5. Дискретные случайные величины

Часто с результатом случайного эксперимента связывают число. В связи с этим возникает важное понятие случайной величины.

Пусть $(\Omega, \mathcal{A}, \mathbb{P})$ — некоторое вероятное пространство. Под *случайной величиной*, определённой на этом вероятностном пространстве, будем понимать числовую функцию $\xi : \mapsto \mathbb{R}$ такую, что

$$\{\omega \,:\, \xi(\omega) \,\le\, x\} \,\in\, \mathcal{A}$$

для всех $x \in \mathbb{R}$. Другими словами, $\{\xi \le x\}$ является событием и мы можем говорить о вероятности этого события. В этом параграфе мы будем изучать, в основном, дискретные случайные величины, которые принимают конечное или счётное множество значений. Если $\xi : \Omega \mapsto \{x_1, x_2, \ldots\}$, то условие того, что ξ — случайная величина, выражается просто: $\{\xi = x_k\} \in \mathcal{A}$, $k = 1, 2, \ldots$.

Вначале рассмотрим случай, когда ξ принимает конечное число значений. Такие случайные величины называют *простыми*. Если ξ принимает только значения $x_1, \ldots, x_n$, то события $D_k = \{\xi = x_k\}$ образуют разбиение $\mathbb{D}_\xi = \{D_1, \ldots, D_n\}$. Через $\mathcal{A}_\xi$ будем обозначать алгебру, порождённую этим разбиением. Заметим также, что $\mathbb{P}(\xi = x_k) = \mathbb{P}(D_k)$.

Простейшей случайной величиной является индикатор события $A \in \mathcal{A}$, который определяется равенством

$$\mathbb{1}_A(\omega) \,=\, \begin{cases} 1, & \text{если } \omega \,\in\, A, \\ 0, & \text{если } \omega \,\notin\, A. \end{cases}$$

Отметим основные свойства индикаторов, которые следуют непосредственно из определения.

$$\mathbb{1}_\Omega(\omega) \,\equiv\, 1, \quad \mathbb{1}_\varnothing(\omega) \,\equiv\, 0, \quad \mathbb{1}_{\overline{A}}(\omega) \,=\, 1 - \mathbb{1}_A(\omega).$$

В дальнейшем мы будем опускать зависимость от элементарного события и писать просто $\mathbb{1}_A$. Имеют место также следующие соотношения

$$\mathbb{1}_{\cap A_i} \,=\, \prod \mathbb{1}_{A_i}, \quad \mathbb{1}_{\cup A_i} \,=\, 1 - \mathbb{1}_{\overline{\cup A_i}} \,=\, 1 - \mathbb{1}_{\cap \overline{A_i}} \,=\, 1 - \prod \left(1 - \mathbb{1}_{A_i}\right).$$

Пусть ξ — случайная величина, принимающая значения $x_1, \ldots, x_n$ соответственно на $D_1, \ldots, D_n$, т. е. $\mathbb{D}_\xi = \{D_1, \ldots, D_n\}$. Тогда

$$\xi \,=\, \sum_{i=1}^{n} x_i \mathbb{1}_{D_i}. \tag{5.1}$$

Другими словами, простую случайную величину можно представить как линейную комбинацию индикаторов.

Допустим теперь, что проводится большое количество N экспериментов, соответствующих вероятностному пространству $(\Omega, \mathcal{A}, \mathbb{P})$. Тогда значение x_i будет приниматься случайной величиной ξ, определяемой равенством (5.1), приблизительно в Np_i, $p_i = \mathbb{P}(D_i)$, случайных экспериментах. Следовательно, среднее наблюдавшихся значений случайной величины ξ будет приблизительно равно величине

$$\frac{1}{N}\left(x_1 N p_1 + \ldots + x_n N p_n\right) = \sum_{i=1}^{n} x_i p_i.$$

Это делает интуитивно понятным следующее определение.

