

Лекция 4

4. Случайные величины и функции распределения

4.1. Определения

Пусть (Ω, \mathcal{F}, P) — произвольное вероятностное пространство. Числовая функция $\xi(\omega)$ на множестве элементарных событий Ω называется *случайной величиной*, если для всякого $x \in \mathbb{R}$

$$\{\xi < x\} = \{\omega \in \Omega : \xi(\omega) < x\} \in \mathcal{F}. \quad (1.1)$$

Из (1.1) следует, что

$$\begin{aligned} \{\xi \geq x\} &= \overline{\{\xi < x\}} \in \mathcal{F}, \\ \{x_1 \leq \xi < x_2\} &= \{\xi < x_2\} \setminus \{\xi < x_1\} \in \mathcal{F}, \\ \{\xi = x\} &= \bigcap_{n=1}^{\infty} \left\{ x \leq \xi < x + \frac{1}{n} \right\}. \end{aligned} \quad (1.2)$$

Функция

$$F(x) = F_{\xi}(x) = P\{\xi < x\}, \quad (1.3)$$

определённая при всех $x \in \mathbb{R}$, называется *функцией распределения*.

Очевидно, что функция распределения $F(x)$ удовлетворяет неравенству

$$0 \leq F(x) \leq 1 \quad (1.4)$$

при всяком $x \in \mathbb{R}$

Введём обозначения

$$F(x-0) = \lim_{y \nearrow x} F(y), \quad F(x+0) = \lim_{y \searrow x} F(y), \quad F(\pm\infty) = \lim_{y \rightarrow \pm\infty} F(y).$$

Справедливы равенства:

- i) $P\{x_1 \leq \xi < x_2\} = F_{\xi}(x_2) - F_{\xi}(x_1),$
- ii) $P\{\xi = x\} = F_{\xi}(x+0) - F_{\xi}(x),$
- iii) $P\{x_1 \leq \xi \leq x_2\} = F_{\xi}(x_2+0) - F_{\xi}(x_1),$
- iv) $P\{x_1 < \xi < x_2\} = F_{\xi}(x_2) - F_{\xi}(x_1+0),$
- v) $P\{x_1 < \xi \leq x_2\} = F_{\xi}(x_2+0) - F_{\xi}(x_1+0).$

Доказательство.

i) Так как

$$\{\xi < x_2\} = \{\xi < x_1\} \cup \{x_1 \leq \xi < x_2\},$$

то из аддитивности P получаем

$$P\{\xi < x_2\} = P\{\xi < x_1\} + P\{x_1 \leq \xi < x_2\}.$$

Отсюда следует (i).

ii) Из непрерывности P , (1.2), (i) получим

$$P\{\xi = x\} = \lim_{n \rightarrow \infty} \left(F_{\xi}\left(x + \frac{1}{n}\right) - F_{\xi}(x) \right) = F_{\xi}(x+0) - F_{\xi}(x).$$

Равенства (iii), (iv), (v) доказываются аналогично. ■

Пример 1. Рассмотрим схему Бернулли, состоящую из n испытаний с вероятностью успеха p . Обозначим через μ число успехов. Случайная величина μ принимает все целочисленные значения от 0 до n включительно. Согласно предыдущей главе

$$P(\mu = m) = P_n(m) = C_n^m p^m q^{n-m}, \quad m = 0, \dots, n.$$

Функция распределения случайной величины μ равна:

$$F(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ \sum_{k \leq x} P_n(k) & \text{при } 0 < x \leq n, \\ 1 & \text{при } x > n. \end{cases}$$

Функция распределения представляет собой ступенчатую функцию со скачками в точках $x = 0, \dots, n$; скачок в точке $x = k$ равен $P_n(k)$.

Каждая случайная величина однозначно определяет свою функцию распределения. Обратное неверно, т.е. одной функцию распределения могут соответствовать сколь угодно различных случайных величин.

