Г Л А В А III

СЛУЧАЙНЫЕ ВЕКТОРЫ

§ 9. Двумерные случайные величины

9.1. Функция распределения.

О п р е д е л е н и е 9.1. *Двумерным случайным вектором* (*или двумерной СВ*) $Z \triangleq \mathrm{col}(X, Y)$ называется вектор, компонентами которого являются СВ $X = X(\omega)$ и $Y = Y(\omega)$, определенные на одном и том же пространстве Ω элементарных событий ω.

П р и м е р 9.1. Скорость ветра в текущий момент времени в конкретной точке на поверхности Земли можно представить в виде двумерной СВ.

О п р е д е л е н и е 9.2. Функция

$$F(x, y) \triangleq \mathbf{P}(\{\omega : X(\omega) \leqslant x\}\{\omega : Y(\omega) \leqslant y\}) \triangleq \mathbf{P}\{X \leqslant x,\, Y \leqslant y\},$$

определенная для всех $(x, y) \in \mathbb{R}^2$, называется *функцией распределения* двумерной СВ $Z = \mathrm{col}(X, Y)$.

П р и м е р 9.2. Пусть опыт G состоит в одновременном подбрасывании двух монет. Элементарным событием будем считать положение упавших монет. Пусть СВ $X(\omega)$ — число выпавших «гербов», а СВ $Y(\omega)$ — расстояние между монетами. В этом случае $X(\omega)$ — дискретная СВ со значениями $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, а $Y(\omega)$ — непрерывная СВ с реализациями в $\mathbb{R}^1$.

Значение функции распределения $F(x_1, y_1)$ равно вероятности попадания двумерной СВ (X, Y) в квадрант D_{11} с вершиной в точке (x_1, y_1) и сторонами, параллельными осям координат (рис. 9.1).

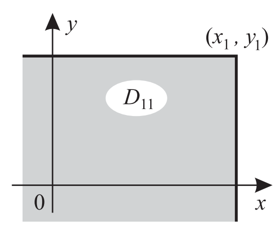

Рис. 9.1

Свойства $F(x,y)$

1) $F(x,y)$ определена для всех $(x,y)\in\mathbb{R}^2$, так как вероятность $\mathbf{P}\{X\leqslant x,\ Y\leqslant y\}$ определена для всех $(x,y)\in\mathbb{R}^2$.

2) $0\leqslant F(x,y)\leqslant 1$ для всех $x,y\in\mathbb{R}^1$. По аксиоме **A1** и свойству 5)P для любого события выполняется $0\leqslant \mathbf{P}(A)\leqslant 1$, поэтому по определению $F(x,y)\in[0,1]$.

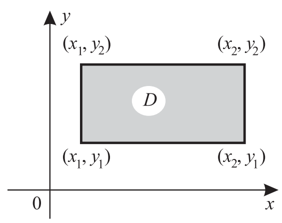

Рис. 9.2

3) $F(-\infty,y)=F(x,-\infty)=F(-\infty,-\infty)=0$ для всех $x,y\in\mathbb{R}^1$. Действительно, рассмотрим следующую последовательность событий $B_n\triangleq\{\omega\colon\ X(\omega)\leqslant -n\}$, где $n=1,2,\ldots$ По аналогии с доказательством свойства 3) $F(x)$ устанавливаем, что

$$F(-\infty,y)\leqslant\lim_{n\to\infty}\mathbf{P}(B_n)=\mathbf{P}(\varnothing)=0.$$

4) $F_Y(y)=F(+\infty,y)$, $F_X(x)=F(x,+\infty)$ для всех $x,y\in\mathbb{R}^1$, где $F_X(y)$ и $F_Y(x)$ — функции распределения СВ X и Y соответственно. Это свойство можно установить, следуя доказательству свойства 3) $F(x)$, так как $\{\omega: X(\omega)\leqslant+\infty\}=\{\omega: Y(\omega)\leqslant+\infty\}=\Omega$.

5) $F(+\infty,+\infty)=1$. В силу свойств 4) $F(x,y)$ и 3) $F(x)$ имеем

$$F(+\infty,+\infty)=F_X(+\infty)=1.$$

6) $F(x,y)$ — монотонно неубывающая по каждому из аргументов. Так для $\Delta x>0$ имеем

$$F(x+\Delta x,y)\triangleq\mathbf{P}\{X\leqslant x+\Delta x,\ Y\leqslant y\}=$$
$$=\mathbf{P}\{X\leqslant x,\ Y\leqslant y\}+\mathbf{P}\{x<X\leqslant x+\Delta x,\ Y\leqslant y\}\geqslant F(x,y),$$

Монотонность $F(x,y)$ по y доказывается аналогично.

7) Вероятность попадания СВ Z в прямоугольник $D=\{x_1<x\leqslant x_2,\ y_1<y\leqslant y_2\}$ равна

$$\mathbf{P}(D)=F(x_2,y_2)+F(x_1,y_1)-F(x_1,y_2)-F(x_2,y_1),$$

где $F(x_i,y_j)$, $i,j=1,2$, — вероятности попадания СВ Z в квадранты D_{ij} с соответствующими вершинами (x_i,y_j), $i,j=1,2$ (рис. 9.2).

Представим квадрант D_{22} тремя способами в виде суммы непересекающихся множеств:

$$D_{22}=D+D_{12}\backslash D_{11}+D_{21},\quad D_{22}=D+D_{21}\backslash D_{11}+D_{12},$$
$$D_{22}=D+D_{12}\backslash D_{11}+D_{21}\backslash D_{11}+D_{11}.$$

Тогда по формуле сложения вероятностей имеем

$$\mathbf{P}(D_{22}) = \mathbf{P}(D) + \mathbf{P}(D_{12} \backslash D_{11}) + \mathbf{P}(D_{21}),$$
$$\mathbf{P}(D_{22}) = \mathbf{P}(D) + \mathbf{P}(D_{21} \backslash D_{11}) + \mathbf{P}(D_{12}),$$
$$\mathbf{P}(D_{22}) = \mathbf{P}(D) + \mathbf{P}(D_{12} \backslash D_{11}) + \mathbf{P}(D_{21} \backslash D_{11}) + \mathbf{P}(D_{11}).$$

Из последних двух равенств следует, что

$$\mathbf{P}(D_{12} \backslash D_{11}) = \mathbf{P}(D_{12}) - \mathbf{P}(D_{11}).$$

Подставляя это выражение в первое равенство, получаем

$$\mathbf{P}(D) = \mathbf{P}(D_{22}) + \mathbf{P}(D_{11}) - \mathbf{P}(D_{12}) - \mathbf{P}(D_{21}).$$

Но $\mathbf{P}(D_{ij}) = F(x_i, y_j)$, $i = \overline{1,2}$, $j = \overline{1,2}$. Подставляя данные значения функции распределения $F(x,y)$ в формулу для $\mathbf{P}(D)$, приходим к требуемому утверждению.

Если функция распределения $F(x,y)$ непрерывна, то доказанная формула верна и для замкнутого прямоугольника $D = \{x_1 \leqslant x \leqslant x_2,\, y_1 \leqslant y \leqslant y_2\}$.

Определение 9.3. Двумерная СВ $Z = \operatorname{col}(X,Y)$ называется *дискретной*, если СВ X и Y дискретны.

Простейшим способом задания закона распределения дискретной двумерной СВ Z является таблица распределения (см. табл. 9.1).

Здесь $p_{ij} \triangleq \mathbf{P}(\{X = x_i\}\{Y = y_j\}) = \mathbf{P}\{X = x_i,\, Y = y_j\}$, $i = \overline{0,n}$, $j = \overline{0,m}$, с условием нормировки: $\sum\limits_{i=0}^{n} \sum\limits_{j=0}^{m} p_{ij} = 1$. Функция распределения имеет вид

Таблица 9.1

$X \backslash Y$	y_0	y_1	$\cdots$	y_m
x_0	p_{00}	p_{01}	$\cdots$	p_{0m}
x_1	p_{10}	p_{11}	$\cdots$	p_{1m}
$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
x_n	p_{n0}	p_{n1}	$\cdots$	p_{nm}

$$F(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} p_{ij} l(x - x_i) l(y - y_j),$$

где $l(\cdot)$ — единичные ступенчатые функции (см. определение 5.6).

Определение 9.4. СВ X и Y называются *независимыми*, если

$$F(x,y) = F_X(x) F_Y(y) \ \text{для всех} \ x \in \mathbb{R}^1, y \in \mathbb{R}^1.$$

Пусть $Z = \operatorname{col}(X,Y)$ — дискретная СВ, т. е. X и Y — одномерные дискретные СВ с реализациями $x_0, \ldots, x_n$ и $y_0, \ldots, y_m$. Можно показать, что СВ X и Y независимы тогда и только тогда, когда для всех $i = \overline{0,n}$; $j = \overline{0,m}$

$$\mathbf{P}\{X = x_i,\, Y = y_j\} = \mathbf{P}\{X = x_i\}\mathbf{P}\{Y = y_j\}.$$

9.2. Плотность распределения.

Определение 9.5. Неотрицательная кусочно непрерывная функция $f(x,y)$ называется *плотностью распределения* (*плотностью вероятности*) двумерной СВ $Z = \mathrm{col}(X,Y)$, если

$$F(x,y) = \int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y} f(t,\tau)\,dt\,d\tau,$$

где использована символическая запись для двойного интеграла по области $D \triangleq \{t \leqslant x, \tau \leqslant y\}$. При этом двумерная СВ Z называется *непрерывной*.

Из математического анализа известен факт, что в точках непрерывности функции $f(x,y)$ она равна второй смешанной производной функции $F(x,y)$, т. е.:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x\,\partial y}.$$

Свойства $f(x,y)$

1) $f(x,y) \geqslant 0$ для всех $x, y \in \mathbb{R}^1$. Это вытекает из определения 9.5.

2) $\mathbf{P}(D) = \int\limits_{x_1}^{x_2}\Big(\int\limits_{y_1}^{y_2} f(x,y)\,dy\Big)\,dx$, где $D \triangleq \{x_1 \leqslant x \leqslant x_2,\ y_1 \leqslant y \leqslant y_2\}$.

По свойству 7) $F(x,y)$ и определению 9.5 имеем

$$\mathbf{P}(D) = F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1) \triangleq$$
$$\triangleq \int\limits_{-\infty}^{x_2}\Big(\int\limits_{-\infty}^{y_2} f(x,y)\,dy\Big)dx - \int\limits_{-\infty}^{x_2}\Big(\int\limits_{-\infty}^{y_1} f(x,y)\,dy\Big)dx - \int\limits_{-\infty}^{x_1}\Big(\int\limits_{-\infty}^{y_2} f(x,y)\,dy\Big)dx +$$
$$+ \int\limits_{-\infty}^{x_1}\Big(\int\limits_{-\infty}^{y_1} f(x,y)\,dy\Big)dx = \int\limits_{x_1}^{x_2}\Big(\int\limits_{y_1}^{y_2} f(x,y)\,dy\Big)dx.$$

3) $\mathbf{P}(D) = \iint\limits_{D} f(x,y)\,dx\,dy$, где D — произвольная квадрируемая область на плоскости $\mathbb{R}^2$. Доказательство проведем для непрерывной $f(x,y)$. Рассмотрим бесконечно малый прямоугольник $\Delta D \triangleq \{x \leqslant X \leqslant x + \Delta x,\ y \leqslant Y \leqslant y + \Delta y\}$. Согласно свойству 2) $f(x,y)$ можно записать

$$\mathbf{P}(\Delta D) = \int\limits_{x}^{x+\Delta x}\Big(\int\limits_{y}^{y+\Delta y} f(t,\tau)\,d\tau\Big)\,dt = \left\|\begin{matrix}\text{по теореме}\\ \text{о среднем}\\ \text{значении}\end{matrix}\right\| \approx f(x,y)\Delta y \Delta x.$$

Таким образом, элемент вероятности $f(x,y)\,dx\,dy$ с точностью до бесконечно малых первого порядка равен вероятности попадания двумерной СВ $Z=\mathrm{col}(X,Y)$ в бесконечно малый прямоугольник, прилегающий к точке (x,y), со сторонами, параллельными осям координат. Так как квадрируемую область $D\subset\mathbb{R}^2$ можно представить с любой степенью точности в виде объединения конечного числа непересекающихся бесконечно малых прямоугольников ΔD, то из аксиомы **А3** следует формула для вероятности попадания СВ (X,Y) в D.

