Лекция 6

6. Случайные векторы

6.1. Законы распределения случайных векторов

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ заданы n случайных величин ξ_1, …, ξ_n. Данные случайные величины можно рассматривать как n-мерный вектор:

$$\xi : \Omega \to \mathbb{R}^n, \quad \xi(\omega) = (\xi_1(\omega), \ldots, \xi_n(\omega)). \tag{1.1}$$

Величина ξ называется *случайным вектором.*

Функцией распределения случайного вектора (1.1) или *совместной функцией распределения* случайных величин ξ_1, …, ξ_n называется функция n переменных

$$F(x_1, \ldots, x_n) = \mathsf{P}\{\, \xi_1 < x_1, \ldots, \xi_n < x_n \,\}. \tag{1.2}$$

Величину (1.2) ещё называют *многомерной функцией распределения.*

Случайные величины ξ_1, …, ξ_n называются *независимыми* (*в совокупности*), если

$$F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = F_1(x_1) \ldots F_n(x_n), \tag{1.3}$$

где $F_i(x_i) = \mathsf{P}(\xi_i < x_i)$ — одномерная функция распределения ξ_i.

Далее мы будем рассматривать только двумерный случай. Для двумерного вектора $\zeta = (\xi, \eta)$, имеем

$$F_\zeta(x, y) = F(x, y) = \mathsf{P}(\xi < x, \eta < y). \tag{1.4}$$

Пусть $Q = [a, b) \times [c, d)$ — полуоткрытый прямоугольник. Очевидно, что $|Q| > 0$ тогда и только тогда, когда $a < b$ и $c < d$.

Предполагая, что $|Q| > 0$, легко найти вероятность

$$\mathsf{P}\{\zeta \in Q\} = \mathsf{P}\{a \le \xi < b,\ c \le \eta < d\}.$$

Обозначим

$$\begin{aligned} Q &= [a, b) \times [c, d), & Q_1 &= [a, b) \times (-\infty, c), \\ Q_2 &= (-\infty, a) \times [c, d), & Q_3 &= (-\infty, a) \times (-\infty, c). \end{aligned}$$

Очевидно, что $(-\infty, b) \times (-\infty, d) = Q \sqcup Q_1 \sqcup Q_2 \sqcup Q_3$. Следовательно,

$$\mathsf{P}\{\zeta \in Q\} + \mathsf{P}\{\zeta \in Q_1\} + \mathsf{P}\{\zeta \in Q_2\} + \mathsf{P}\{\zeta \in Q_3\} = F_\zeta(b, d).$$

Далее, так как

$$Q_1 \sqcup Q_3 = (-\infty, b) \times (-\infty, c), \quad Q_2 \sqcup Q_3 = (-\infty, a) \times (-\infty, d),$$

то

$$\mathsf{P}\{\zeta \in Q_1\} + \mathsf{P}\{\zeta \in Q_3\} = F_\zeta(b, c), \quad \mathsf{P}\{\zeta \in Q_2\} + \mathsf{P}\{\zeta \in Q_3\} = F_\zeta(a, d).$$

В результате получаем

$$\mathsf{P}\{a \le \xi < b,\ c \le \eta < d\} = F_\zeta(b, d) - F_\zeta(b, c) - F_\zeta(a, d) + F_\zeta(a, c). \tag{1.5}$$

Функция распределения $F_{\xi\eta}(x, y)$ обладает следующими свойствами:

R1. $F_{\xi\eta}(x, y)$ есть неубывающая функция по каждому аргументу x и y.

R2. $F_{\xi\eta}(x, y)$ непрерывна слева по каждому аргументу x и y.

R3. $F_{\xi\eta}(x, y)$ удовлетворяет соотношениям

$$F(+\infty, +\infty) = 1, \quad F(-\infty, y) = 0, \quad F(x, -\infty) = 0,$$

при произвольных значениях x и y.