Определение 5.1 *Математическим ожиданием случайной величины ξ, определяемой равенством* (5.1), *называется число*

$$\mathbb{E}\xi = \sum_{i=1}^{n} x_i \mathbb{P}(D_i) = \sum_{i=1}^{n} x_i \mathbb{P}(\xi = x_i). \qquad (5.2)$$

Замечание 5.1 *Если случайная величина ξ, определяемая равенством* (5.1), *допускает также представление*

$$\xi = \sum_{j=1}^{N} y_j \mathbb{1}_{H_j},$$

где $\{H_1, \ldots, H_N\}$ — разбиение, а $y_1, \ldots, y_N$ не обязательно все все различные, то

$$\mathbb{E}\xi = \sum_{j=1}^{N} y_j \mathbb{P}(H_j).$$

Доказательство *Действительно, множество индексов $\{1, \ldots, N\}$ можно разбить на подмножества $J_k = \{j : y_j = x_k\}$, $k = 1, \ldots, n$. Заметим, что при этом $D_k = \cup_{j \in J_k} H_j$. Следовательно,*

$$\sum_{j=1}^{N} y_j \mathbb{P}(H_j) = \sum_{k=1}^{n} \sum_{j \in J_k} x_k \mathbb{P}(H_j) = \sum_{k=1}^{n} x_k \mathbb{P}(D_k) = \mathbb{E}\xi.$$

□

Замечание 5.2 *Если ξ определена равенством* (5.1), *а $\varphi : \mathbb{R} \mapsto \mathbb{R}$ — некоторая функция, то математическое ожидание случайной величины $\eta = \varphi(\xi)$ можно вычислить по формуле*

$$\mathbb{E}\eta = \sum_{i=1}^{n} \varphi(x_i) \mathbb{P}(D_i).$$

Свойства математического ожидания простых случайных величин.

1°. Для любого $A \in \mathcal{A}$ имеет место равенство $\mathbb{E}\mathbb{1}_A = \mathbb{P}(A)$;

2°. (Линейность.) Если ξ, η — случайные величины и $\alpha \in \mathbb{R}$, то $\mathbb{E}(\alpha\xi) = \alpha\mathbb{E}\xi$ и $\mathbb{E}(\xi + \eta) = \mathbb{E}\xi + \mathbb{E}\eta$;

3°. (Монотонность.) Если $\xi \geq 0$, то $\mathbb{E}\xi \geq 0$ и равенство $\mathbb{E}\xi = 0$ возможно лишь в случае $\mathbb{P}(\{\omega : \xi(\omega) \neq 0\}) = 0$, т. е. случайная величина ξ почти наверное равна нулю;

4°. Для любой случайной величины ξ имеет место неравенство $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$;

5°. (Неравенство Шварца.) Для любых случайных величин ξ и η имеет место неравенство $(\mathbb{E}\xi\eta)^2 \leq (\mathbb{E}\xi^2)(\mathbb{E}\eta^2)$.

Доказательство *Свойство* $1°$ *очевидно. Равенство* $\mathbb{E}(\alpha\xi) = \alpha\mathbb{E}\xi$ *следует из замечания 5.2. Допустим теперь, что* $\xi = \sum_{i=1}^n x_i \mathbb{1}_{D_i}$, $\eta = \sum_{j=1}^m y_j \mathbb{1}_{H_j}$ *— две случайные величины, а* $\{D_1, \ldots, D_n\}$, $\{H_1, \ldots, H_m\}$ *— соответствующие им разбиения. Тогда случайная величина* $\xi + \eta$ *может быть представлена в виде*

$$\xi + \eta = \sum_{i,j} (x_i + y_j) \mathbb{1}_{D_i H_j}.$$

Но тогда, в силу замечания 5.1 имеем

$$\begin{aligned}
\mathbb{E}(\xi + \eta) &= \sum_{i,j} (x_i + y_j)\mathbb{P}(D_i H_j) = \\
&= \sum_{i,j} x_i \mathbb{P}(D_i H_j) + \sum_{i,j} y_j \mathbb{P}(D_i H_j) = \\
&= \sum_i x_i \sum_j \mathbb{P}(D_i H_j) + \sum_j y_j \sum_i \mathbb{P}(D_i H_j) = \\
&= \sum_i \mathbb{P}(D_i) + \sum_j y_j \mathbb{P}(H_j) = \\
&= \mathbb{E}\xi + \mathbb{E}\eta
\end{aligned}$$