4) $\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)\,dx\,dy=1$, поскольку по свойству 5) $F(x,y)$

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)\,dx\,dy \triangleq F(+\infty,+\infty)=1.$$

5) $F_X(x)=\int\limits_{-\infty}^{x}\Big(\int\limits_{-\infty}^{+\infty} f(t,y)\,dy\Big)\,dt$, $F_Y(y)=\int\limits_{-\infty}^{y}\Big(\int\limits_{-\infty}^{+\infty} f(x,\tau)\,dx\Big)\,d\tau$,

где $F_X(x)$, $F_Y(y)$ — функции распределения СВ X и Y. Например, по свойству 4) $F(x,y)$

$$F_X(x)=F(x,+\infty)\triangleq\int\limits_{-\infty}^{x}\Big(\int\limits_{-\infty}^{+\infty} f(t,y)\,dy\Big)\,dt.$$

Для $F_Y(y)$ утверждение доказывается аналогично.

6) $f_X(x)=\int\limits_{-\infty}^{+\infty} f(x,y)\,dy$, $f_Y(y)=\int\limits_{-\infty}^{+\infty} f(x,y)\,dx$. Это вытекает из свойства 5) $f(x,y)$ и определения 5.8. Свойство 6) $f(x,y)$ позволяет по плотности вероятности двумерной СВ Z найти плотности вероятности СВ X и Y.

7) Пусть СВ $V\triangleq\varphi(X,Y)$, где X, Y — непрерывные СВ с совместной плотностью $f(x,y)$, а $\varphi(x,y)$ — заданная скалярная функция аргументов $x,y\in\mathbb{R}^1$, такая, что $\iint\limits_{-\infty}^{+\infty} |\varphi(x,y)|f(x,y)\,dx\,dy<$ $<\infty$. Тогда можно показать, что

$$\mathbf{M}[V]=\iint\limits_{-\infty}^{+\infty}\varphi(x,y)f(x,y)\,dx\,dy.$$

Причем, если $\varphi(x,y)\triangleq\sum\limits_{k=1}^{m}\varphi_k(x,y)$, то, пользуясь линейными свойствами интеграла, можно показать, что $\mathbf{M}[V] = \sum_{k=1}^{m} \mathbf{M}[V_k]$, где $V_k \triangleq \varphi_k(X, Y)$, $k = \overline{1, m}$.

8) Для независимости непрерывных СВ X и Y достаточно, чтобы выполнялось равенство

$$f(x, y) = f_X(x) f_Y(y)$$

во всех точках непрерывности этих функций. Действительно, по определению плотности

$$\int_{-\infty}^{x} \int_{-\infty}^{y} f(t, \tau)\, dt\, d\tau = F(x, y) = \left\| \begin{array}{l} \text{в силу незави-} \\ \text{симости } X \text{ и } Y \end{array} \right\| = F_X(x) F_Y(y) =$$

$$= \int_{-\infty}^{x} f_X(t)\, dt \int_{-\infty}^{y} f_Y(\tau)\, d\tau = \int_{-\infty}^{x} \int_{-\infty}^{y} f_X(t) f_Y(\tau)\, dt\, d\tau.$$

Откуда и следует данное свойство.

9) Если непрерывные СВ X и Y независимы, то справедлива *формула свертки плотностей*, т. е. плотность распределения СВ $V \triangleq X + Y$ имеет вид

$$f_V(v) = \int_{-\infty}^{+\infty} f_X(x) f_Y(v - x)\, dx,$$

где $f_X(x)$, $f_Y(y)$ — плотность распределения СВ X и Y. Действительно, пусть $D \triangleq \{x, y\colon\ x + y \leqslant v\}$. Тогда по свойствам 2) $f(x, y)$ и 8) $f(x, y)$ получаем

$$F_V(v) \triangleq \mathbf{P}\{X + Y \leqslant v\} = \iint_D f(x, y)\, dx\, dy =$$

$$= \iint_D f_X(x) f_Y(y)\, dx\, dy = \int_{-\infty}^{+\infty} f_X(x) \Big(\int_{-\infty}^{v-x} f_Y(y)\, dy \Big)\, dx = \| y \triangleq t - x \| =$$

$$= \int_{-\infty}^{+\infty} f_X(x) \Big(\int_{-\infty}^{v} f_Y(t - x)\, dt \Big)\, dx = \int_{-\infty}^{v} \Big(\int_{-\infty}^{+\infty} f_X(x) f_Y(t - x)\, dx \Big)\, dt.$$

Отсюда, согласно определению 5.8, вытекает формула свертки.

Можно показать также, что верно равенство

$$f_V(v) = \int_{-\infty}^{+\infty} f_X(v - y) f_Y(y)\, dy.$$

Пример 9.3. Пусть имеются равномерно распределенные СВ $X \sim \mathbf{R}(a;b)$ и $Y \sim \mathbf{R}(a;b)$, которые независимы.

Пользуясь формулой свертки, найдем плотность вероятности СВ $Z \triangleq X+Y$. Учтем, что в данном случае подынтегральное выражение отлично от нуля лишь когда $2a \leqslant v \leqslant 2b$, а именно

$$f_X(x) = \frac{1}{b-a}, \quad \text{если } a \leqslant x \leqslant b,$$

$$f_Y(v-x) = \frac{1}{b-a}, \quad \text{если } a \leqslant v-x \leqslant b.$$

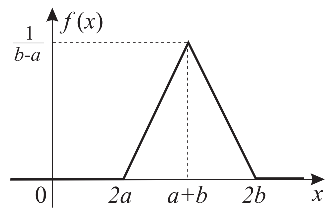

Рис. 9.3

Рассматривая два случая взаимного расположения отрезков, на которых плотности одновременно отличны от нуля, получаем

$$f_V(v) = \frac{1}{(b-a)^2} \int\limits_a^{v-a} dx = \frac{v-2a}{(b-a)^2}, \quad \text{если } 2a \leqslant v \leqslant a+b,$$

$$f_V(v) = \frac{1}{(b-a)^2} \int\limits_{v-b}^{b} dx = \frac{2b-v}{(b-a)^2}, \quad \text{если } a+b < v \leqslant 2b.$$

Таким образом, получаем выражение для плотности *треугольного распределения* (*распределения Симпсона*) (рис. 9.3):

$$f_V(v) = \begin{cases} 0, & v < 2a, \ v > 2b, \\ \dfrac{v-2a}{(b-a)^2}, & 2a \leqslant v \leqslant a+b, \\ \dfrac{2b-v}{(b-a)^2}, & a+b \leqslant v \leqslant 2b. \end{cases}$$

9.3. Типовые задачи.

Задача 9.1. Пусть $f(x,y)$ — плотность вероятности некоторого случайного вектора $Z = \mathrm{col}(X,Y)$. Может ли функция $f_1(u,v) \triangleq \int\limits_{-\infty}^{+\infty} f(u,v)\,du \int\limits_{-\infty}^{+\infty} f(u,v)\,dv$ быть плотностью вероятности какой-либо двумерной случайной величины $T = \mathrm{col}(U,V)$?

Решение. Непосредственно проверяется, что условия неотрицательности и нормировки для функции $f_1(u,v)$ выполнены, т. е. функция $f_1(u,v)$ является плотностью. Укажем двумерную случайную величину T, для которой $f_1(u,v)$ является плотностью. Согласно свойству 6) $f(x,y)$ имеем

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx.$$

Поэтому $f_1(x,y) = f_X(x)\cdot f_Y(y)$. Рассмотрим двумерный случайный вектор $T = \mathrm{col}(U,V)$ с независимыми компонентами U и V, где СВ U и V имеют следующие плотности: $f_U(u) = f_X(u)$, $f_V(v) = f_Y(v)$. Тогда

$$f_T(u,v) = f_U(u)\cdot f_V(v)$$

во всех точках непрерывности этих функций. Поэтому $f_T(u,v) = f_1(u,v)$. Следовательно, $f_1(u,v)$ является плотностью распределения построенной СВ T.

Ответ. Функция $f_1(u,v)$ является плотностью вероятности двумерной СВ T.

Задача 9.2. Предприятие имеет две поточные линии по сборке некоторой продукции. Технологические процессы на линиях связаны между собой. Рассмотрим случайные величины: X — количество единиц продукции, собранной за день на первой линии, а Y — на второй линии. Совместное распределение этих величин задано табл. 9.2.

Таблица 9.2

$Y\backslash X$	0	1
0	1/8	0
1	1/4	1/8
2	1/8	3/8

Найти:

а) частные распределения случайных величин X и Y;

б) распределение, математическое ожидание и дисперсию случайной величины $Z = X + Y$ — суммарного количества единиц продукции, выпускаемой предприятием за день.

Решение. *а*) Согласно правилам построения частных распределений компонент случайного вектора, получим

$$\mathbf{P}\{X = x_i\} = \sum_{j=1}^{3} \mathbf{P}\{X = x_i, Y = y_j\}, \qquad i = \overline{1,\,2}. \tag{9.1}$$

Таблица 9.3

X	0	1
$\mathbf{P}$	1/2	1/2

Таблица 9.4

Y	0	1	2
$\mathbf{P}$	1/8	3/8	1/2

Таким образом, получаем ряд распределения СВ X, представленный в табл. 9.3. Аналогичным образом строим ряд распределения для случайной величины Y (см. табл. 9.4).

б) Случайная величина Z может принимать целые значения от 0 до 3. Определим вероятности, с которыми Z принимает эти значения:

$$\mathbf{P}\{Z=0\} = \mathbf{P}\{X=0, Y=0\} = \frac{1}{8},$$

$$\mathbf{P}\{Z=1\} = \mathbf{P}\Big(\{X=1, Y=0\} + \{X=0, Y=1\}\Big).$$

Так как события $\{X=1, Y=0\}$ и $\{X=0, Y=1\}$ несовместны, то согласно формуле сложения вероятностей имеем

$$\mathbf{P}\{Z=1\} = \mathbf{P}\{X=1, Y=0\} + \mathbf{P}\{X=0, Y=1\} = \frac{1}{4}.$$

Аналогично, для других возможных значений Z

$$\mathbf{P}\{Z=2\} = \mathbf{P}\{\{X=0, Y=2\} + \{X=1, Y=1\}\} = \\ = \mathbf{P}\{X=0, Y=1) + \mathbf{P}\{X=1, Y=1\} = \frac{1}{8} + \frac{1}{8} = \frac{1}{4},$$

$$\mathbf{P}\{Z=3\} = \mathbf{P}\{X=1, Y=2\} = \frac{3}{8}.$$

Таким образом, ряд распределения случайной величины Z может быть представлен табл. 9.5.