В одномерном случае перечисленные свойства необходимы и достаточны, чтобы функция $F(x)$ была функцией распределения некоторой случайной величины. В многомерном случае этих свойств уже недостаточно. Для того, чтобы функция $F(x, y)$ была функцией распределения, к перечисленным свойствам R1 – R3, нужно добавить следующее:

R4. Для любых чисел $a < b$, $c < d$ справедливо неравенство:

$$F_\zeta(b, d) - F_\zeta(b, c) - F_\zeta(a, d) + F_\zeta(a, c) \geq 0.$$

Пример. Пусть

$$F(x, y) = \begin{cases} 0, & \text{при } x \leq 0, \text{или } y \leq 0, \text{или } x + y \leq 1, \\ 1, & \text{в остальных точках плоскости.} \end{cases}$$

Функция $F(x, y)$ удовлетворяет свойствам R1 – R3. При этом

$$F(1, 1) - F\left(1, \frac{1}{2}\right) - F\left(\frac{1}{2}, 1\right) + F\left(\frac{1}{2}, \frac{1}{2}\right) = -1.$$

Следовательно, R4 не выполнено.

Рассмотрим случайные величины ξ, η с функциями распределения $F_\xi(x)$, $F_\eta(y)$. Пусть $F_{\xi\eta}(x, y)$ совместная функция распределения ξ, η. Тогда справедливы равенства:

$$F_{\xi\eta}(x, +\infty) = F_\xi(x), \quad F_{\xi\eta}(+\infty, y) = F_\eta(y). \tag{1.6}$$

Равенства (1.6) обычно называют *условиями согласованности.*

6.2. Случайные векторы дискретного типа

Двумерный случайный вектор (ξ, η) называется *случайным вектором дискретного типа* (сокращённо СВДТ), если множество его значений не более, чем счётно. *Законом распределения* дискретной двумерной случайной величины (ξ, η) называется перечень возможных значений этой величины, т. е. пар (x_i, y_j) и их вероятностей

$$p_{ij} = p(x_i, y_j) = \mathsf{P}\{\, \xi = x_i, \eta = y_j \,\}. \tag{2.1}$$

Очевидно, что величины $p_{i,j}$ удовлетворяют условию

$$\sum_{i,j} p_{ij} = 1. \tag{2.2}$$

Пусть $\varphi(x, y)$ некоторая функция двух переменных. Тогда $\varphi(\xi, \eta)$ является случайной величиной и справедлива формула

$$\mathsf{M}\varphi(\xi, \eta) = \sum_{i,j} \varphi(x_i, y_j)\, p(x_i, y_j). \tag{2.3}$$

Если множество значений СВДТ конечно, то закон распределения этой величины удобно представлять в виде следующей таблицы.

$\xi \setminus \eta$	y_1	$\dots$	y_m
x_1	p_{11}	$\dots$	p_{1m}
$\dots$	$\dots$	$\dots$	$\dots$
x_n	p_{n1}	$\dots$	p_{nm}

Одномерные законы распределения отдельных компонент СВДТ выражаются через вероятности совместных значений $p_{i,j}$ по формулам

$$p_{i\cdot} = \mathsf{P}\{\,\xi = x_i\} = \sum_j p_{ij}, \quad p_{\cdot j} = \mathsf{P}\{\,\eta = y_j\} = \sum_i p_{ij}. \tag{2.4}$$

Дискретные случайные величины ξ, η независимы тогда и только тогда, когда

$$p_{ij} = \mathsf{P}\{\,\xi = x_i, \eta = y_j\,\} = \mathsf{P}\{\,\xi = x_i\}\,\mathsf{P}\{\,\eta = y_j\,\} = p_{i\cdot}p_{\cdot j} \tag{2.5}$$

для любых x_i и y_j.

Пример. Пусть дан случайный вектор (ξ, η) дискретного типа с законом распределения

$\xi \setminus \eta$	0	1	2
0	1/4	1/3	1/9
1	0	1/6	1/9
2	0	0	1/36

Найдём одномерные законы распределения по формулам (2.4).

$$\begin{aligned} \mathsf{P}\{\,\xi = 0\,\} &= 1/4 + 1/3 + 1/9 = 25/36, \\ \mathsf{P}\{\,\xi = 1\,\} &= 0 + 1/6 + 1/9 = 10/36, \\ \mathsf{P}\{\,\xi = 2\,\} &= 0 + 0 + 1/36 = 1/36. \end{aligned}$$

Далее

$$\begin{aligned} \mathsf{P}\{\,\eta = 0\,\} &= 1/4 + 0 + 0 = 1/4, \\ \mathsf{P}\{\,\eta = 1\,\} &= 1/3 + 1/6 + 0 = 1/2, \\ \mathsf{P}\{\,\eta = 2\,\} &= 1/9 + 1/9 + 1/36 = 1/4. \end{aligned}$$