и свойство $2°$ *доказано.*

Условие $\xi \geq 0$ *означает, что в представлении* (5.1) *все* x_i, $i = 1, \ldots, n$, *неотрицательны. Но тогда и* $\mathbb{E}\xi$, *как сумма неотрицательных слагаемых, также будет неотрицательной. Равенство* $\mathbb{E}\xi = 0$ *возможно лишь в* *случае, когда $\mathbb{P}(D_i) = 0$ для тех индексов i, которые соответствуют строго положительным значениям x_i. Это доказывает свойство 3°.*

Свойство 4° является непосредственным следствием неравенства треугольника. Для доказательства 5° заметим вначале, что в случае равенства нулю правой части неравенства Шварца обращается в нуль и левая его часть. Если, например, $\mathbb{E}\xi^2 = 0$, то $\mathbb{P}(\xi = 0) = 1$ и $\mathbb{E}\xi\eta = 0$. Допустим теперь, что $\mathbb{E}\xi^2 > 0$ и $\mathbb{E}\eta^2 > 0$. Тогда корректно определены случайные величины

$$\hat{\xi} = \frac{|\xi|}{\sqrt{\mathbb{E}\xi^2}}, \qquad \hat{\eta} = \frac{|\eta|}{\sqrt{\mathbb{E}\eta^2}}.$$

Замечая, что $\mathbb{E}\hat{\xi}^2 = \mathbb{E}\hat{\eta}^2 = 1$, и используя очевидное неравенство $2\hat{\xi}\hat{\eta} \le \hat{\xi}^2 + \hat{\eta}^2$, получаем

$$2\mathbb{E}\hat{\xi}\hat{\eta} \le \mathbb{E}\hat{\xi}^2 + \mathbb{E}\hat{\eta}^2 = 2, \qquad \mathbb{E}\hat{\xi}\hat{\eta} \le 1,$$

что эквивалентно неравенству Шварца. □

Другой важной числовой характеристикой случайной величины является дисперсия, которая характеризует степень "<разброса">" значений случайной величины.

Определение 5.2 *Дисперсией случайной величины ξ называется число*

$$\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2.$$

Величина $\sigma_\xi = \sqrt{\mathbb{D}\xi}$ называется стандартным (или среднеквадратическим) отклонением.

Свойства дисперсии.

1°. $\mathbb{D}\xi = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2$;

2°. Для любого вещественного числа c и случайной величины ξ имеют место равенства

$$\mathbb{D}(c\xi) = c^2\mathbb{D}\xi, \qquad \mathbb{D}(\xi + c) = \mathbb{D}\xi;$$

3°. Равенство $\mathbb{D}\xi = 0$ возможно лишь в случае $\mathbb{P}(\xi = \mathbb{E}\xi) = 1$, т. е. случайная величина ξ почти наверное равна постоянной.

Доказательство *Используя свойство линейности математического ожидания, получаем*

$$\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - 2\mathbb{E}\xi \cdot \mathbb{E}\xi + (\mathbb{E}\xi)^2 = \mathbb{E}\xi^2 - (\mathbb{E}\xi)^2.$$

Пусть теперь $c \in \mathbb{R}$ и ξ — случайная величина. Тогда

$$\mathbb{D}(c\xi) \,=\, \mathbb{E}(c\xi - c\mathbb{E}\xi)^2 \,=\, c^2\mathbb{E}(\xi - \mathbb{E}\xi)^2 \,=\, c^2\mathbb{D}\xi,$$

а также

$$\mathbb{D}(\xi + c) \,=\, \mathbb{E}(\xi + c - \mathbb{E}(\xi + c))^2 \,=\, \mathbb{E}(\xi - \mathbb{E}\xi)^2 \,=\, \mathbb{D}\xi.$$

Из свойства монотонности математического ожидания 3° следует, что $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2 \geq 0$ и равенство $\mathbb{D}\xi = 0$ возможно лишь в случае $\xi - \mathbb{E}\xi = 0$ почти наверное. □

Независимость случайных величин.