Таблица 9.5

Z	0	1	2	3
P	1/8	1/4	1/4	3/8

Найдем математическое ожидание и дисперсию СВ Z:

$$\mathbf{M}[Z] = \sum_{i=1}^{4} z_i p_i = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{3}{8} = \frac{15}{8},$$

$$\mathbf{D}[Z] = \sum_{i=1}^{4} (z_i - \mathbf{M}[Z])^2 p_i = \\ = \left(-\frac{15}{8}\right)^2 \frac{1}{8} + \left(1 - \frac{15}{8}\right)^2 \frac{1}{4} + \left(2 - \frac{15}{8}\right)^2 \frac{1}{4} + \left(3 - \frac{15}{8}\right)^2 \frac{3}{8} = \frac{71}{64}.$$

Ответ. *а*) Ряд распределения СВ X имеет вид табл. 9.3, а ряд распределения СВ Y имеет вид табл. 9.4; *б*) ряд распределения СВ Z имеет вид табл. 9.5, $\mathbf{M}[Z] = \frac{15}{8}$, $\mathbf{D}[Z] = \frac{71}{64}$.

Задача 9.3. Пусть двухмесячные объемы продаж продукции некоторого предприятия удовлетворительно описываются двумерным случайным вектором $Z = \mathrm{col}(X, Y)$ с плотностью распределения вероятности

$$f(x,y) = \begin{cases} c, & 100 \leqslant x \leqslant 150,\ 50 \leqslant y \leqslant 100, \\ 0 — & \text{в остальных случаях.} \end{cases}$$

Найти:

а) константу c;

б) функцию распределения $F(x,y)$;

в) исследовать случайные величины X и Y на независимость.

Решение. *а*) Для нахождения константы c воспользуемся условием нормировки

$$\int_{-\infty}^{+\infty}\Big(\int_{-\infty}^{+\infty} f(x,y)\,dx\Big)\,dy = 1.$$

Вычисляя интеграл, получим

$$\int_{-\infty}^{+\infty}\Big(\int_{-\infty}^{+\infty} f(x,y)\,dx\Big)\,dy = \int_{50}^{100}\Big(\int_{100}^{150} c\,dx\Big)\,dy = \int_{50}^{100} 50c\,dy = 2500c.$$

Следовательно, $c = \dfrac{1}{2500}$.

б) По определению

$$F(x,y) = \int_{-\infty}^{y}\Big(\int_{-\infty}^{x} f(t,\tau)\,dt\Big)\,d\tau.$$

В нашем случае при $x \leqslant 100$ или $y \leqslant 50$ имеем $f(x,y) = 0$, а следовательно, и $F(x,y) = 0$. Если $x \in [100,150]$ и $y \in [50,100]$, то, подставив найденную плотность вероятности в выражение для $F(x,y)$, получим

$$\int_{50}^{y}\Big(\int_{100}^{x} \frac{1}{2500}\,d\tau\Big)\,dt = \int_{50}^{y} \frac{1}{2500}\,(x-100)\,d\tau = \frac{1}{2500}\,(x-100)(y-50).$$

Если $x \in [50,150]$, а $y \geqslant 100$, то

$$F(x,y) = \int_{100}^{x}\Big(\int_{50}^{y} f(t,\tau)\,d\tau\Big)\,dt =$$
$$= \int_{100}^{x}\Big(\int_{50}^{100} \frac{1}{2500}\,d\tau\Big)\,dt + \int_{100}^{x}\Big(\int_{100}^{y} 0\,d\tau\Big)\,dt = \frac{1}{50}(x-100).$$

Если $x \geqslant 150$, а $y \in [50,100]$, имеем

$$F(x,y) = \int_{100}^{x}\Big(\int_{50}^{y} f(t,\tau)\,d\tau\Big)\,dt =$$
$$= \int_{100}^{150}\Big(\int_{50}^{y} \frac{1}{2500}\,d\tau\Big)\,dt + \int_{150}^{x}\Big(\int_{50}^{y} 0\,d\tau\Big)\,dt = (y-50)\frac{1}{50}.$$

При $x \geqslant 150, y \geqslant 100$

$$F(x,y) = \int\limits_{100}^{150}\left(\int\limits_{50}^{100} \frac{1}{2500}\, d\tau\right) dt + \int\limits_{150}^{x}\left(\int\limits_{100}^{y} 0\, d\tau\right) dt = 1.$$

Таким образом,

$$F(x,y) = \begin{cases} 0, & x \leqslant 100 \quad \text{или} \quad y \leqslant 50, \\ (x-100)(y-50)/2500, & 100 \leqslant x \leqslant 150, 50 \leqslant y \leqslant 100, \\ (x-100)/50, & 100 \leqslant x \leqslant 150, y \geqslant 100, \\ (y-50)/50, & 50 \leqslant y \leqslant 100, x \geqslant 150, \\ 1, & 150 \geqslant x, 100 \geqslant y. \end{cases} \tag{9.2}$$

в) Чтобы исследовать компоненты случайного вектора Z на независимость, найдем их частные распределения:

$$F_X(x) = F(x, +\infty) = \begin{cases} 0, & x \leqslant 100, \\ (x-100)/50, & 100 \leqslant x \leqslant 150, \\ 1, & x \geqslant 150, \end{cases}$$

$$F_Y(y) = F(+\infty, y) = \begin{cases} 0, & y \leqslant 50, \\ (y-50)/50, & 50 \leqslant y \leqslant 100, \\ 1, & y \geqslant 100. \end{cases}$$

Очевидно, что $F(x,y) = F_X(x) \cdot F_Y(y)$ для всех $x, y \in \mathbb{R}^1$. Следовательно, СВ X и Y независимы.

Ответ. *а*) $c = \dfrac{1}{2500}$; *б*) см. формулу (9.2); *в*) СВ X и Y независимы.

Задача 9.4. Случайные величины X и Y независимы. Случайная величина X распределена по нормальному закону с параметрами $m_X = 0$, $\sigma_X^2 = 1/2$. Случайная величина Y равномерно распределена на отрезке $[0,1]$. Требуется найти:

а) плотность вероятности случайного вектора $Z = \operatorname{col}(X,Y)$;

б) функцию распределения случайного вектора Z;

в) вероятность попадания СВ Z в область $D = \{-1 \leqslant X \leqslant 1,\ |Y| \leqslant 0,5\}$.

Решение. *а*) Плотности вероятностей СВ X и Y выглядят следующим образом:

$$f_X(x) = \frac{1}{\sqrt{\pi}}\, e^{-x^2}, \qquad f_Y(y) = \begin{cases} 1, & y \in [0,1], \\ 0, & y \notin [0,1]. \end{cases}$$

Так как СВ X и Y независимы, то плотность вероятности случайного вектора Z представляется в виде $f_Z(x,y) = f_X(x) f_Y(y)$. В данном случае

$$f_Z(x,y)=\begin{cases} \dfrac{1}{\sqrt{\pi}}\,e^{-x^2}, & y\in[0,1],\\ 0, & y\notin[0,1]. \end{cases} \tag{9.3}$$

б) Найдем функцию распределения. Двумерная функция распределения независимых случайных величин, по определению, имеет вид: $F_Z(x,y)=F_X(x)F_Y(y)$. Выпишем одномерные функции распределения СВ X и Y

$$F_X(x)=\int\limits_{-\infty}^{x}\frac{1}{\sqrt{\pi}}\,e^{-t^2}dt,\qquad F_Y(y)=\begin{cases} 0, & y\leqslant 0,\\ y, & 0<y<1,\\ 1, & y\geqslant 1.\end{cases}$$

Тогда

$$F_Z(x,y)=\begin{cases} 0, & y\leqslant 0,\\ y\displaystyle\int\limits_{-\infty}^{x}\frac{1}{\sqrt{\pi}}\,e^{-t^2}dt, & 0<y<1,\\ \displaystyle\int\limits_{-\infty}^{x}\frac{1}{\sqrt{\pi}}\,e^{-t^2}dt, & y\geqslant 1. \end{cases} \tag{9.4}$$

Пользуясь представлением гауссовской функции распределения через функцию Лапласа, можно записать выражение (9.4) следующим образом:

$$F_Z(x,y)=\begin{cases} 0, & y\leqslant 0,\\ y\left(\Phi_0(\sqrt{2}x)+\dfrac{1}{2}\right), & 0<y<1,\\ \Phi_0(\sqrt{2}x)+\dfrac{1}{2}, & y\geqslant 1. \end{cases} \tag{9.5}$$

в) Область $D=\{-1\leqslant x\leqslant 1,\ |y|\leqslant 0,5\}$ представляет собой прямоугольник. Вероятность попадания в него вычисляется следующим образом:

$$\mathbf{P}\{Z\in D\}=\int\limits_{-1}^{1}\Bigl(\int\limits_{-0,5}^{0,5} f_Z(x,y)\,dy\Bigr)dx=\int\limits_{-1}^{1}\frac{1}{\sqrt{\pi}}\,e^{-x^2}\,dx\int\limits_{0}^{0,5}1\,dy=$$

$$=\frac{1}{2}\cdot 2\Phi_0\left(\sqrt{2}\cdot 1\right)=\Phi_0\left(\sqrt{2}\right)\approx 0{,}42.$$

Вероятность попадания СВ Z в прямоугольник может быть найдена и другим способом:

$$\mathbf{P}\{-1 \leqslant X \leqslant 1, -0{,}5 \leqslant Y \leqslant\ 0{,}5\} = \mathbf{P}\{X \leqslant 1, Y \leqslant 0{,}5\} -$$

$$-\mathbf{P}\{X \leqslant -1, Y \leqslant 0{,}5\} - \mathbf{P}\{X \leqslant 1, Y \leqslant -0{,}5\} + \mathbf{P}\{X \leqslant -1, Y \leqslant -0{,}5\} =$$

$$= F_Z(1, 0{,}5) - F_Z(-1, 0{,}5) - F_Z(1, -0{,}5) + F_Z(-1, -0{,}5) =$$

$$= \frac{1}{2}\left(\Phi_0(1 \cdot \sqrt{2}) + 0{,}5\right) - \frac{1}{2}\left(0{,}5 - \Phi_0(1 \cdot \sqrt{2})\right) = \Phi_0\left(\sqrt{2}\right) \approx 0{,}42.$$

Ответ. *а*) Плотность распределения СВ Z имеет вид формулы (9.3); *б*) функция распределения СВ Z имеет вид выражения (9.5); *в*) $\mathbf{P}\{Z \in D\} \approx 0{,}42$.

§ 10. Условные распределения

10.1. Условная функция распределения. Рассмотрим следующую задачу: найти закон распределения двумерной СВ $Z \triangleq \operatorname{col}(X, Y)$, зная законы распределения двух одномерных СВ X и Y. Оказывается, что в общем случае эту задачу решить нельзя, так как необходимо знать связь между СВ X и Y, которая характеризуется условным распределением. Заметим, что свойства 6) $f(x, y)$ и 4) $F(x, y)$ позволяют решить обратную задачу: найти законы распределения одномерных СВ X и Y, зная закон распределения двумерной СВ $Z = \operatorname{col}(X, Y)$.

Определение 10.1. Пусть СВ Y является дискретной с реализациями y_j, $j = \overline{0, m}$. *Условной функцией распределения* СВ X при условии, что СВ $Y = y_j$, называется условная вероятность

$$F_X(x|y_j) \triangleq \frac{\mathbf{P}\{X \leqslant x, Y = y_j\}}{\mathbf{P}\{Y = y_j\}}, \quad j = \overline{0, m},\ x \in \mathbb{R}^1,$$

(см. определение 2.6), если $\mathbf{P}\{Y = y_j\} \neq 0$.

Рассмотрим двумерную СВ $Z = \operatorname{col}(X, Y)$ с плотностью распределения $f(x, y)$.

Определение 10.2. Пусть для всех $x \in \mathbb{R}^1$ плотности $f(x, y)$ и $f_Y(y)$ непрерывны в точке $y \in \mathbb{R}^1$ и $f_Y(y) \neq 0$. *Условной функцией распределения* СВ X при условии, что СВ $Y = y$, называется предел при $\Delta y > 0$ следующей условной вероятности:

$$F_X(x|y) \triangleq \lim_{\Delta y \to 0} \frac{\mathbf{P}\{X \leqslant x, |Y - y| \leqslant \Delta y\}}{\mathbf{P}\{|Y - y| \leqslant \Delta y\}}.$$

Аналогично определяется условная функция распределения $F_Y(y|x)$ СВ Y при условии, что СВ $X = x$.