Так как

$$\mathsf{P}\{\,\xi = 1, \eta = 1\,\} = 1/6 \neq 5/36 = \mathsf{P}\{\,\xi = 1\,\} \cdot \mathsf{P}\{\,\eta = 1\,\},$$

то случайные величины зависимы. Найдём математические ожидания и дисперсии.

$$\begin{aligned} \mathsf{M}\xi &= 0 \cdot 25/36 + 1 \cdot 10/36 + 2 \cdot 1/36 = 1/3, \\ \mathsf{M}\eta &= 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/4 = 1, \\ \mathsf{M}\xi^2 &= 0 \cdot 25/36 + 1 \cdot 10/36 + 4 \cdot 1/36 = 7/18, \\ \mathsf{M}\eta^2 &= 0 \cdot 1/4 + 1 \cdot 1/2 + 4 \cdot 1/4 = 3/2. \end{aligned}$$

Таким образом,

$$\begin{aligned} \mathsf{D}\xi &= \mathsf{M}\xi^2 - (\mathsf{M}\xi)^2 = 7/18 - 1/9 = 5/18, \\ \mathsf{D}\eta &= \mathsf{M}\eta^2 - (\mathsf{M}\eta)^2 = 3/2 - 1 = 1/2. \end{aligned}$$

6.3. Случайные векторы непрерывного типа

Двумерный случайный вектор (ξ, η) называется *случайным вектором непрерывного типа* (сокращённо СВНТ), если функция распределения $F_{\xi,\eta}(x, y)$ непрерывна на всей плоскости и существует такая неотрицательная интегрируемая функция $p_{\xi,\eta}(x, y)$, называемая *плотностью распределения вероятностей* случайного вектора (ξ, η), что

$$F_{\xi,\eta}(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} p_{\xi,\eta}(s, t)\, ds\, dt. \tag{3.1}$$

Очевидно, что плотность распределения вероятностей удовлетворяет равенству:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{\xi,\eta}(x, y)\, dx\, dy = 1. \tag{3.2}$$

Равенство (3.2) называется *условием нормировки.* Плотность распределения вероятностей отдельных компонент СВНТ выражаются в виде интегралов от совместной плотности:

$$p_\xi(x) = \int_{-\infty}^{\infty} p_{\xi,\eta}(x, y)\, dy, \quad p_\eta(y) = \int_{-\infty}^{\infty} p_{\xi,\eta}(x, y)\, dx. \tag{3.3}$$

Если (ξ, η) — СВНТ, то вероятность попадания случайной точки в произвольную квадрируемую область $G \subset \mathbb{R}^2$ определяется по формуле

$$\mathsf{P}\{\, (\xi, \eta) \in G \,\} = \iint_G p_{\xi,\eta}(x, y)\, dx\, dy. \tag{3.4}$$

Пусть $\varphi(x, y)$ непрерывная функция двух переменных. Тогда $\varphi(\xi, \eta)$ является случайной величиной и справедлива формула

$$\mathsf{M}\varphi(\xi, \eta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi(x, y)\, p_{\xi,\eta}(x, y)\, dx\, dy, \tag{3.5}$$

если интеграл в правой части (3.5) существует.

Непрерывные случайные величины ξ, η независимы тогда и только тогда, когда

$$p_{\xi,\eta}(x, y) = p_\xi(x) \cdot p_\eta(y) \tag{3.6}$$

для любых x и y.

П р и м е р. Пусть дан случайный вектор (ξ, η) непрерывного типа с законом распределения

$$p_{\xi,\eta}(x, y) = \begin{cases} \dfrac{1}{6\pi}, & \text{если } x^2/9 + y^2/4 \le 1, \\ \\ 0, & \text{если иначе.} \end{cases}$$

Найдём одномерные законы распределения по формулам (3.3).

$$p_\xi(x) = \frac{1}{6\pi} \int_{-2\sqrt{1-x^2/9}}^{2\sqrt{1-x^2/9}} dy = \frac{4}{6\pi}\sqrt{1 - x^2/9} = \frac{2}{9\pi}\sqrt{9 - x^2}.$$

Следовательно,

$$p_\xi(x) = \begin{cases} \dfrac{2}{9\pi}\sqrt{9-x^2}, & \text{если } |x| \le 3, \\ \\ 0, & \text{если } |x| > 3. \end{cases}$$