Пусть ξ — случайная величина с представлением (5.1). Информация о том, какое из значений $x_1, \ldots, x_n$ приняла случайная величина ξ, эквивалентна тому, что мы знаем, какое из событий $D_1, \ldots, D_n$ разбиения $\mathbb{D}_\xi$ произошло. Но тогда мы можем сказать о любом событии из алгебры $\mathcal{A}_\xi$ произошло оно или нет. Другими словами, алгебра $\mathcal{A}_\xi$ отражает информацию, связанную со случайной величиной ξ. В связи с этим естественно ввести следующее определение.

Определение 5.3 *Случайные величины ξ и η, определенные на вероятностном пространстве $(\Omega, \mathcal{A}, \mathbb{P})$, называются независимыми, если независимы алгебры $\mathcal{A}_\xi$ и $\mathcal{A}_\eta$.*

Напомним, что условие независимости алгебр $\mathcal{A}_\xi$ и $\mathcal{A}_\eta$ эквивалентно тому, что для любых $A \in \mathcal{A}_\xi$ и $B \in \mathcal{A}_\eta$ выполняется равенство $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$. Кроме того, в силу теоремы 3.4 для того, чтобы ξ и η были независимы, достаточно выполнения условия $\mathbb{P}(DH) = \mathbb{P}(D)\mathbb{P}(H)$ для любых $D \in \mathbb{D}_\xi$ и $H \in \mathbb{D}_\eta$.

Теорема 5.1 *Если ξ и η — независимые случайные величины, то*

$$\mathbb{E}\xi\eta \,=\, \mathbb{E}\xi \cdot \mathbb{E}\eta, \qquad \mathbb{D}(\xi + \eta) \,=\, \mathbb{D}\xi + \mathbb{D}\eta.$$

Доказательство *Пусть*

$$\xi \,=\, \sum_{i=1}^{n} x_i \mathbb{1}_{D_i}, \qquad \eta \,=\, \sum_{j=1}^{m} y_j \mathbb{1}_{H_j}.$$

Тогда, в силу независимости случайных величин, выполняются равенства $\mathbb{P}(D_i H_j) = \mathbb{P}(D_i)\mathbb{P}(H_j)$ и, следовательно,

$$\begin{aligned}\mathbb{E}(\xi\eta) \,&=\, \sum_{i,j} x_i y_j \mathbb{P}(D_i H_j) \,=\, \sum_{i,j} x_i y_j \mathbb{P}(D_i)\mathbb{P}(H_j) \\ &=\, \left(\sum_{i=1}^{n} x_i \mathbb{P}(D_i)\right)\left(\sum_{j=1}^{m} y_j \mathbb{P}(H_j)\right) \,=\, \mathbb{E}\xi\mathbb{E}\eta.\end{aligned}$$

Далее,

$$\mathbb{D}(\xi+\eta) = \mathbb{E}(\xi+\eta)^2-(\mathbb{E}(\xi+\eta))^2 = \mathbb{E}\xi^2+2\mathbb{E}(\xi\eta)+\mathbb{E}\eta^2-(\mathbb{E}\xi)^2-2\mathbb{E}\xi\mathbb{E}\eta-(\mathbb{E}\eta)^2.$$

Отсюда с использованием равенства $\mathbb{E}\xi\eta = \mathbb{E}\xi\mathbb{E}\eta$ *получаем, что* $\mathbb{D}(\xi+\eta) = \mathbb{D}\xi + \mathbb{D}\eta$ *и теорема доказана.* □

Замечание 5.3 *Если* ξ *и* η *— две независимые случайные величины, а* $\varphi : \mathbb{R} \mapsto \mathbb{R}$ *— некоторая функция, то* $\varphi(\xi)$ *и* η *также являются независимыми случайными величинами.*

Действительно, это сразу же следует из включения $\mathcal{A}_{\varphi(\xi)} \subset \mathcal{A}_\xi$.

Ковариация и коэффициент корреляции. Эти числовые характеристики можно рассматривать как меру зависимости случайных величин.