Убедимся в корректности определения 10.2. Пусть СВ $Z \triangleq \triangleq \operatorname{col}(X, Y)$ непрерывна, причем плотности $f(x, y)$, $f_Y(y)$ непрерывны в точке y для всех $x \in \mathbb{R}^1$ и $f_Y(y) \neq 0$. Покажем, что в этом случае предел в определении 10.2 существует и он равен

$$F_X(x|y) = \frac{1}{f_Y(y)} \int\limits_{-\infty}^{x} f(t, y)\, dt.$$

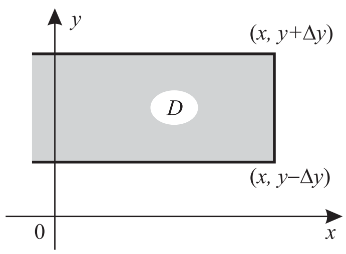

Рис. 10.1

Рассмотрим вероятность

$$\mathbf{P}(D) \triangleq \mathbf{P}\{X \leqslant x, |Y - y| \leqslant \Delta y\}$$

попадания СВ Z в полуполосу D (рис. 10.1). По свойству 17)Р получаем

$$\mathbf{P}(D) = \mathbf{P}\{|Y - y| \leqslant \Delta y\} \cdot \mathbf{P}\{X \leqslant x | y - \Delta y \leqslant Y \leqslant y + \Delta y\}.$$

Тогда, из свойств 2) $f(x, y)$, 3) $f_X(x)$ вытекает, что

$$\mathbf{P}\{X \leqslant x | y - \Delta y \leqslant Y \leqslant y + \Delta y\} = \frac{\mathbf{P}\{X \leqslant x, |Y - y| \leqslant \Delta y\}}{\mathbf{P}\{|Y - y| \leqslant \Delta y\}} =$$

$$= \frac{\int\limits_{-\infty}^{x} \left(\int\limits_{y-\Delta y}^{y+\Delta y} f(t, \tau)\, d\tau \right) dt}{\int\limits_{y-\Delta y}^{y+\Delta y} f_Y(\tau)\, d\tau} = \left\| \begin{array}{l} \text{по теореме} \\ \text{о среднем} \\ \text{значении} \end{array} \right\| = \frac{1}{f_Y(\approx{y})} \int\limits_{-\infty}^{x} f(t, \widetilde{y})\, dt,$$

где $\widetilde{y}$ и $\overset{\approx}{y}$ некоторые точки из интервала $(y - \Delta y, y + \Delta y)$. Учитывая непрерывность $f(x, y)$ и $f(y)$ и переходя к пределу при $\Delta y \to 0$, имеем

$$\widetilde{y} \to y, \quad \overset{\approx}{y} \to y, \quad \mathbf{P}\{X \leqslant x | y - \Delta y \leqslant Y \leqslant y + \Delta y\} \to F_X(x|y).$$

Отсюда следует требуемая формула, которая определяет условную функцию распределения через плотности распределения двумерной СВ и одномерной СВ.

Свойства $F(x\,|y)$

1) $F_X(x|y)$ определена для всех $x \in \mathbb{R}^1$. Это следует из определения $F_X(x|y)$.

2) $0 \leqslant F_X(x|y) \leqslant 1$ для всех $x \in \mathbb{R}^1$. Это объясняется тем, что условная вероятность принимает значения из отрезка $[0,1]$, а значит, и ее предел при $\Delta y \to 0$ также лежит в данном отрезке.

3) $F_X(-\infty|y) = 0$ по определению 10.2.

4) $F_X(+\infty|y) = 1$, так как, например, из определения 10.2 вытекает

$$F_X(+\infty|y) = \frac{\int\limits_{-\infty}^{+\infty} f(t,y)\,dt}{f_Y(y)} = \frac{f_Y(y)}{f_Y(y)} = 1.$$

5) $F(x|y)$ — функция монотонно неубывающая по $x \in \mathbb{R}^1$ для любого фиксированного y. Например, в силу определения 10.2 имеем

$$F_X(x+\Delta x|y) = \frac{\int\limits_{-\infty}^{x+\Delta x} f(t,y)\,dt}{f_Y(y)} \geqslant \frac{\int\limits_{-\infty}^{x} f(t,y)\,dt}{f_Y(y)} = F_X(x|y).$$

Аналогичные свойства имеет условная функция распределения $F_Y(y|x)$ СВ Y при условии, что СВ $X = x$.

10.2. Условная плотность распределения.

Определение 10.3. *Условной плотностью распределения* (*вероятности*) непрерывной СВ X при условии, что непрерывная СВ Y с плотностью $f_Y(y) \neq 0$ приняла значение y, называется функция

$$f(x|y) = \frac{f(x,y)}{f_Y(y)}, \qquad x \in \mathbb{R}^1.$$

Аналогично определяется условная плотность вероятности $f(y|x)$ СВ Y при условии, что $X = x$.

Для построения двумерной плотности распределения в общем случае требуется знание не только плотностей распределения одномерных СВ, но еще и условных плотностей, например $f(x,y) = f_Y(y)f_X(x|y)$.

Свойства $f(x\,|y)$

1) $f_X(x|y) \geqslant 0$, так как $f(x,y) \geqslant 0,\ \ f_Y(y) \geqslant 0$.

2) $F_X(x|y) \triangleq \int\limits_{-\infty}^{x} f_X(t|y)\,dt$, если плотности $f(x,y)$, $f_Y(y)$ непрерывны. В этом можно убедиться, сравнивая выражения для условной плотности и условной функции распределения.

3) $\int\limits_{-\infty}^{+\infty} f_X(t|y)\,dt = 1$ по свойствам 2) $f(x|y)$, 4) $F(x|y)$.

4) $f_X(x|y) = \dfrac{f(x,y)}{f_Y(y)}$, если плотности $f(x,y)$, $f_Y(y)$ непрерывны. Действительно, имеем

$$f_X(x|y) \triangleq \frac{\partial F_X(x|y)}{\partial x} = \frac{1}{f_Y(y)}\,\frac{\partial}{\partial x}\Big(\int\limits_{-\infty}^{x} f(t,y)\,dt\Big) = \frac{f(x,y)}{f_Y(y)}.$$

5) Если непрерывные СВ X и Y независимы, то $f_X(x|y) = f_X(x)$. Согласно свойству 8) $f(x,y)$ для независимых СВ X и Y имеем равенство $f(x,y) = f_X(x)f_Y(y)$, из которого следует по определению условной плотности, что $f_X(x|y) = f_X(x)$.

6) $\mathbf{P}\{x_1 \leqslant X \leqslant x_2\} = \int\limits_{-\infty}^{\infty} f_Y(y)\Big(\int\limits_{x_1}^{x_2} f_X(x|y)\,dx\Big)\,dy.$

Действительно, по свойствам 3) $f(x)$, 6) $f(x,y)$ и определению условной плотности имеем

$$\mathbf{P}\{x_1 \leqslant X \leqslant x_2\} = \int\limits_{x_1}^{x_2} f_X(x)\,dx = \int\limits_{x_1}^{x_2}\Big(\int\limits_{-\infty}^{\infty} f(x,y)\,dy\Big)\,dx =$$
$$= \int\limits_{x_1}^{x_2}\Big(\int\limits_{-\infty}^{\infty} f_Y(y)f_X(x|y)\,dy\Big)\,dx = \int\limits_{-\infty}^{\infty} f_Y(y)\Big(\int\limits_{x_1}^{x_2} f_X(x|y)\,dx\Big)\,dy.$$

Последнее свойство аналогично формуле полной вероятности в случае дискретных СВ X и Y (см. теорему 3.3):

$$\mathbf{P}\{x_1 \leqslant X \leqslant x_2\} = \sum_{j=0}^{m} \mathbf{P}\{Y = y_j\}\mathbf{P}\{x_1 \leqslant X \leqslant x_2 | Y = y_j\}.$$

Такие же свойства имеет условная плотность распределения $f_Y(y|x)$ СВ Y при условии, что $X = x$.

Пример 10.1. Пусть имеется СВ $Z = \mathrm{col}(X,Y)$, где X — время появления первого покупателя в понедельник, а Y — время появления первого покупателя во вторник. Положим, что $f(x,y) = e^{-x-y}$, $x, y \geqslant 0$. Требуется найти:

1) $F(x,y)$; 2) $F_X(x)$, $F_Y(y)$; 3) $f_X(x)$, $f_Y(y)$; 4) $F_X(x|y)$, $F_Y(y|x)$; 5) $f_X(x|y)$, $f_Y(y|x)$.

Установить, зависимы ли СВ X и Y.

Решение.

1) $F(x,y) = \int\limits_{0}^{y}\Big(\int\limits_{0}^{x} f(t,\tau)\,dt\Big)\,d\tau = \int\limits_{0}^{y} e^{-\tau}\,d\tau \int\limits_{0}^{x} e^{-t}\,dt = (1-e^{-x})(1-e^{-y})$

при $x \geqslant 0,\ y \geqslant 0;\ F(x,y) = 0$ в остальных случаях.
2) $F_X(x) = F(x,\infty) = 1 - e^{-x}$ при $x > 0,\ F_Y(y) = 1 - e^{-y}$ при $y > 0$.
3) $f_X(x) = \dfrac{dF_X(x)}{dx} = e^{-x}$ при $x > 0,\ f_Y(y) = \dfrac{dF_Y(y)}{dy} = e^{-y}$ при $y > 0$.
4) $F_X(x|y) = \dfrac{1}{f_Y(y)} \int\limits_0^x f(t,y)\,dt = e^y \int\limits_0^x e^{-t}e^{-y}\,dt = 1 - e^{-x}$ при $x > 0$.
Аналогично, $F_Y(y|x) = 1 - e^{-y}$ при $y > 0$.
5) $f_X(x|y) = \dfrac{\partial}{\partial x} F_X(x|y) = e^{-x}$ при $x > 0,\ f_Y(y|x) = e^{-y}$ при $y > 0$.

Так как $f_X(x|y) = f_X(x) = e^{-x}$ и $f_Y(y|x) = f_Y(y) = e^{-y}$, то $f(x,y) = f_X(x)f_Y(y)$. Поэтому СВ X и Y независимы. Это означает, что появление первого покупателя во вторник не зависит от того, когда пришел первый покупатель в понедельник.

10.3. Условное математическое ожидание.

Определение 10.4. *Условным математическим ожиданием* непрерывной СВ X при условии, что непрерывная СВ Y приняла значение y, называется функция

$$\mathbf{M}[X|y] \triangleq \int\limits_{-\infty}^{+\infty} x f_X(x|y)\,dx.$$

Условное МО существует, если последний интеграл сходится абсолютно.

В случае дискретных СВ X и Y условное МО СВ X при условии, что $Y = y_j$, $j = \overline{0,m}$, определяется формулой

$$\mathbf{M}[X|y_j] \triangleq \sum_{i=0}^{n} x_i \frac{p_{ij}}{p_j}, \quad j = \overline{0,m},$$

где $p_{ij} \triangleq \mathbf{P}\{X = x_i, Y = y_j\}$, $p_j \triangleq \mathbf{P}\{Y = y_j\}$.

Определение 10.5. Условное математическое ожидание $\mathbf{M}[X|y]$ СВ X как функция параметра $y \in \mathbb{R}^1$ называется *регрессией X на y*. График функции $x = \mathbf{M}[X|y]$ называется *кривой регрессии X на y*.