Пример 2. Пусть случайная величина ξ принимает два значения -1 и $+1$, каждое с вероятностью $1/2$. Случайная величина $\nu = -\xi$ всегда отлична от ξ . При этом обе эти случайные величины имеют одну и ту же функцию распределения

$$F(x) = \begin{cases} 0 & \text{при } x \leq -1, \\ 1/2 & \text{при } -1 < x \leq 1, \\ 1 & \text{при } x > 1. \end{cases}$$

4.2. Свойства функции распределения

Пусть ξ — случайная величина. Функция распределения $F(x) = F_\xi(x)$ обладает следующими свойствами:

F1. $F(x)$ не убывает.

F2. $F(x)$ непрерывна слева.

F3. $F(-\infty) = 0$, $F(+\infty) = 1$.

Доказательство.

1. Пусть $x_1 < x_2$. Тогда $\{\xi < x_1\} \subset \{\xi < x_2\}$ и, следовательно, $F(x_1) \leq F(x_2)$.

2. Пусть числовая последовательность $\{y_n\}$ возрастает и $\lim_{n \rightarrow \infty} y_n = x_0$. Тогда

$$\{\xi < y_n\} \subset \{\xi < y_{n+1}\}, \quad \bigcup_{n=1}^{\infty} \{\xi < y_n\} = \{\xi < x_0\}.$$

Из непрерывности P и монотонности функции распределения получаем

$$F(x_0 - 0) = \lim_{n \rightarrow \infty} P\{\xi < y_n\} = P\{\xi < x_0\} = F(x_0).$$

3. В силу п. 1, F монотонна и поэтому существуют пределы $F(\pm\infty) = \lim_{x \rightarrow \pm\infty} F(x)$. Пусть $A_k = \{k-1 \leq \xi < k\}$. Ясно, что $\Omega = \sum_{k=-\infty}^{\infty} A_k$ и $P(A_k) = F(k) - F(k-1)$. Получаем

$$\begin{aligned} 1 = P(\Omega) &= \sum_{k=-\infty}^{\infty} P(A_k) = \lim_{N \rightarrow \infty} \sum_{k=-N+1}^N P(A_k) = \\ &= \lim_{N \rightarrow \infty} (F(N) - F(-N)) = F(+\infty) - F(-\infty). \end{aligned}$$

Из неравенства (1.4) следует, что $0 \leq F(\pm\infty) \leq 1$. Следовательно, $F(-\infty) = 0$ и $F(+\infty) = 1$. ■

Теорема 2.1. Пусть Функция $F(x)$ обладает свойствами F1, F2 и F3. Тогда существует вероятностное пространство (Ω, \mathcal{F}, P) и случайная величина ξ на этом пространстве такая, что $F_{\xi}(x) = F(x)$.

4.3. Дискретные и абсолютно непрерывные распределения

Распределение случайной величины ξ называется *дискретным*, если существует конечное или счётное множество чисел x_1, x_2, \dots таких, что

$$P\{\xi = x_n\} = p_n, \quad n = 1, 2, \dots; \quad \sum_{n=1}^{\infty} p_n = 1. \quad (3.1)$$

Распределение случайной величины ξ называется *абсолютно непрерывным*, если существует неотрицательная функция $p_{\xi}(x)$ такая, что для всякого $x \in \mathbb{R}$

$$F_{\xi}(x) = P\{\xi < x\} = \int_{-\infty}^x p_{\xi}(u) du. \quad (3.2)$$

Функция $p_{\xi}(x)$ называется *плотностью распределения вероятностей*.

Очевидно

$$\begin{aligned} P\{a \leq \xi < b\} &= \int_a^b p_{\xi}(x) dx, \\ P\{\xi = a\} &= \lim_{n \rightarrow \infty} \int_a^{a+\frac{1}{n}} p_{\xi}(x) dx = 0. \end{aligned}$$

Отсюда следует, что для абсолютно непрерывных величин

$$P\{a \leq \xi \leq b\} = P\{a < \xi \leq b\} = P\{a \leq \xi < b\} = P\{a < \xi < b\}.$$

Приведём часто встречающиеся распределения. Сначала перечислим дискретные распределения.