Определение 5.4 *Пусть* ξ *и* η *— две случайные величины, определенные на одном вероятностном пространстве. Под ковариацией этих случайных величин понимается число*

$$\operatorname{cov}(\xi,\eta) = \mathbb{E}(\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta).$$

Для вычисления ковариации нужно знать совместное распределение случайных величин. Пусть $\xi = \sum_{i=1}^n x_i \mathbb{1}_{D_i}$, $\eta = \sum_{j=1}^m y_j \mathbb{1}_{H_j}$, т. е. $\mathbb{D}_\xi = \{D_1, \dots, D_n\}$, $\mathbb{D}_\eta = \{H_1, \dots, H_m\}$. Рассмотрим случайный вектор (ξ, η), который принимает значение (x_i, y_j) на $D_i H_j$, $i = 1, \dots, n$, $j = 1, \dots, m$. Другими словами, случайному вектору (ξ, η) можно сопоставить разбиение

$$\mathbb{D}_{\xi,\eta} = \{D_i H_j : i = 1, \dots, n, j = 1, \dots, m\}.$$

Вероятности $p_{ij} = \mathbb{P}(D_i H_j) = \mathbb{P}(\xi = i, \eta = j)$ определяют *совместное распределение* случайных величин ξ и η. Совместное распределение двух простых случайных величин удобно представить таблицей.

$\xi \diagdown \eta$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{11}	p_{12}	$\dots$	p_{1m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

Отметим некоторые особенности этой таблицы. Поскольку $\mathbb{D}_{\xi\eta}$ является разбиением, то $\sum_{i,j} p_{ij} = 1$. Кроме того,

$$\sum_{j=1}^m p_{ij} = p_i^x = \mathbb{P}(\xi = x_i), \qquad \sum_{i=1}^n p_{ij} = p_j^y = \mathbb{P}(\eta = y_j).$$

Таким образом, совместное распределение случайных величин ξ и η позволяет получить их индивидуальные распределения. Условие независимости случайных величин ξ и η эквивалентна выполнению равенств $p_{ij} = p_i^x p_j^y$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Заметим также, что любая такая таблица с выполнением условий $p_{ij} \geq 0$, $\sum_{i,j} p_{ij} = 1$ определяет совместное распределение двух случайных величин.

Наряду с ковариацией также рассматривают коэффициент корреляции. Его вводят в рассмотрение при условии, что $\mathbb{D}\xi > 0$ и $\mathbb{D}\eta > 0$. Заметим, что если дисперсия одной из случайных величин равна нулю, например $\mathbb{D}\xi = 0$, то $\xi = \mathbb{E}\xi$ почти наверное и в этом случае $\operatorname{cov}(\xi, \eta) = 0$. Допустим теперь, что $\mathbb{D}\xi > 0$ и $\mathbb{D}\eta > 0$. Тогда корректно определены случайные величины

$$\hat{\xi} = \frac{\xi - \mathbb{E}\xi}{\sqrt{\mathbb{D}\xi}}, \qquad \hat{\eta} = \frac{\eta - \mathbb{E}\eta}{\sqrt{\mathbb{D}\eta}},$$

которые удовлетворяют условиям

$$\mathbb{E}\hat{\xi} = \mathbb{E}\hat{\eta} = 0, \qquad \mathbb{D}\hat{\xi} = \mathbb{D}\hat{\eta} = 1.$$

Коэффициент корреляции $\varrho(\xi, \eta)$ определяется равенством

$$\varrho(\xi, \eta) = \mathbb{E}(\hat{\xi}\hat{\eta}) = \frac{\operatorname{cov}(\xi, \eta)}{\sqrt{\mathbb{D}\xi} \cdot \sqrt{\mathbb{D}\eta}}.$$

Свойства ковариации и коэффициента корреляции.

1°. $\operatorname{cov}(\xi, \eta) = \mathbb{E}\xi\eta - \mathbb{E}\xi \cdot \mathbb{E}\eta$;

2°. Если ξ и η независимы, то $\operatorname{cov}(\xi, \eta) = 0$;

3°. $|\varrho(\xi, \eta)| \leq 1$ и $|\varrho(\xi, \eta)| = 1$ в том и только том случае, если $\mathbb{P}(\eta = a\xi + b) = 1$ при некоторых $a, b \in \mathbb{R}$.