Аналогично определяется условное МО СВ $V \triangleq \psi(X)$ при условии, что $Y = y$. Например, для непрерывных X и Y:

$$\mathbf{M}[\psi(X)|y] \triangleq \int\limits_{-\infty}^{+\infty} \psi(x) f_X(x|y)\,dx.$$

Рассмотрим функцию $\varphi(y) \triangleq \mathbf{M}[X|y]$.

Определение 10.6. *Условным математическим ожиданием* СВ X относительно СВ Y называется СВ $V \triangleq \varphi(Y) \triangleq \mathbf{M}[X|Y]$.

Аналогично можно определить и другие, более высокого порядка, условные моменты СВ.

Свойства $\mathbf{M}[X\,|y]$

1) $\mathbf{M}[\psi(Y)|y] = \psi(y)$, где $\psi(y)$ — некоторая функция.

2) $\mathbf{M}[\psi(Y)X|y] = \psi(y)\mathbf{M}[X|y]$. Действительно, например, в случае непрерывных СВ X и Y имеем

$$\mathbf{M}[\psi(Y)X|y] = \int\limits_{-\infty}^{+\infty} \psi(y)xf_X(x|y)\,dx = \psi(y)\mathbf{M}[X|y].$$

3) $\mathbf{M}[X + \psi(Y)|y] = \mathbf{M}[X|y] + \psi(y)$. Это свойство доказывается аналогично свойству 2)M$[X|y]$.

4) $\mathbf{M}[X|y] = \mathbf{M}[X]$, если X и Y — независимы. Пусть, например, СВ X и Y — непрерывны, тогда по свойству 5) $f(x|y)$

$$\mathbf{M}[X|y] \triangleq \int\limits_{-\infty}^{+\infty} xf_X(x|y)\,dx = \int\limits_{-\infty}^{+\infty} xf(x)\,dx \triangleq \mathbf{M}[X].$$

5) $\mathbf{M}[X] = \mathbf{M}\big[\mathbf{M}[X|Y]\big]$ — это равенство называется *формулой полного математического ожидания*. Пусть СВ X и Y непрерывны, тогда по свойству 6) $f(x,y)$

$$\mathbf{M}[X] \triangleq \int\limits_{-\infty}^{+\infty} xf_X(x)\,dx = \int\limits_{-\infty}^{+\infty} x\Big(\int\limits_{-\infty}^{+\infty} f(x,y)dy\Big)\,dx =$$
$$= \int\limits_{-\infty}^{+\infty} x\Big(\int\limits_{-\infty}^{+\infty} f_Y(y)f_X(x|y)\,dy\Big)\,dx = \int\limits_{-\infty}^{+\infty} f_Y(y)\Big(\int\limits_{-\infty}^{+\infty} xf_X(x|y)\,dx\Big)\,dy =$$
$$= \int\limits_{-\infty}^{+\infty} f_Y(y)\mathbf{M}[X|y]\,dy = \mathbf{M}\left[\mathbf{M}[X|Y]\right].$$

В случае дискретной СВ Y при конечном m формула полного МО приобретает следующий вид:

$$\mathbf{M}[X] = \sum_{j=0}^{m} p_j\mathbf{M}[X|y_j], \quad \text{где } p_j \triangleq \mathbf{P}\{Y = y_j\},\ j = \overline{0,m}.$$

Пример 10.2. Число N радиотехнических приборов, сдаваемых покупателями в гарантийную мастерскую в течении дня, можно представить в виде СВ, хорошо описываемой распределением Пуассона $\mathbf{\Pi}(a)$, где a является средним числом радиоприборов, сданных за день. Вероятность того, что сданный прибор потребует длительного ремонта, равна p. Найдем среднее число сданных приборов, требующих длительного ремонта. Так как при фиксированном числе n поступивших приборов количество приборов, требующих капитального ремонта, представляет собой СВ X с биномиальным распределением $\mathbf{Bi}(n;p)$, то $\mathbf{M}[X|n] = np$, $n = 0, 1, 2, \ldots$ Поскольку СВ X имеет распределение Пуассона $\mathbf{\Pi}(a)$, то $\mathbf{M}[N] = a$. Тогда по формуле полного МО

$$\mathbf{M}[X] = \mathbf{M}\left[\mathbf{M}[X|N]\right] = \mathbf{M}[pN] = p\,\mathbf{M}[N] = p\,a.$$

10.4. Корреляционная зависимость.

Определение 10.7. *Ковариацией* (*корреляционным моментом*) k_{XY} непрерывных СВ X и Y называется второй центральный смешанный момент

$$k_{XY} \triangleq \mathbf{M}[(X-m_X)(Y-m_Y)] = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(x-m_X)(y-m_Y)f(x,y)\,dx\,dy.$$

Корреляционный момент k_{XY} дискретных СВ X и Y с конечными числами возможных значений равен

$$k_{XY} \triangleq \sum_{i=0}^{n}\sum_{j=0}^{n}(x_i-m_X)(y_j-m_Y)p_{ij}.$$

Величина k_{XY} зависит от единиц измерения СВ X и Y. Во избежание этого неудобства ковариацию часто вычисляют не для X и Y, а для соответствующих им *нормированных* СВ $\overset{*}{X} \triangleq \dfrac{X-m_X}{\sigma_X}$, $\overset{*}{Y} \triangleq \dfrac{Y-m_Y}{\sigma_Y}$, если у СВ X и Y существуют дисперсии $\sigma_X^2 > 0$ и $\sigma_Y^2 > 0$.

Определение 10.8. Ковариация нормированных СВ $\overset{*}{X}$ и $\overset{*}{Y}$ называется *коэффициентом корреляции*:

$$r_{XY} \triangleq k_{\overset{*}{X}\overset{*}{Y}} = \mathbf{M}[\overset{*}{X}\overset{*}{Y}] = \frac{\mathbf{M}[(X-m_X)(Y-m_Y)]}{\sigma_X\sigma_Y} \triangleq \frac{k_{XY}}{\sigma_X\sigma_Y}.$$

Определение 10.9. СВ X и Y называют *коррелированными* (*корреляционно зависимыми*), если $r_{XY} \neq 0$ ($k_{XY} \neq 0$), и *некоррелированными*, если $r_{XY} = 0$ ($k_{XY} = 0$).

Определение 10.10. СВ X и Y называют *положительно коррелированными*, если если $r_{XY} > 0$, и *отрицательно коррелированными*, если $r_{XY} < 0$.

Пример 10.3. Рассмотрим две СВ X и $Y \triangleq X^2$, где $X \sim \mathbf{R}(-a; a)$. Тогда ковариация между X и Y равна

$$k_{XY} = \mathbf{M}[(X - m_X)(Y - m_Y)] = \mathbf{M}[X(Y - m_Y)] = \\ = \mathbf{M}[X(X^2 - m_Y)] = \frac{1}{2a}\int_{-a}^{a} x(x^2 - m_Y)\,dx = 0.$$

Таким образом, СВ X и Y некоррелированы. В то же время СВ X и Y связаны самой «жесткой» зависимостью — функциональной.

Свойства k_{XY}

1) $|r_{XY}| \leqslant 1$, т.е. $|k_{XY}| \leqslant \sigma_X\sigma_Y$. Это свойство доказано ниже для более общего случая (см. свойство 5)M$[X]$ в § 11).

2) Если $Y = aX + b$, где $a \neq 0$ и b — постоянные и $0 < \sigma_X < +\infty$, то $|r_{XY}| = 1$. Действительно,

$$m_Y \triangleq \mathbf{M}[Y] = \mathbf{M}[aX + b] = a\mathbf{M}[X] + b = am_X + b,$$

$$\mathbf{D}[Y] \triangleq \mathbf{M}[(Y - m_Y)^2] = a^2 d_X, \quad \sigma_Y = \sqrt{d_Y} = |a|\sqrt{d_X} = |a|\sigma_X,$$

$$k_{XY} \triangleq \mathbf{M}[(X - m_X)(Y - m_Y)] = \mathbf{M}[a(X - m_X)^2] \triangleq a\mathbf{D}[X] = a\sigma_X^2,$$

$$r_{XY} \triangleq \frac{k_{XY}}{\sigma_X\sigma_Y} = \frac{a\sigma_X^2}{\sigma_X^2|a|} = \frac{a}{|a|}, \quad \text{т.е.} \quad |r_{XY}| = 1.$$

Следовательно, если $a > 0$, то $r_{XY} = 1$ и, если $a < 0$, то $r_{XY} = -1$.

3) Из независимости СВ X и Y следует их некоррелированность. Пусть, например, СВ X и Y непрерывны, тогда по свойствам 8) $f(x,y)$ и 5) m_X имеем

$$k_{XY} \triangleq \mathbf{M}[(X - m_X)(Y - m_Y)] = \iint_{-\infty}^{+\infty} (x - m_X)(y - m_Y) f(x,y)\,dx\,dy = \\ = \int_{-\infty}^{\infty} (x - m_X) f_X(x)\,dx \int_{-\infty}^{\infty} (y - m_Y) f_Y(y)\,dy \triangleq \mathbf{M}[\overset{\circ}{X}]\mathbf{M}[\overset{\circ}{Y}] = 0.$$

4) $k_{XY} = \mathbf{M}[XY] - m_X m_Y$. Пусть, например, СВ $Z \triangleq (X, Y)$ непрерывна. Тогда, согласно свойству 7) $f(x,y)$, имеем

$$k_{XY} \triangleq \mathbf{M}[(X - m_X)(Y - m_Y)] = \\ = \mathbf{M}[XY] + m_X m_Y - 2m_X m_Y = \mathbf{M}[XY] - m_X m_Y.$$

10.5. Двумерное нормальное распределение.

Определение 10.11. Плотность *двумерной нормально распределенной* СВ Z с параметрами c_{11}, c_{22}, c_{12} и m_1, m_2, где $c_{11}c_{22} - c_{12}^2 > 0$, $c_{11} > 0$, $c_{22} > 0$, определяется формулой

$$f(x,y) = \frac{\sqrt{c_{11}c_{22} - c_{12}^2}}{2\pi}\, e^{-\frac{1}{2}[c_{11}(x-m_1)^2 + 2c_{12}(x-m_1)(y-m_2) + c_{22}(y-m_2)^2]}.$$

Свойства двумерного нормального распределения

1) В соответствии со свойством 6) $f(x,y)$ найдем плотность СВ X

$$f_X(x) = \sqrt{\frac{c_{11}c_{22} - c_{12}^2}{2\pi c_{22}}} \exp\left\{-\frac{1}{2}\,\frac{c_{11}c_{22} - c_{12}^2}{c_{22}}\,(x - m_1)^2\right\}.$$

Сравнивая эту формулу с определением 7.3, заключаем, что СВ X имеет нормальное распределение $\mathbf{N}(m_X; \sigma_X^2)$ с параметрами

$$m_X = m_1, \qquad \mathbf{D}[X] \triangleq \sigma_X^2 = \frac{c_{22}}{c_{11}c_{22} - c_{12}^2}.$$

По аналогии можно получить плотность СВ Y:

$$f_Y(y) = \sqrt{\frac{c_{11}c_{22} - c_{12}^2}{2\pi c_{11}}} \exp\left\{-\frac{1}{2}\,\frac{c_{11}c_{22} - c_{12}^2}{c_{11}}\,(y - m_2)^2\right\},$$

т. е. $Y \sim \mathbf{N}(m_Y; \sigma_Y^2)$, где $m_Y = m_2$, $\mathbf{D}[Y] \triangleq \sigma_Y^2 = c_{11}/(c_{11}c_{22} - c_{12}^2)$.