Далее

$$p_\eta(y) = \frac{1}{6\pi}\int\limits_{-3\sqrt{1-y^2/4}}^{3\sqrt{1-y^2/4}} dx = \frac{6}{6\pi}\sqrt{1-y^2/4} = \frac{1}{2\pi}\sqrt{4-y^2}.$$

Поэтому

$$p_\eta(y) = \begin{cases} \dfrac{1}{2\pi}\sqrt{4-y^2}, & \text{если } |y| \le 2, \\ \\ 0, & \text{если } |y| > 2. \end{cases}$$

Так как равенство (3.6), очевидно, не выполнено, то случайные величины не независимы.

Найдём математические ожидания и дисперсии. Так как $p_\xi(x)$, $p_\eta(y)$ чётные функции своих аргументов, то

$$\mathsf{M}\xi = \int\limits_{-\infty}^{\infty} x p_\xi(x)\,dx = 0, \quad \mathsf{M}\eta = \int\limits_{-\infty}^{\infty} y p_\eta(y)\,dy = 0.$$

Следовательно,

$$\mathsf{D}\xi = \mathsf{M}\xi^2 = \int\limits_{-\infty}^{\infty} x^2 p_\xi(x)\,dx = \frac{2}{9\pi}\int\limits_{-3}^{3} x^2\sqrt{9-x^2}\,dx = \frac{18}{\pi}\int\limits_{-\pi/2}^{\pi/2}\sin^2 t\cos^2 t\,dt =$$

$$= \frac{9}{2\pi}\int\limits_{-\pi/2}^{\pi/2}\sin^2 2t\,dt = \frac{9}{4\pi}\int\limits_{-\pi/2}^{\pi/2}(1-\cos 4t)\,dt = \frac{9}{4}.$$

Здесь мы сделали замену $x = 3\sin t$. Аналогично для $\mathsf{D}\eta$:

$$\mathsf{D}\eta = \mathsf{M}\eta^2 = \int\limits_{-\infty}^{\infty} y^2 p_\eta(y)\,dy = \frac{1}{2\pi}\int\limits_{-2}^{2} y^2\sqrt{4-y^2}\,dy = \frac{8}{\pi}\int\limits_{-\pi/2}^{\pi/2}\sin^2 t\cos^2 t\,dt = 1.$$

6.4. Ковариация. Коэффициент корреляции

Ковариацией случайных величин ξ_1, ξ_2 называется величина

$$\mathsf{cov}(\xi_1,\xi_2) = \mathsf{M}\left[(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2)\right]. \tag{4.1}$$

Так как

$$(\xi_1 - \mathsf{M}\xi_1)(\xi_2 - \mathsf{M}\xi_2) = \xi_1\xi_2 - \xi_1\mathsf{M}\xi_2 - \xi_2\mathsf{M}\xi_1 + \mathsf{M}\xi_1\mathsf{M}\xi_1,$$

то ковариацию можно записать в виде

$$\operatorname{cov}(\xi_1,\xi_2) = \mathsf{M}\xi_1\xi_2 - \mathsf{M}\xi_1\mathsf{M}\xi_2. \tag{4.2}$$

Теорема 4.1. *Если для случайных величин* $\xi_1, \ldots, \xi_n$ *существуют* $\operatorname{cov}(\xi_i, \xi_j) = \sigma_{ij}$, $i, j = 1, \ldots, n$, *то при любых постоянных* $c_1, \ldots, c_n$ *справедливо равенство*

$$\mathsf{D}\Big(\sum_{i=1}^{n} c_i \xi_i\Big) = \sum_{i,j=1}^{n} \sigma_{ij} c_i c_j \tag{4.3}$$

Д о к а з а т е л ь с т в о. Положим

$$\eta = \sum_{i=1}^{n} c_i \xi_i.$$

Нетрудно проверить, что

$$\eta - \mathsf{M}\eta = \sum_{i=1}^{n} c_i (\xi_i - \mathsf{M}\xi_i)$$

и

$$(\eta - \mathsf{M}\eta)^2 = \sum_{i,j=1}^{n} c_i c_j (\xi_i - \mathsf{M}\xi_i)(\xi_j - \mathsf{M}\xi_j).$$