Доказательство *Свойства 1° и 2° очевидны. Остается доказать свойство 3°. В силу неотрицательности дисперсии имеем*

$$0 \leq \mathbb{D}(\hat{\xi} \pm \hat{\eta}) = \mathbb{D}\hat{\xi} + \mathbb{D}\hat{\eta} \pm 2\varrho(\xi, \eta) = 2(1 \pm \varrho(\xi, \eta)).$$

Отсюда следует, что $-1 \leq \varrho(\xi, \eta) \leq 1$. *Если же* $\varrho(\xi, \eta) = 1$, *то* $\mathbb{D}(\hat{\xi} - \hat{\eta}) = 0$ *и тогда* $\hat{\xi} - \hat{\eta} \equiv \mathrm{const}$ *почти наверное. Это влечёт линейную зависимость* ξ *и* η. *Аналогично, если* $\varrho(\xi, \eta) = -1$, *то* $\mathbb{D}(\hat{\xi} + \hat{\eta}) = 0$ *и тогда* $\hat{\xi} + \hat{\eta} \equiv \mathrm{const}$ *почти наверное.* □

Пример. Пусть ζ — случайная величина, которая с равными вероятностями $1/3$ принимает значения $0, \pi/2$ и π. Тогда случайные величины $\xi = \cos\zeta$ и $\eta = \sin\zeta$ функционально связаны, но $\operatorname{cov}(\xi, \eta) = 0$.

Определение 5.5 *Случайные величины ξ и η называются некоррелированными, если* $\operatorname{cov}(\xi,\eta)=0$.

Для вычисления ковариации случайных величин достаточно знать таблицу их совместного распределения, поскольку

$$\mathbb{E}\xi = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} p_{ij}, \qquad \mathbb{E}\eta = \sum_{j=1}^{m} y_j \sum_{i=1}^{n} p_{ij}, \qquad \mathbb{E}\xi\eta = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j p_{ij}.$$

Отметим также, что непосредственно из определений следуют равенства

$$\mathbb{D}(\xi+\eta) = \mathbb{D}\xi + \mathbb{D}\eta + 2\operatorname{cov}(\xi,\eta), \qquad \operatorname{cov}(\xi,\eta) = \sqrt{\mathbb{D}\xi}\cdot\sqrt{\mathbb{D}\eta}\cdot\varrho(\xi,\eta).$$

Ковариационная матрица.

Пусть $\xi_1,\ldots,\xi_n$ — случайные величины, определённые на одном вероятностном пространстве. Тогда степень их попарной зависимости отражает ковариационная матрица $\mathbb{V} = (v_{ij})$, $v_{ij} = \operatorname{cov}(\xi_i,\xi_j)$, $i,j = 1,\ldots,n$. В случае $i=j$ элемент v_{ij} является дисперсией $\mathbb{D}\xi$. Отметим некоторые особенности ковариационной матрицы. Это симметрическая матрица, на диагонали которой стоят неотрицательные числа. Кроме того, она неотрицательно определена, т. е.

$$\mathbf{x}^t\mathbb{V}\mathbf{x} = \sum_{i,j=1}^{n} v_{ij}x_ix_j \geq 0$$

для всех $\mathbf{x}=(x_1,\ldots,x_n)$ из $\mathbb{R}^n$. Будем $\mathbf{x}\in\mathbb{R}^n$ представлять как вектор-столбец. Также сформируем случайный вектор-столбец $\vec{\xi} = (\xi_1,\ldots,\xi_n)^t$ и пусть $\mathbb{E}\vec{\xi} = (\mathbb{E}\xi,\ldots,\mathbb{E}\xi_n)^t$ — его математическое ожидание. Тогда

$$\mathbb{V} = \mathbb{E}\{(\vec{\xi}-\mathbb{E}\vec{\xi})(\vec{\xi}-\mathbb{E}\vec{\xi})^t\},$$