2) Вычисляя коэффициент корреляции СВ X и Y, получаем

$$r_{XY} \triangleq \frac{1}{\sigma_X \sigma_Y} \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x - m_X)(y - m_Y) f(x,y)\, dx\, dy = -\frac{c_{12}}{\sqrt{c_{11}c_{22}}}.$$

3) Найдем условную плотность

$$f_X(x|y) \triangleq \frac{f(x,y)}{f_Y(y)} = \sqrt{\frac{c_{11}}{2\pi}} \exp\left\{-\frac{1}{2}\, c_{11}\left[x - m_1 + \frac{c_{12}}{c_{11}}\,(y - m_2)\right]^2\right\}.$$

Сравнивая эту формулу с определением 7.3, приходим к выводу, что $f_X(x|y)$ является нормальной плотностью, причем

$$\mathbf{M}[X|y] = m_1 - \frac{c_{12}}{c_{11}}\,(y - m_2), \ \mathbf{D}[X|y] = \frac{1}{c_{11}}.$$

Используя выражения, полученные выше для σ_X, σ_Y, r_{XY}, условное МО $\mathbf{M}[X|y]$ можно представить в виде соотношения

$$\mathbf{M}[X|y] = \mathbf{M}[X] + r_{XY}\,\frac{\sigma_X}{\sigma_Y}\,(y - \mathbf{M}[Y]),$$

которое называется *теоремой о нормальной корреляции.*

4) Пусть СВ X и Y независимы. Тогда

$$f(x,y) = f_X(x) f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_X}\,e^{-\frac{(x-m_X)^2}{2}}\,\frac{1}{\sqrt{2\pi}\sigma_Y}\,e^{-\frac{(y-m_Y)^2}{2}}.$$

Пусть теперь СВ X и Y некоррелированы, т.е. $r_{XY} = 0$. Поэтому, $c_{12} = 0$ и тогда $\sigma_X^2 = 1/c_{11}$, $\sigma_Y^2 = 1/c_{22}$. Поэтому, $f_X(x|y) = f_X(x)$, откуда вытекает, что $f(x,y) = f_X(x) f_Y(y)$. Таким образом, в случае нормального распределения вектора Z из некоррелированности СВ X и Y следует их независимость. При этом $\mathbf{M}[X|y] = \mathbf{M}[X]$.

Пример 10.4. Двумерное нормальное распределение хорошо описывает, например, скорость ветра в районе аэропорта, прогнозируемые координаты падения метеорита на поверхность Земли и т.п.

10.6. Типовые задачи.

Задача 10.1. Пусть случайный вектор $Z = \mathrm{col}(X, Y)$ равномерно распределен в круге радиуса R:

$$f(x,y) = \begin{cases} 1/(\pi R^2), & x^2 + y^2 \leqslant R^2, \\ 0, & x^2 + y^2 > R^2. \end{cases}$$

Требуется найти r_{XY} и ответить на вопрос: зависимы ли СВ X и Y.

Решение. Из симметрии плотности вероятности следует $m_X = m_Y = 0$. Действительно, из свойства 6) $f(x,y)$ находим $f_X(x) = 0$, если $|x| > R$, а в случае $|x| \leqslant R$ получаем

$$f_X(x) = \int\limits_{-\infty}^{+\infty} f(x,y)\,dy = \frac{1}{\pi R^2}\int\limits_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy = \frac{2}{\pi R^2}\sqrt{R^2 - x^2}.$$

Исходя из этого,

$$m_X = \frac{2}{\pi R^2}\int\limits_{-R}^{R} x\sqrt{R^2 - x^2}\,dx = -\frac{2}{3\pi R^2}\,(R^2 - x^2)^{3/2}\Big|_{-R}^{R} = 0.$$

Аналогично можно показать, что $f_Y(y) = 0$, если $|y| > R$, и $f_Y(y) = \dfrac{2}{\pi R^2}\sqrt{R^2 - x^2}$, если $|y| \leqslant R$, поэтому $m_Y = 0$. Таким образом, имеем

$$k_{XY} = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf(x,y)\,dx\,dy = \frac{2}{\pi R^2}\int\limits_{-R}^{R} x\Bigg(\int\limits_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} y\,dy\Bigg)\,dx = 0,$$

т. е. $k_{XY} = r_{XY} = 0$. Следовательно, СВ X и Y некоррелированы. При этом СВ X и Y зависимы, так как плотность $f(x,y)$ нельзя представить в данном случае в виде произведения $f(x,y) = f_X(x)f_Y(y)$.

Ответ. $r_{XY} = 0$, но СВ X и Y зависимы.

Задача 10.2. Пусть дисперсии случайных величин X и Y равны между собой, $\mathbf{D}[X] = \mathbf{D}[Y]$. Чему равна ковариация случайных величин $X_1 = (X+Y)$ и $X_2 = (X-Y)$?

Решение. Вычислим $k_{X_1X_2}$. По определению $k_{X_1X_2} = \mathbf{M}[X_1X_2] - \mathbf{M}[X_1]\mathbf{M}[X_2]$, следовательно,

$$\begin{aligned} k_{X_1X_2} = \mathbf{M}[(X+Y)(X-Y)] - \mathbf{M}[X+Y]\mathbf{M}[X-Y] = \\ = \mathbf{M}[X^2 - XY + XY - Y^2] - \\ - \Big((\mathbf{M}[X])^2 - \mathbf{M}[X]\mathbf{M}[Y] + \mathbf{M}[X]\mathbf{M}[Y] - (\mathbf{M}[Y])^2\Big). \end{aligned}$$

Пользуясь свойствами математического ожидания, имеем

$$k_{X_1X_2} = \mathbf{M}[X^2] - (\mathbf{M}[X])^2 - (\mathbf{M}[Y^2] - (\mathbf{M}[Y])^2) = \mathbf{D}[X] - \mathbf{D}[Y] = 0.$$

Ответ. $k_{X_1X_2} = 0$.

Задача 10.3. Шесть футбольных команд, участвующих в турнире, разбиты на две группы по три команды. Победители групп выходят в финал. Эксперты оценивают вероятности появления финальных пар согласно данным табл. 10.1. Найти наиболее вероятных претендентов на выход в финал и коэффициент корреляции между претендентами на выход в финал из двух групп.

Таблица 10.1

$X\backslash Y$	1	2	3
1	2/9	1/9	0
2	1/9	0	1/9
3	2/9	1/9	1/9

Решение. Чтобы определить наиболее вероятных претендентов, найдем частные распределения компонент X (номер команды в первой подгруппе) и Y (номер команды в первой подгруппе). Для этого воспользуемся формулой (9.1). Ряды распределения случайных величин X и Y заданы соответственно в табл. 10.2 и 10.3.

Таблица 10.2

X	1	2	3
$\mathbf{P}$	1/3	2/9	4/9

Таблица 10.3

Y	1	2	3
$\mathbf{P}$	5/9	2/9	2/9

Таким образом, наиболее вероятным претендентом на выход в финал из первой подгруппы является третья команда, а из второй подгруппы — первая.

Для нахождения коэффициента корреляции найдем дисперсии случайных величин X и Y и их ковариацию:

$$\mathbf{M}[X] = \sum_{i=1}^{3} x_i p_i = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{9} + 3 \cdot \frac{4}{9} = \frac{19}{9},$$

$$\mathbf{M}[Y] = \sum_{i=1}^{3} y_i p_i = 1 \cdot \frac{5}{9} + 2 \cdot \frac{2}{9} + 3 \cdot \frac{2}{9} = \frac{5}{3},$$

$$\mathbf{D}[X] = \sum_{i=1}^{3} (x_i - \mathbf{M}[X])^2 p_i = \\ = \left(1 - \frac{19}{9}\right)^2 \cdot \frac{1}{3} + \left(2 - \frac{19}{9}\right)^2 \cdot \frac{2}{9} + \left(3 - \frac{19}{9}\right)^2 \cdot \frac{4}{9} \approx 0{,}765,$$

$$\mathbf{D}[Y] = \sum_{i=1}^{3} (y_i - \mathbf{M}[Y])^2 p_i = \\ = \left(1 - \frac{5}{3}\right)^2 \cdot \frac{5}{9} + \left(2 - \frac{5}{3}\right)^2 \cdot \frac{2}{9} + \left(3 - \frac{5}{3}\right)^2 \cdot \frac{2}{9} \approx 0{,}667,$$

$$\sigma_X = \sqrt{\mathbf{D}[X]} \approx 0{,}875,$$

$$\sigma_Y = \sqrt{\mathbf{D}[Y]} \approx 0{,}817,$$

$$k_{XY} = \sum_{j=1}^{3} \sum_{i=1}^{3} (x_i - \mathbf{M}[X])(y_j - \mathbf{M}[Y]) p_{ij} \approx 0{,}148.$$

Коэффициент корреляции r_{XY} вычисляется следующим образом:

$$r_{XY} = \frac{k_{XY}}{\sigma_X \sigma_Y} \approx 0{,}207.$$

Ответ. Наиболее вероятным претендентом на выход в финал из первой подгруппы является третья команда, а из второй подгруппы — первая. Искомый коэффициент корреляции $r_{XY} \approx 0{,}207$.

Задача 10.4. Пусть $Y_1 = aX_1 + b, Y_2 = cX_2 + d$. Доказать, что $k_{Y_1Y_2} = a\,c\,k_{X_1X_2}$, где X_1 и X_2 — некоторые случайные величины.

Решение. Для решения этой задачи воспользуемся определением:

$$\overset{\circ}{Y}_1 = aX_1 + b - \mathbf{M}[aX_1 + b] = a(X_1 - \mathbf{M}[X_1]) = a\overset{\circ}{X}_1,$$

$$\overset{\circ}{Y}_2 = aX_2 + b - \mathbf{M}[aX_2 + b] = a(X_2 - \mathbf{M}[X_2]) = a\overset{\circ}{X}_2\,.$$

Используя свойства математического ожидания, получим

$$k_{Y_1Y_2} = \mathbf{M}[\overset{\circ}{Y}_1\overset{\circ}{Y}_2] = \mathbf{M}[a\,c\,\overset{\circ}{X}_1\overset{\circ}{X}_2] = a\,c\,k_{X_1X_2},$$

что и требовалось доказать.

Задача 10.5. Игральная кость размечена таким образом, что сумма очков на противоположных гранях равна 7 (т.е. 1 и 6, 2 и 5, 3 и 4). Пусть X — число очков на верхней грани; Y — число очков на нижней грани. Построить совместный закон распределения случайных величин X и Y и найти коэффициент корреляции между ними.

Решение. Ясно, что выполняется равенство $X + Y = 7$, и поэтому $\mathbf{P}\{X + Y \neq 7\} = 0$. Следовательно, для построения таблицы распределения случайного вектора $Z = \mathrm{col}(X, Y)$ остается вычислить следующие вероятности:

$$\mathbf{P}\{X = 1, Y = 6\} = \mathbf{P}(\{X = 1\} \cdot \{Y = 6\}) = \\ = \mathbf{P}(\{X = 1\} \cdot \{7 - X = 6\}) = \mathbf{P}(\{X = 1\}) = 1/6,$$

$$\mathbf{P}\{X = 6, Y = 1\} = \mathbf{P}\{X = 6\} = 1/6.$$

Аналогично можно показать, что

$$\mathbf{P}\{X = 2, Y = 5\} = \mathbf{P}\{X = 5, Y = 2\} = \mathbf{P}\{X = 2\}) = 1/6,$$

$$\mathbf{P}\{X = 3, Y = 4\} = \mathbf{P}\{X = 4, Y = 3\} = \mathbf{P}\{X = 3\} = 1/6.$$

Таблица 10.4

$X \backslash Y$	1	2	3	4	5	6
1	0	0	0	0	0	1/6
2	0	0	0	0	1/6	0
3	0	0	0	1/6	0	0
4	0	0	1/6	0	0	0
5	0	1/6	0	0	0	0
6	1/6	0	0	0	0	0

Таблица распределения случайного вектора Z представлена в табл. 10.4.