1. *Вырожденное распределение:*

$$P\{\xi = a\} = 1, \quad a — \text{постоянная}.$$

2. *Гипергеометрическое распределение* (N, M, n — натуральные числа, $M \leq N$, $n \leq N$):

$$P\{\xi = m\} = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}, \quad m = 0, 1, \dots, \min\{M, n\}.$$

3. *Биномиальное распределение* (n — натуральное число, $0 < p < 1$):

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

4. *Распределение Пуассона* ($\lambda > 0$):

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

5. *Геометрическое распределение* ($0 < p < 1$):

$$P\{\xi = k\} = (1-p)^{k-1} p, \quad k = 1, 2, \dots$$

Теперь перечислим некоторые абсолютно непрерывные распределения, указав их плотность.

1. *Равномерное распределение на отрезке* $[a, b]$, $a < b$:

$$p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

2. *Нормальное распределение с параметрами* (a, σ^2) ($\sigma > 0$, $a \in \mathbb{R}$):

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Нормальное распределение также называется *гауссовым*. Нормальное распределение с параметрами $(0, 1)$ называется *стандартным нормальным распределением*.

3. *Показательное распределение с параметром* $\lambda > 0$:

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

4.4. Математическое ожидание

Пусть задано вероятностное пространство (Ω, \mathcal{F}, P) . *Математическим ожиданием* случайной величины ξ , заданной на (Ω, \mathcal{F}, P) , называется величина

$$M\xi = \int_{\Omega} \xi(\omega) P(d\omega), \quad (4.1)$$

если интеграл Лебега, стоящий в правой части равенства, существует.

Говорят, что некоторое свойство выполнено *P-почти наверное*, если существует множество $N \in \mathcal{F}$ с $P(N) = 0$ такое, что это свойство выполнено для всех $\omega \in \Omega \setminus N$. Вместо слов *P-почти наверное* часто говорят *P-почти всюду* или просто *почти наверное* (п.н.), *почти всюду* (п.в.).

Следующие свойства математического ожидания являются следствиями теории интеграла Лебега.

M1. Если c — постоянная, то $Mc = c$.

M2. Если c — постоянная, то $M(c\xi) = cM\xi$.

М3. Для любой случайной величины ξ

$$|M\xi| \leq M|\xi|.$$

М4. Для любых случайных величин ξ_1 и ξ_2

$$M(\xi_1 + \xi_2) = M\xi_1 + M\xi_2.$$

М5. Если случайные величины ξ_1 и ξ_2 независимы, то

$$M\xi_1\xi_2 = M\xi_1 \cdot M\xi_2.$$

М6. Если случайная величина ξ неотрицательна, то $M\xi \geq 0$.

М7. Если случайная величина ξ неотрицательна и $M\xi = 0$, то $\xi = 0$ почти наверное.

Неравенство Коши-Буняковского. Для любых двух случайных величин ξ , η таких, что $M\xi^2 < \infty$, $M\eta^2 < \infty$, справедливо неравенство

$$|M\xi\eta| \leq \sqrt{M\xi^2 \cdot M\eta^2} \quad (4.2)$$

Доказательство. Если $M\xi^2 = 0$ или $M\eta^2 = 0$, то из М7 следует, что либо $\xi = 0$ п.н. либо $\eta = 0$ п.н. и неравенство (4.2) доказано. Поэтому будем предполагать, что $M\xi^2 > 0$, $M\eta^2 > 0$.

Пусть x — константа. Рассмотрим случайную величину $(\xi + x\eta)^2 \geq 0$. Из М2, М6 получаем

$$0 \leq M(\xi + x\eta)^2 = M\xi^2 + 2xM\xi\eta + x^2M\eta^2.$$

Так как это неравенство справедливо для любого x , то дискриминант \mathcal{D} стоящего справа квадратичного многочлена неположителен. Таким образом,

$$\frac{\mathcal{D}}{4} = (M\xi\eta)^2 - M\xi^2 M\eta^2 \leq 0.$$

Неравенство (4.2) доказано. ■

Переформулируем определение математического ожидания в наиболее важных случаях.

Математическим ожиданием $M\xi$ случайной величины ξ , заданной на дискретном вероятностном пространстве (Ω, \mathcal{F}, P) , $\Omega = \{\omega_1, \omega_2, \dots\}$, называется число

$$M\xi = \sum_{k=1}^{\infty} \xi(\omega_k) P(\omega_k), \quad (4.3)$$

если ряд (4.3) абсолютно сходится. Если ряд (4.3) не сходится абсолютно, то говорят, что математическое ожидание случайной величины ξ не существует.