а значение квадратичной формы на векторе $\mathbf{x}\in\mathbb{R}^n$ запишется в виде

$$\mathbf{x}^t\mathbb{V}\mathbf{x} = \mathbb{E}\{\mathbf{x}^t(\vec{\xi}-\mathbb{E}\vec{\xi})(\vec{\xi}-\mathbb{E}\vec{\xi})^t\mathbf{x}\} = \mathbb{E}(\mathbf{x}^t(\vec{\xi}-\mathbb{E}\vec{\xi}))^2 \geq 0.$$

Замечание 5.4 *Симметричность и неотрицательная определённость являются характеристическими свойствами ковариационной матрицы.*

Если $\mathbb{D}\xi_i = v_{ii} > 0$ при всех $i=1,\ldots,n$, то определена также корреляционная матрица $\mathbb{K} = (\varrho_{ij})$, $\varrho_{ij} = v_{ij}/(\sigma_i\sigma_j)$, где $\sigma_i = \sqrt{\mathbb{D}\xi_i}$. Корреляционная матрица также является симметрической и неотрицательно определённой. Её отличительной чертой является то, что на главной диагонали стоят единицы.

Задача линейного оценивания.

Пусть ξ и η — две случайные величины, из которых лишь ξ является наблюдаемой. Если ξ и η коррелированы, то естественно предположить, что знание значения ξ позволит получить некоторые выводы о значениях ненаблюдаемой случайной величины η. Всякую функцию $\varphi : \mathbb{R} \mapsto \mathbb{R}$ в контексте этой задачи называют оценкой для η. Оценка φ^* называется оптимальной в смысле среднеквадратического отклонения в классе оценок Φ, если

$$\mathbb{E}(\eta - \varphi^*(\xi))^2 \;=\; \inf_{\varphi \in \Phi} \mathbb{E}(\eta - \varphi(\xi))^2.$$

Оказывается, что для отыскания оптимальной оценки в классе линейных функций $\varphi(x) = ax + b$ достаточно знания ковариации $\mathrm{cov}(\xi, \eta)$.

Рассмотрим функцию двух переменных

$$\psi(a, b) \;=\; \mathbb{E}(\eta - (a\xi + b))^2,$$

которая представляет собой неотрицательную квадратичную форму относительно переменных a и b. Поэтому она достигает минимума в некоторой точке (a^*, b^*), которая является решением системы уравнений

$$\frac{\partial \psi}{\partial a}(a, b) \;=\; 0, \qquad \frac{\partial \psi}{\partial b}(a, b) \;=\; 0.$$

Вычисляя частные производные функции ψ, получаем

$$\begin{cases} \mathbb{E}\xi\eta - a\mathbb{E}\xi^2 - b\mathbb{E}\xi \;=\; 0, \\ \quad \mathbb{E}\eta - a\mathbb{E}\xi - b \;=\; 0. \end{cases}$$

Умножая второе равенство на $\mathbb{E}\xi$ и вычитая из первого, приходим к соотношению

$$\mathrm{cov}(\xi, \eta) - a\mathbb{D}\xi \;=\; 0,$$

откуда находим

$$a^* \;=\; \frac{\mathrm{cov}(\xi, \eta)}{\mathbb{D}\xi}$$

и далее

$$b^* \;=\; \mathbb{E}\eta - \frac{\mathrm{cov}(\xi, \eta)}{\mathbb{D}\xi}\mathbb{E}\xi.$$

Следовательно, оптимальной оценкой в классе линейных функций является

$$\eta^* \;=\; \varphi^*(\xi) \;=\; \mathbb{E}\eta + \frac{\mathrm{cov}(\xi, \eta)}{\mathbb{D}\xi}(\xi - \mathbb{E}\xi).$$

Это равенство называют также уравнением регрессии η на ξ.