Поскольку между случайными величинами X и Y имеется линейная связь $X = 7 - Y$, то, согласно свойству 2) k_{XY}, $r_{XY} = -1$.

Ответ. $r_{XY} = -1$.

Задача 10.6. Случайные величины X и Y независимы и распределены по нормальному закону. Известно, что $\mathbf{M}[X] = a$, $\mathbf{M}[Y] = b$, $\mathbf{D}[X] = \mathbf{D}[Y] = \sigma^2$. Найти радиус R круга с центром в точке (a, b), вероятность попадания в который случайной точки (X, Y) равна $0{,}997$.

Решение. Обозначим $Z \triangleq \mathrm{col}(X, Y)$. Поскольку СВ X и Y независимы, то

$$f_Z(x,y) = f_X(x) f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-a)^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(y-b)^2}{2\sigma^2}}.$$

В этом случае искомая вероятность $\mathbf{P}(D)$, где $D = B_R(a,b)$ — круг радиуса R с центром в (a, b), вычисляется согласно свойству 3) $f(x,y)$:

$$\mathbf{P}(B_R(a,b)) = \iint\limits_{B_R(a,b)} f_Z(t,\tau)\, dt\, d\tau = \iint\limits_{(x-a)^2+(y-b)^2 \leqslant R^2} f_X(t) f_Y(\tau)\, dt\, d\tau =$$

$$= \frac{1}{\sigma^2}\frac{1}{2\pi} \iint\limits_{(x-a)^2+(y-b)^2 \leqslant R^2} \exp\left\{-\frac{1}{2\sigma^2}\left((t-a)^2 + (\tau-b)^2\right)\right\} dt\, d\tau =$$

$$= \left\| \begin{array}{l} u = \dfrac{t-a}{\sigma}, \\ v = \dfrac{\tau - b}{\sigma} \end{array} \right\| = \frac{1}{2\pi} \iint\limits_{\widetilde{x}^2+\widetilde{y}^2 \leqslant (R/\sigma)^2} \exp\left\{-\frac{1}{2}\left(u^2+v^2\right)\right\} du\, dv =$$

$$= \left\| \begin{array}{l} u = r\cos\varphi, \\ v = r\sin\varphi \end{array} \right\| = \frac{1}{2\pi} \int\limits_0^{R/\sigma} \left(\int\limits_0^{2\pi} r\, e^{-r^2/2}\, d\varphi \right) dr =$$

$$= \int\limits_0^{R/\sigma} e^{-r^2/2}\, d\left(\frac{r^2}{2}\right) = 1 - \exp\left\{\frac{-R^2}{2\sigma^2}\right\}.$$

Теперь, решая уравнение

$$1 - \exp\left\{\frac{-R^2}{2\sigma^2}\right\} \approx 0{,}997,$$

получим $R^2 = -2\sigma^2 \cdot \ln 0{,}003 \approx 11{,}6\, \sigma^2$. Отсюда $R \approx 3{,}41\sigma$.

Ответ. $R \approx 3{,}41\sigma$.

§ 11. Многомерные случайные величины

11.1. Основные характеристики многомерных СВ.

Определение 11.1. Для непрерывной n*-мерной случайной величины* (*случайного вектора*) $X \triangleq \mathrm{col}(X_1, \ldots, X_n)$, где $X_1, \ldots, X_n$ — скалярные СВ, определенные на одном и том же пространстве элементарных событий, *функция распределения* $F(x_1, \ldots, x_n)$ и *плотность распределения* $f(x_1, \ldots, x_n)$, которая неотрицательна, определяются следующим образом:

$$F(x_1, \ldots, x_n) \triangleq \mathbf{P}\{X_1 \leqslant x_1, \ldots, X_n \leqslant x_n\} \triangleq \\ \triangleq \int\limits_{-\infty}^{x_1} \ldots \int\limits_{-\infty}^{x_n} f(t_1, \ldots, t_n)\, dt_1 \ldots dt_n.$$

Основные результаты, полученные для двумерной СВ, переносятся и на n-мерную СВ. В частности, в точках непрерывности плотности выполнено равенство

$$f(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F(x_1, \ldots, x_n).$$

Определение 11.2. *Математическим ожиданием* (МО) случайного вектора X называется вектор $\mathbf{M}[X] \triangleq m_X \triangleq \mathrm{col}(m_1, \ldots$ $\ldots, m_n)$, где $m_i \triangleq \mathbf{M}[X_i]$, $i = \overline{1, n}$.

Замечание 11.1. МО случайного вектора X есть вектор координат «средней» точки $(m_1, \ldots, m_n)$ в n-мерном пространстве $\mathbb{R}^n$, вокруг которой группируются реализации случайного вектора X.

Определение 11.3. Матрицу K размерности $n \times n$ с элементами $k_{ij} \triangleq \mathbf{M}[(X_i - m_i)(X_j - m_j)]$ называют *ковариационной матрицей.* Элементы k_{ij} ковариационной матрицы являются ковариациями СВ X_i и X_j при $i \neq j$, а диагональные элементы k_{ii} — дисперсиями СВ X_i, т. е.

$$k_{ii} \triangleq d_i \triangleq \sigma_i^2 = \mathbf{M}[(X_i - m_i)^2], \quad i = \overline{1, n}.$$

Дисперсии d_i, $i = \overline{1, n}$, характеризуют рассеивание реализаций компонент случайного вектора относительно средней точки $m_X \triangleq \mathrm{col}(m_1, \ldots, m_n)$, а при фиксированных дисперсиях ковариации k_{ij} характеризуют степень линейной зависимости между СВ X_i

и X_j. В частности, по свойству 2) k_{XY} при линейной связи между X_i и X_j ковариация между ними равна $k_{ij} = \pm\sigma_i\sigma_j$, а по свойству 1) k_{XY} всегда $|k_{ij}| \leqslant \sigma_i\sigma_j$.

Определение 11.4. Нормированную ковариационную матрицу R, элементами которой являются коэффициенты корреляции r_{ij}, называют *корреляционной матрицей.*

Матрицы K и R неотрицательно определены и, кроме того симметричны, так как

$$k_{ij} \triangleq \mathbf{M}[(X_i - m_i)(X_j - m_j)] = \mathbf{M}[(X_j - m_j)(X_i - m_i)] \triangleq k_{ji}.$$

Определение 11.5. СВ $X_1, \ldots, X_n$ называются *независимыми*, если для любых $x_1, \ldots, x_n \in \mathbb{R}^1$: $F(x_1, \ldots, x_n) = \prod\limits_{i=1}^{n} F_i(x_i)$, где $F_i(x)$ — функция распределения СВ X_i.

Определение 11.6. СВ $X_1, \ldots, X_n$ называются *попарно некоррелированными*, если $k_{ij} = 0$ при всех $i \neq j$.

Если СВ $X_1, \ldots, X_n$ независимы, то они являются и попарно некоррелированными. Обратное утверждение неверно.

Если СВ $Z \triangleq \operatorname{col}(X_1, \ldots, X_n)$ является непрерывной, то для независимости $X_1, \ldots, X_n$ необходимо и достаточно, чтобы $f(x_1, \ldots, x_n) = \prod\limits_{i=1}^{n} f_i(x_i)$ во всех точках непрерывности этих функций. Этот факт доказывается по индукции на основе свойства 8) $f(x, y)$ для двух случайных величин.

Свойства $\mathrm{M}[X]$ и $\mathrm{D}[X]$

1) $\mathbf{M}[\sum\limits_{i=1}^{n} X_i] = \sum\limits_{i=1}^{n} \mathbf{M}[X_i]$. Пусть $Z \triangleq \operatorname{col}(X, Y)$ — непрерывная двумерная СВ, где $X \triangleq X_1$, $Y \triangleq X_2$. По свойствам 6) $f(x, y)$ и 7) $f(x, y)$ имеем

$$\mathbf{M}[X+Y] = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}(x+y)f(x,y)\,dx\,dy = \int\limits_{-\infty}^{+\infty} x\Big(\int\limits_{-\infty}^{+\infty} f(x,y)\,dy\Big)\,dx + {} \\ + \int\limits_{-\infty}^{+\infty} y\Big(\int\limits_{-\infty}^{+\infty} f(x,y)\,dx\Big)\,dy = \int\limits_{-\infty}^{+\infty} x f_X(x)\,dx + \int\limits_{-\infty}^{+\infty} y f_Y(y)\,dy \triangleq \mathbf{M}[X] + \mathbf{M}[Y].$$

Аналогично доказывается формула и для дискретных СВ X и Y. Общая формула доказывается по индукции.

2) Если СВ X_i, $i = \overline{1,n}$, независимы, то $\mathbf{M}[\prod\limits_{i=1}^{n} X_i] = \prod\limits_{i=1}^{n} \mathbf{M}[X_i]$.

Пусть $Z \triangleq \mathrm{col}(X,Y)$ — непрерывная двумерная СВ, тогда по свойствам 7) $f(x,y)$ и 8) $f(x,y)$ имеем

$$\mathbf{M}[XY] = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf(x,y)\,dx\,dy = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} xyf_X(x)f_Y(y)\,dx\,dy =$$

$$= \int\limits_{-\infty}^{+\infty} xf_X(x)\,dx \int\limits_{-\infty}^{+\infty} yf_Y(y)\,dy = \mathbf{M}[X]\mathbf{M}[Y].$$

Общая формула доказывается по индукции.

3) $\mathbf{D}[\sum\limits_{i=1}^{n} X_i] = \sum\limits_{i=1}^{n} \mathbf{D}[X_i] + 2\sum\limits_{i=1}^{n-1}\sum\limits_{j=i+1}^{n} k_{ij}$. Пусть $Z \triangleq \mathrm{col}(X,Y)$, тогда по свойству 1)M $[X]$

$$\begin{aligned}
\mathbf{D}[X+Y] &= \mathbf{M}[(X+Y-\mathbf{M}[X+Y])^2] = \\
&= \mathbf{M}[(X+Y-(m_X+m_Y))^2] = \mathbf{M}[((X-m_X)+(Y-m_Y))^2] = \\
&= \mathbf{M}[(X-m_X)^2] + \mathbf{M}[(Y-m_Y)^2] + 2\mathbf{M}[(X-m_X)(Y-m_Y)] = \\
&= \mathbf{D}[X] + \mathbf{D}[Y] + 2k_{XY}.
\end{aligned}$$

Аналогично доказывается формула и для дискретных СВ X и Y. Общая формула доказывается по индукции.

4) Если СВ X_i, $i = \overline{1,n}$, попарно некоррелированы, т. е. $k_{ij} = 0$ при $i \neq j$, то из свойства 3)M $[X]$ следует

$$\mathbf{D}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mathbf{D}[X_i].$$

5) $|r_{ij}| \leqslant 1$ при $i \neq j$. Обозначим $X \triangleq X_i$, $Y \triangleq X_j$. Пусть $\sigma_X > 0$, $\sigma_Y > 0$. Рассмотрим нормированные СВ $\overset{*}{X} \triangleq \dfrac{X - m_X}{\sigma_X}$, $\overset{*}{Y} \triangleq \dfrac{Y - m_Y}{\sigma_Y}$, для которых $\mathbf{D}[\overset{*}{X}] = \mathbf{D}[\overset{*}{Y}] = 1$. По определению 10.8 получаем $r_{XY} = k_{\overset{*}{X}\overset{*}{Y}}$. Поэтому, по доказанному выше свойству

$$\mathbf{D}[\overset{*}{X} + \overset{*}{Y}] = \mathbf{D}[\overset{*}{X}] + \mathbf{D}[\overset{*}{Y}] + 2k_{\overset{*}{X}\overset{*}{Y}} = 2 + 2r_{XY}.$$

Так как дисперсия не может быть отрицательной, т. е. $\mathbf{D}[\overset{*}{X} + \overset{*}{Y}] \geqslant 0$, то получаем неравенство $2 + 2r_{XY} > 0$. Отсюда следует, что $r_{XY} \geqslant -1$.