Математическим ожиданием $M\xi$ случайной величины ξ , заданной на абсолютно непрерывном вероятностном пространстве (Ω, \mathcal{F}, P) , называется число

$$M\xi = \int_{\Omega} \xi(x_1, \dots, x_n) \pi(x_1, \dots, x_n) dx_1 \dots dx_n, \quad (4.4)$$

если интеграл (4.4) абсолютно сходится. Если интеграл (4.4) не сходится абсолютно, то говорят, что математическое ожидание случайной величины ξ не существует.

Теорема 4.1. Пусть случайная величина ξ имеет дискретное распределение. Если

$$\sum_{n=1}^{\infty} |g(x_n)| P\{\xi = x_n\} < \infty,$$

тогда случайная величина $g(\xi)$ имеет математическое ожидание

$$Mg(\xi) = \sum_{n=1}^{\infty} g(x_n) P\{\xi = x_n\} \quad (4.5)$$

В частности

$$M\xi = \sum_{n=1}^{\infty} x_n P\{\xi = x_n\} \quad (4.6)$$

Теорема 4.2. Пусть случайная величина ξ имеет абсолютно непрерывное распределение с плотностью $p(x)$. Если

$$\int_{-\infty}^{+\infty} |g(x)| p(x) dx < \infty,$$

тогда случайная величина $g(\xi)$ имеет математическое ожидание

$$Mg(\xi) = \int_{-\infty}^{+\infty} g(x) p(x) dx. \quad (4.7)$$

В частности

$$M\xi = \int_{-\infty}^{+\infty} xp(x) dx. \quad (4.8)$$

4.5. Дисперсия

Дисперсией случайной величины ξ называется величина

$$D\xi = M(\xi - M\xi)^2, \quad (5.1)$$

если математическое ожидание, стоящее в правой части равенства, существует. Дисперсия является мерой рассеяния значений случайной величины около её математического ожидания. Величина $\sigma = \sqrt{D\xi}$ называется *средним квадратическим отклонением*.

Преобразуем правую часть (5.1). Имеем

$$M(\xi - M\xi)^2 = M(\xi^2 - 2\xi M\xi + (M\xi)^2) = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2.$$

Подставляя в (5.1), получаем

$$D\xi = M\xi^2 - (M\xi)^2. \quad (5.2)$$

Если случайная величина ξ имеет дискретное распределение, то из (4.5) при $g(x) = (x - M\xi)^2$ получаем

$$D\xi = \sum_{n=1}^{\infty} (x_n - M\xi)^2 P\{\xi = x_n\} \quad (5.3)$$

Аналогично для случайной величины ξ с абсолютно непрерывным распределением из (4.7) будем иметь

$$D\xi = \int_{-\infty}^{+\infty} (x - M\xi)^2 p(x) dx. \quad (5.4)$$

Теорема 5.1.

1. Для любой случайной величины ξ справедливо неравенство $D\xi \geq 0$.
2. Если c — постоянная, то $Dc = 0$.
3. Если c — постоянная, то

$$D(c\xi) = c^2 D\xi, \quad D(\xi + c) = D\xi$$

4. Если случайные величины ξ_1 и ξ_2 независимы, то

$$D(\xi_1 + \xi_2) = D\xi_1 + D\xi_2.$$

Доказательство. Свойства 1–3 следуют из определения и свойств математического ожидания. Докажем свойство 4. По определению (5.1)

$$\begin{aligned} D(\xi_1 + \xi_2) &= M[(\xi_1 + \xi_2) - M(\xi_1 + \xi_2)]^2 = M[(\xi_1 - M\xi_1) + (\xi_2 - M\xi_2)]^2 = \\ &= M[(\xi_1 - M\xi_1)]^2 + M[(\xi_2 - M\xi_2)]^2 + 2M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]. \end{aligned}$$

Так как случайные величины ξ_1 и ξ_2 независимы, то

$$M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)] = 0.$$

Поэтому

$$D(\xi_1 + \xi_2) = M[(\xi_1 - M\xi_1)]^2 + M[(\xi_2 - M\xi_2)]^2 = D\xi_1 + D\xi_2. \quad \blacksquare$$

4.6. Примеры

Вычислим математические ожидания и дисперсии некоторых распределений.