Заметим также, что среднеквадратическая ошибка линейного оценивания вычисляется по формуле

$$\begin{aligned}\mathbb{E}(\eta-\eta^*)^2 &= \mathbb{E}\left((\eta-\mathbb{E}\eta)-\frac{\mathrm{cov}(\xi,\eta)}{\mathbb{D}\xi}(\xi-\mathbb{E}\xi)\right)^2 \\ &= \mathbb{D}\eta-2\frac{(\mathrm{cov}(\xi,\eta))^2}{\mathbb{D}\xi}+\frac{(\mathrm{cov}(\xi,\eta))^2}{\mathbb{D}\xi} \\ &= \mathbb{D}\eta\left(1-(\varrho(\xi,\eta))^2\right).\end{aligned}$$

Отсюда видно, что оценка тем точнее, чем ближе коэффициент корреляции $\varrho(\xi,\eta)$ по модулю к единице.

Целочисленные случайные величины и производящие функции.

С точки зрения распределения вероятностей легко ввести в рассмотрение случайные величины, принимающие счётное число значений: $\mathbb{P}(\xi = x_k) = p_k$, $p_k \geq 0$ и $\sum_{k=1}^{\infty} p_k = 1$. Однако в этом случае математическое ожидание и дисперсия не всегда определены.

Определение 5.6 *Пусть случайная величина ξ принимает счётное число значений: $\mathbb{P}(\xi = x_k) = p_k$, $k = 1, 2, \ldots$. Будем говорить, что для ξ определено математическое ожидание, если сходится числовой ряд $\sum_{k=1}^{\infty} |x_k| p_k$. В этом случае определяется*

$$\mathbb{E}\xi = \sum_{k=1}^{\infty} x_k p_k.$$

Дискретную случайную величину ξ, принимающую только целые неотрицательные значения, называют *целочисленной* случайной величиной. Её распределение вероятностей $\mathbb{P}(\xi = k) = p_k$, $k = 0, 1, 2, \ldots$, удобно представлять *производящей функцией*

$$g_\xi(x) = \mathbb{E}x^\xi = \sum_{k=0}^{\infty} p_k x^k.$$

Заметим, что степенной ряд, определяющий производящую функцию, сходится при $|x| \leq 1$. При этом $g_\xi(1) = 1$ и

$$p_k = \frac{1}{k!} g_\xi^{(k)}(0), \qquad k = 0, 1, 2, \ldots.$$

Если $p_k \neq 0$ лишь для конечного числа индексов, то $g_\xi(x)$ представляет собой полином, а ξ принимает конечное число значений. В терминах производных от производящей функции легко вычисляются математическое ожидание и дисперсия случайной величины. Действительно,

$$\mathbb{E}\xi = \sum_{k=1}^{\infty} kp_k = g'_\xi(1),$$

и аналогично

$$\mathbb{D}\xi = \sum_{k=1}^{\infty} k^2 p_k - \left(\sum_{k=1}^{\infty} kp_k\right)^2 = g''_\xi(1) + g'_\xi(1) - \big(g'_\xi(1)\big)^2.$$

Наиболее часто встречающиеся в приложениях дискретные случайные величины являются целочисленными. Перечислим некоторые из них и найдём их производящие функции и числовые характеристики.

◇ Бернуллиевское распределение.

$$\mathbb{P}(\xi = 1) = p, \qquad \mathbb{P}(\xi = 0) = q, \quad q = 1-p, \quad 0 < p < 1.$$

$$g_\xi(x) = px + q, \quad \mathbb{E}\xi = p, \quad \mathbb{D}\xi = pq.$$

◇ Биномиальное распределение.

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

$$g_\xi(x) = (px+q)^n, \qquad \mathbb{E}\xi = np, \qquad \mathbb{D}\xi = npq.$$

◇ Пуассоновское распределение.

$$\mathbb{P}(\xi = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad \lambda > 0, \quad k = 0, 1, 2, \ldots.$$

$$g_\xi(x) = e^{\lambda(x-1)}, \qquad \mathbb{E}\xi = \lambda, \qquad \mathbb{D}\xi = \lambda.$$

◇ Геометрическое распределение.

$$\mathbb{P}(\xi = k) = pq^k, \qquad 0 < p < 1, \qquad q = 1-p, \qquad k = 0, 1, \ldots.$$

$$g_\xi(x) = \frac{p}{1-qx}, \qquad \mathbb{E}\xi = \frac{q}{p}, \qquad \mathbb{D}\xi = \frac{q}{p^2}.$$