поточечными пределами равномерно ограниченной последовательности непрерывных функций.

Задача. Доказать, что любая ограниченная борелевская функция является поточечным пределом равномерно ограниченной последовательности непрерывных функций.

Доказательство Предложения 1.1. Для доказательства " \Leftarrow " выберем

$$g(x) = \begin{cases} 1 & \text{если } x \leq \alpha \\ 0 & \text{если } x > \alpha \end{cases}$$

Тогда

$$E[g(X)] = P(X \leq \alpha)$$

и

$$\int_{-\infty}^{\infty} g(x)f(x)dx = \int_{-\infty}^{\alpha} f(x)dx.$$

Таким образом, $P(X \leq \alpha) = \int_{-\infty}^{\alpha} f(x)dx$. Для доказательства " \Rightarrow " заметим, что

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)dF(x) = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Пример 1. $X \sim N(0,1)$, то есть имеет плотность

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Пусть

$$Y = \begin{cases} X, & \text{если } |X| \leq c \\ -X, & \text{если } |X| > c \end{cases}$$

$c > 0$. Найти плотность распределения с.в. Y .

Решение. Пусть g - произвольная ограниченная борелевская функция.

$$E[g(Y)] = E[g(Y)I_{|X| \leq c} + g(Y)I_{|X| > c}] =$$

$$E[g(Y)I_{|X| \leq c}] + E[g(Y)I_{|X| > c}] =$$

$$E[g(X)I_{|X| \leq c}] + E[g(-X)I_{|X| > c}] =$$

$$\int_{-\infty}^{\infty} g(x)I_{|x| \leq c} \phi(x)dx + \int_{-\infty}^{\infty} g(-x)I_{|x| > c} \phi(x)dx =$$

$$\int_{-c}^c g(x) \phi(x) dx + \int_c^{\infty} g(y) \phi(y) dy +$$

$$\int_{-\infty}^{-c} g(y) \phi(y) dy = \int_{-\infty}^{\infty} g(x) \phi(x) dx.$$

$$\Rightarrow Y \sim N(0,1).$$

Замечания.

- а) Часто можно услышать утверждение, что если $X \sim N(0,1), Y \sim N(0,1)$, то с.в. $X + Y$ распределена нормально. **Это неверно.** Из рассмотренного выше примера следует

$$X + Y = \begin{cases} 2X, & \text{если } |X| \leq c, \\ 0, & \text{если } |X| > c. \end{cases}$$

- б) Отметим, что в этом примере

$$E[XY] = E[XYI_{|X| \leq c}] + E[XYI_{|X| > c}] =$$

$$E[X^2I_{|X| \leq c}] + E[-X^2I_{|X| > c}] =$$

$$\int_{-c}^c x^2 \phi(x) dx + \int_{-\infty}^{-c} -x^2 \phi(x) dx + \int_c^{\infty} -x^2 \phi(x) dx,$$

т.е. $E[XY] \rightarrow 1$, если $c \rightarrow \infty$, и $\rightarrow -1$, если $c \rightarrow 0 +$.

Поэтому из непрерывности по c следует, что существует $c^* > 0$ такое, что $E[XY] = 0$, тогда

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0,$$

откуда следует, что при таком c^* случайные величины X и Y некоррелированы, но, очевидно, они не являются независимыми.

с) Для того, чтобы утверждение п. а) было верно, необходимо, чтобы X и Y имели **совместное** нормальное распределение.

d) Если

$$\begin{bmatrix} X \\ Y \end{bmatrix} = A \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix},$$

где Z_1 и Z_2 независимые нормально распределённые с.в., а A - 2×2 - матрица, то X и Y имеют совместное нормальное распределение.

Независимость случайных величин.