Вычисляя математическое ожидание от обеих частей последнего равенства, получим утверждение теоремы. ■

Полагая в (4.3) $c_i = 1$, $i = 1, \ldots, n$, получим

$$\mathsf{D}\Big(\sum_{i=1}^{n} \xi_i\Big) = \sum_{i=1}^{n} \mathsf{D}\xi_i + 2 \sum_{1 \le i < j \le n} \operatorname{cov}(\xi_i, \xi_j). \tag{4.4}$$

Из неравенства Коши-Буняковского (5.2.1) вытекает, что

$$|\,\mathsf{M}(\xi - \mathsf{M}\xi)(\eta - \mathsf{M}\eta)\,| \le \sqrt{\mathsf{M}(\xi - \mathsf{M}\xi)^2 \cdot \mathsf{M}(\eta - \mathsf{M}\eta)^2}.$$

Следовательно,

$$|\operatorname{cov}(\xi\eta)| \le \sqrt{\mathsf{D}\xi\,\mathsf{D}\eta} \tag{4.5}$$

Из (4.2) и свойства математического ожидания M5 следует, что если случайные величины ξ_1, ξ_2 независимы, то $\operatorname{cov}(\xi_1, \xi_2) = 0$. Таким образом, если $\operatorname{cov}(\xi_1, \xi_2) \ne 0$, то случайные величины ξ_1, ξ_2 зависимы.

В качестве количественной характеристики степени зависимости случайных величин ξ_1 и ξ_2 используется *коэффициент корреляции* $\rho(\xi_1, \xi_2)$, определяемый равенством:

$$\rho(\xi_1, \xi_2) = \frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{\mathsf{D}\xi_1\,\mathsf{D}\xi_2}}. \tag{4.6}$$

Теорема 4.2. *Если для случайных величин* ξ_1, ξ_2 *существуют конечные дисперсии, отличные о нуля, то*

COR1. $|\,\rho(\xi_1, \xi_2)\,| \le 1$;

COR2. *если* ξ_1, ξ_2 *независимы, то* $\rho(\xi_1, \xi_2) = 0$;

COR3. *равенство* $|\,\rho(\xi_1, \xi_2)\,| = 1$ *справедливо тогда и только тогда, когда* ξ_1 *и* ξ_2 *зависят друг от друга линейно.*

Случайные величины ξ_1 и ξ_2 называются *некоррелироваными*, если $\rho(\xi_1, \xi_2) = 0$. Если $\rho(\xi_1, \xi_2) \ne 0$, то ξ_1 и ξ_2 называются *коррелироваными.*

П р и м е р 1. Рассмотрим пример из п.2. Вычислим ковариацию случайных величин ξ и η, используя формулу (2.3). Имеем

$$\mathsf{M}\xi\eta = 1 \cdot 1 \cdot 1/6 + 1 \cdot 2 \cdot 1/9 + 2 \cdot 1 \cdot 0 + 2 \cdot 2 \cdot 1/36 = 1/2.$$

Следовательно,

$$\mathsf{cov}(\xi, \eta) = \mathsf{M}\xi\eta - \mathsf{M}\xi\, \mathsf{M}\eta = 1/2 - 1/3 = 1/6.$$

Коэффициент корреляции равен

$$\rho(\xi, \eta) = \frac{\mathsf{cov}(\xi, \eta)}{\sqrt{\mathsf{D}\xi}\,\sqrt{\mathsf{D}\eta}} = \frac{1}{\sqrt{5}}.$$

Таким образом, случайные величины ξ и η коррелированы.

П р и м е р 2. Рассмотрим пример из п.3. Вычислим ковариацию случайных величин ξ и η, используя формулу (3.5). Имеем

$$\mathsf{M}\xi\eta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, p_{\xi,\eta}(x, y)\, dx\, dy = \int_{-2}^{2} y\, dy \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} x\, dx = 0.$$

Следовательно,

$$\mathsf{cov}(\xi, \eta) = \mathsf{M}\xi\eta - \mathsf{M}\xi\, \mathsf{M}\eta = 0.$$

Коэффициент корреляции, очевидно, тоже равен нулю: $\rho(\xi, \eta) = 0$. Таким образом, случайные величины ξ и η некоррелированы.

Из примера 2 вытекает, что из некоррелированности случайных величин ξ и η не следует их независимость.