1. *Вырожденное распределение.* В этом случае очевидно, что

$$M\xi = a, \quad D\xi = 0. \quad (6.1)$$

2. *Биномиальное распределение* ($n \in \mathbb{N}$, $0 < p < 1$). Имеем

$$M\xi = \sum_{k=0}^n k P\{\xi = k\} = \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k} = \sum_{k=1}^n k C_n^k p^k (1-p)^{n-k}.$$

Так как $k C_n^k = n C_{n-1}^{k-1}$, то

$$M\xi = np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} = np(p + (1-p))^{n-1} = np.$$

Найдём $M[\xi(\xi - 1)]$. Запишем

$$M[\xi(\xi - 1)] = \sum_{k=0}^n k(k-1) C_n^k p^k (1-p)^{n-k} = \sum_{k=2}^n k(k-1) C_n^k p^k (1-p)^{n-k}.$$

Учитывая равенство $k(k-1)C_n^k = n(n-1)C_{n-2}^{k-2}$, получим

$$\begin{aligned} M[\xi(\xi-1)] &= n(n-1)p^2 \sum_{k=2}^n C_{n-2}^{k-2} p^{k-2} (1-p)^{n-k} = \\ &= n(n-1)p^2(p+(1-p))^{n-2} = n(n-1)p^2. \end{aligned}$$

$$D\xi = M\xi^2 - (M\xi)^2 = M[\xi(\xi-1)] + M\xi - (M\xi)^2.$$

Следовательно

$$D\xi = n(n-1)p^2 + np - (np)^2 = np(1-p).$$

В результате

$$M\xi = np, \quad D\xi = np(1-p) = npq. \quad (6.2)$$

3. *Распределение Пуассона* ($\lambda > 0$).

Найдём $M\xi$.

$$M\xi = \sum_{k=0}^{\infty} k P\{\xi = k\} = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

Вычислим $M[\xi(\xi-1)]$. Имеем

$$M[\xi(\xi-1)] = \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2.$$

Далее

$$D\xi = M\xi^2 - (M\xi)^2 = M[\xi(\xi-1)] + M\xi - (M\xi)^2 = \lambda.$$

В результате

$$M\xi = \lambda, \quad D\xi = \lambda. \quad (6.3)$$

4. *Равномерное распределение на отрезке* $[a, b]$, $a < b$:

$$M\xi = \int_a^b xp(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \frac{1}{2} (b^2 - a^2) = \frac{a+b}{2}.$$

Далее

$$M\xi^2 = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{b-a} \frac{1}{3} (b^3 - a^3) = \frac{a^2 + ab + b^2}{3}.$$

Найдём дисперсию.

$$D\xi = M\xi^2 - (M\xi)^2 = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

В результате

$$M\xi = \frac{a+b}{2}, \quad D\xi = \frac{(b-a)^2}{12}. \quad (6.4)$$

5. *Нормальное распределение с параметрами* (a, σ^2) ($\sigma > 0$, $a \in \mathbb{R}$).

$$M\xi = \int_{-\infty}^{+\infty} xp(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} xe^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

Сделаем замену $x = a + t\sigma$. Тогда

$$M\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (a + t\sigma) e^{-t^2/2} dt.$$

Учитывая, что

$$\int_{-\infty}^{+\infty} t e^{-t^2/2} dt = 0, \quad \int_{-\infty}^{+\infty} e^{-t^2/2} dt = \sqrt{2\pi}, \quad (6.5)$$

получим $M\xi = a$.

$$D\xi = \int_{-\infty}^{+\infty} (x - M\xi)^2 p(x) dx = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} (x - a)^2 e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

После замены переменной $x = a + t\sigma$

$$D\xi = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt.$$

Значение интеграла сведём к (6.5) интегрированием по частям:

$$\sqrt{2\pi} = \int_{-\infty}^{+\infty} e^{-t^2/2} dt = te^{-t^2/2} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt = \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt.$$

В результате

$$M\xi = a, \quad D\xi = \sigma^2. \quad (6.6)$$