Две с.в. X и Y на (Ω, \mathcal{F}, P) называются независимыми, если

$$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y).$$

Часто используют следующий результат:

X и Y независимы тогда и только тогда, когда для любых двух ограниченных борелевских функций h и g

$$E[h(X) \cdot g(Y)] = E[h(X)] \cdot E[g(Y)].$$

Пример 2. Пусть с.в. X_1, X_2, X_3 независимы и имеют распределение $N(0,1)$. Найти плотность распределения с.в. X

$$X = \frac{X_1 + X_2 X_3}{\sqrt{1 + X_3^2}}.$$

Решение. Пусть g ограниченная борелевская функция. Тогда

$$E[g(X)] =$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g\left(\frac{x_1 + x_2 x_3}{\sqrt{1 + x_3^2}}\right) \phi(x_1) \phi(x_2) \phi(x_3) dx_1 dx_2 dx_3.$$

Пусть $u = \frac{x_1 + x_2 x_3}{\sqrt{1 + x_3^2}}$. Преобразуя последний интеграл (проделайте

соответствующие преобразования в качестве упражнения), получим

$$E[g(X)] = \int_{-\infty}^{\infty} g(u)\phi(u)du.$$

Таким образом, согласно Предложению 1.1, $X \sim N(0,1)$.

Можно дать и другое, более короткое, доказательство того, что $X \sim N(0,1)$, использующее свойства условных математических ожиданий относительно σ -алгебр (Лекция 3).

«Заморозим» с.в. X_3 , тогда $X = aX_1 + bX_2$, $a = \frac{1}{\sqrt{1+X_3^2}}$, $b = \frac{X_3}{\sqrt{1+X_3^2}}$,

$a^2 + b^2 = 1$. Тогда условное (при фиксированном значении X_3) распределение X является стандартным нормальным, т.е.

$X \sim N(0,1)$. Используя свойства условных математических ожиданий, получим

$$E[g(X)] = E[E[g(X)|X_3]] =$$

$$E\left[\int_{-\infty}^{\infty} g(u)\phi(u)du\right] = \int_{-\infty}^{\infty} g(u)\phi(u)du.$$

Таким образом, согласно Предложению 1.1., $X \sim N(0,1)$.

Лемма Дуба-Дынкина.

Случайная величина X является отображением из Ω в R (или из Ω в R^n , если речь идёт о многомерной случайной величине).

Рассмотрим множества

$$A_\alpha = \{\omega \in \Omega | X(\omega) \leq \alpha\}$$

для различных $\alpha \in R$.

Пусть σ_X (будем использовать также обозначение $\sigma(X)$) - наименьшая σ -алгебра, содержащая все множества A_α . Скажем, что $g: \Omega \rightarrow R$ является $\sigma(X)$ -измеримой, если $\{\omega \in \Omega: g(\omega) \leq \alpha\} \in \sigma(X)$. Следующая важная лемма носит название Леммы Дуба-Дынкина.

Предложение 1.2. Случайная величина $g(\omega)$ является $\sigma(X)$ -измеримой тогда и только тогда, когда $g = h \circ X$ для некоторой

измеримой по Борелю (или, кратко, борелевской) функции $h: R \rightarrow R$.

Заметим, что $g = h \circ X$ обозначает супрепозицию, то есть $g(\omega) = h(X(\omega)), \forall \omega \in \Omega$.

Обобщение леммы Дуба-Дынкина на случай n случайных величин очевидно. Для $\alpha_1, \alpha_2, \dots, \alpha_n \in R$ положим

$$A_{\alpha_1, \alpha_2, \dots, \alpha_n} = \{\omega \in \Omega: X_1(\omega) \leq \alpha_1, \dots, X_n(\omega) \leq \alpha_n\} \quad (3)$$

Пусть $\sigma(X_1, \dots, X_n)$ - наименьшая σ - алгебра, содержащая все множества вида (3). Тогда g является $\sigma(X_1, \dots, X_n)$ измеримой тогда и только тогда, когда

$$g(\omega) = h(X_1(\omega), \dots, X_n(\omega)) \text{ для всех } \omega \in \Omega. \quad (4)$$

для некоторой измеримой по Борелю функции $h: R^n \rightarrow R$.

Что можно сказать, если $n = \infty$? Рассмотрим $Y = \int_0^1 B(s)ds$, где B – Броуновское движение. Тогда Y является $\sigma\{B(s): 0 \leq s \leq 1\}$ измеримой, но нет бесконечномерного аналога (4).

Резюмируя, можно сказать, что измеримость тесно связана с функциональной зависимостью. В прикладной математике часто сталкиваются со случайными величинами, зависящими от некоторого числа базовых случайных величин. Это может быть выражено либо формулой (4), либо на языке измеримости относительно σ - алгебр, связанных с базовыми величинами, заданными в виде стохастических процессов.