Рассматривая дисперсию $\mathbf{D}[\overset{*}{X} - \overset{*}{Y}]$, можно аналогично получить, что выполняется неравенство $r_{XY} \leqslant 1$. Таким образом, $|r_{XY}| \leqslant 1$.

11.2. Многомерное нормальное распределение.

Определение 11.7. Говорят, что n-мерная СВ $X \triangleq \mathrm{col}(X_1, \ldots, X_n)$ имеет *нормальное* (*гауссовское*) *распределение*, $X \sim \mathbf{N}(m; K)$, если ее плотность распределения есть

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^n \det K}} \exp\left\{-\frac{1}{2}(x-m)^T K^{-1}(x-m)\right\},$$

где $\det K$ — определитель положительно определенной матрицы K, $m \triangleq \mathrm{col}(m_1, \ldots, m_n)$, $x \triangleq \mathrm{col}(x_1, \ldots, x_n)$.

Плотность нормального распределения можно задать также через элементы обратной матрицы K^{-1} таким образом, как это сделано в определении 10.11 при рассмотрении двумерной СВ.

Свойства нормального распределения $\mathbf{N}(m; K)$

1) $\mathbf{M}[X] = m$, ковариационная матрица СВ X равна $\mathbf{M}[(X - m)(X - m)^T] = K$.

2) Так как матрица K — невырожденная, то каждая i-я компонента X_i вектора X распределена нормально.

3) Если $X \sim \mathbf{N}(m; K)$, то $Y \triangleq AX + b$, где A — матрица размерности $n \times k$ с $\mathrm{rang} A = k$ и b — вектор размерности n, имеет распределение $\mathbf{N}(m_Y; K_Y)$, где $m_Y = Am + b$, $K_Y = AKA^T$.

4) Пусть СВ $Y \triangleq a_1 X_1 + \ldots + a_n X_n$, где $X_1, \ldots, X_n$ распределены нормально, а коэффициенты $a_1, \ldots, a_n \in \mathbb{R}^1$ не все равны нулю. Тогда СВ Y распределена нормально.

5) Если случайный вектор имеет нормальное распределение, а его компоненты попарно некоррелированны, то

$$f(x_1, \ldots, x_n) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left\{-\frac{(x_1 - m_1)^2}{2\sigma_1^2}\right\} \times \ldots$$
$$\ldots \times \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left\{-\frac{(x_n - m_n)^2}{2\sigma_n^2}\right\} = f_1(x_1) \cdot \ldots \cdot f_n(x_n),$$

т. е. СВ $X_1, \ldots, X_n$ независимы. Но из независимости следует некоррелированность (см. свойство 3) k_{XY}), поэтому для нормального распределения условия некоррелированности и независимости эквивалентны.

11.3. Биржевой парадокс. Рассмотрим любопытный экономический пример. Пусть имеется начальный капитал K, который требуется увеличить. Для этого имеются две возможности: вкладывать деньги в надежный банк и покупать на бирже акции некоторой компании. Пусть u — доля капитала, вкладываемая в банк, а v — доля капитала, расходуемая на приобретение акций. Очевидно, что $0 \leqslant u + v \leqslant 1$. Предположим, банк гарантирует $b \times 100\% > 0$ годовых, а акции приносят $X \times 100\%$ годовых. Так как предполагается, что банк абсолютно надежен, то b является неслучайной величиной. Стоимость акций, как правило, меняется в течение года, т. е. X является случайной величиной. Допустим, что приобретение акций в среднем более прибыльно, чем вложение средств в банк, т. е. $m_X \triangleq \mathbf{M}[X] > b > 0$. Но при этом имеется ненулевая вероятность того, что акции обесценятся и мы потеряем все деньги, вложенные в акции, $\mathbf{P}\{X \leqslant -1\} = \varepsilon > 0$.

Таким образом, мы можем надеяться, что через год капитал составит величину $K_1 = K(1 + bu + Xv)$, которая является случайной. Рассмотрим также ожидаемое через год среднее значение капитала:

$$K_0(u, v) \triangleq \mathbf{M}[K(1 + bu + Xv)] = K(1 + bu + m_X v).$$

Поставим задачу распределить капитал таким образом, чтобы максимизировать средний доход за год:

$$K_0(u_0, v_0) = \max_{u+v \leqslant 1,\, u \geqslant 0,\, v \geqslant 0} K_0(u, v).$$

Нетрудно найти решение этой простой *задачи линейного программирования*. Так как по условию задачи $m_X > b > 0$, то, очевидно, все деньги нужно вкладывать в акции, которые в среднем более прибыльны, чем вложение в банк, т. е. $u_0 = 0$, $v_0 = 1$. При такой стратегии среднее значение капитала через год будет максимально:

$$K_0(u_0, v_0) = K(1 + m_X).$$

Выясним, к чему приведет такая стратегия управления капиталом, если применять ее многократно. Пусть X_i, $i = \overline{1, n}$, — ежегодный прирост капитала за счет приобретения акций. Предположим, что СВ X_i, $i = \overline{1, n}$, независимы. Пусть $u_i = 0$, $v_i = 1$, $i = \overline{1, n}$, т. е. ежегодно покупаются только акции, которые в среднем более прибыльны, чем вложение в банк, $\mathbf{M}[X_i] = m_X > b > 0$. Тогда среднее значение капитала через n лет составит величину

$$K_n \triangleq \mathbf{M}\Big[K \prod_{i=1}^{n} (1 + X_i)\Big] = K \prod_{i=1}^{n} (1 + \mathbf{M}[X_i]) = K(1 + m_X)^n.$$

Здесь использовалось свойство 2)M[X]. Так как по предположению $m_X > 0$, то $1 + m_X > 1$. Поэтому при $n \to \infty$ получаем $K_n \to \to \infty$. Образно говоря, при таком управлении капиталом можно стать неограниченно богатым «в среднем».

Посмотрим, что происходит с вероятностью нашего разорения при выбранной стратегии

$$\mathbf{P}(B_n) \triangleq \mathbf{P}(A_1 + \ldots + A_n),$$

где событие $A_i \triangleq \{X_i : 1 + X_i \leqslant 0\}$ характеризует разорение в i-й год, а событие $B_n \triangleq A_1 + \ldots + A_n$ — возможность разорения хотя бы один раз за n лет.

Рассмотрим противоположное событие $\overline{B}_n = \Omega \backslash B_n$. Воспользовавшись свойством $11)A$, находим

$$\overline{B}_n = \prod_{i=1}^{n} \overline{A}_i, \quad \text{где} \quad \overline{A}_i = \{X_i : 1 + X_i > 0\}.$$

Так как СВ X_i независимы, то независимы также события A_i, $i = = \overline{1, n}$. Поэтому имеем

$$\mathbf{P}(\overline{B}_n) = \prod_{i=1}^{n} \mathbf{P}(\overline{A}_i) = \prod_{i=1}^{n} \mathbf{P}\{X_i + 1 > 0\}.$$

Но по предположению $\mathbf{P}(A_i) = \mathbf{P}\{X_i \leqslant -1\} = \varepsilon > 0$, т. е. с ненулевой вероятностью можно потерять весь капитал в каждый i-й год. Поэтому $\mathbf{P}(\overline{A}_i) = 1 - \mathbf{P}(A_i) = 1 - \varepsilon < 1$. Отсюда следует, что

$$\mathbf{P}(\overline{B}_n) = (1 - \varepsilon)^n \to 0 \quad \text{при} \quad n \to \infty.$$

А следовательно, $\mathbf{P}(B_n) = 1 - \mathbf{P}(\overline{B_n}) \to 1$ при $n \to \infty$, т. е. вероятность разорения при выбранной стратегии стремится к единице. И это несмотря на то, что средний доход стремится к бесконечности. В этом и состоит *биржевой парадокс*, к которому мы пришли, решив покупать лишь одни акции, пренебрегая возможностью получения в банке хоть и небольшой, но зато гарантированной, прибыли. Это значит, что не следует «складывать все яйца в одну корзину».

На практике, чтобы избежать этого парадокса, используют так называемую *логарифмическую стратегию*, которая определяется из следующего условия:

$$L_0(u_L, v_L) = \max_{u+v \leqslant 1,\, u \geqslant 0,\, v \geqslant 0} \mathbf{M}[\ln(1 + bu + Xv)],$$

т. е. из условия максимизации средней скорости роста капитала. В частности, иногда предполагают, что СВ $Y \triangleq 1 + X$ имеет логнормальное распределение (см. определение 7.5). При логарифмической стратегии капитал распределяется, как правило, в некоторых пропорциях между покупкой акций и вложением в банк.

Следует также отметить, что рассмотренный пример хорошо иллюстрирует особенность поведения случайной последовательности

$$Z_n \triangleq K \prod_{i=1}^{n} (1 + bu_i + X_i v_i)$$

при $n \to \infty$. В отличие от детерминированной последовательности случайная последовательность может сходится в разных смыслах и к разным значениям. В данном примере если $u_i > 0$ и $v_i = 1$, то $Z_n \to \infty$ в среднем и в то же время $Z_n \to 0$ по вероятности.

11.4. Типовые задачи.

Задача 11.1. Для случайных величин X и Y, рассмотренных в задаче 9.2, проверить справедливость следующих равенств:

$$\mathbf{M}[X+Y] = \mathbf{M}[X] + \mathbf{M}[Y], \ \mathbf{D}[X+Y] = \mathbf{D}[X] + \mathbf{D}[Y] + 2k_{XY}. \quad (11.1)$$

Решение. Для проверки указанных равенств найдем необходимые значения $\mathbf{M}[X]$, $\mathbf{M}[Y]$, $\mathbf{D}[X]$, $\mathbf{D}[Y]$, k_{XY}, воспользовавшись полученными в задаче 9.2 рядами распределения СВ X и Y:

$$\mathbf{M}[X] = \sum_{i=1}^{2} x_i p_i = \frac{1}{2}, \quad \mathbf{M}[Y] = \sum_{i=1}^{3} y_j p_j = 1 \cdot \frac{3}{8} + 2 \cdot \frac{1}{2} = \frac{11}{8},$$

$$\mathbf{D}[X] = \sum_{i=1}^{2} (x_i - \mathbf{M}[X])^2 p_i = \left(-\frac{1}{2}\right)^2 \cdot \frac{1}{2} + \left(1 - \frac{1}{2}\right)^2 \cdot \frac{1}{2} = \frac{1}{4},$$

$$\mathbf{D}[Y] = \sum_{i=1}^{3} (y_j - \mathbf{M}[Y])^2 p_j =$$
$$= \left(-\frac{11}{8}\right)^2 \cdot \frac{1}{8} + \left(1 - \frac{11}{8}\right)^2 \cdot \frac{3}{8} + \left(2 - \frac{11}{8}\right)^2 \cdot \frac{1}{2} = \frac{31}{64},$$

$$r_{XY} = \sum_{i=1}^{2} \sum_{i=1}^{3} (x_i - \mathbf{M}[X])(y_j - \mathbf{M}[Y]) p_{ij} =$$
$$= \sum_{i=1}^{2} \sum_{j=1}^{3} x_i y_i p_{ij} - \mathbf{M}[X]\mathbf{M}[Y] = 1 \cdot \frac{1}{8} + 2 \cdot 1 \cdot \frac{3}{8} - \frac{1}{2} \cdot \frac{11}{8} = \frac{3}{16}.$$

Таким образом, имеем

$$\mathbf{M}[X] + \mathbf{M}[Y] = \frac{1}{2} + \frac{11}{8} = \frac{15}{8},$$
$$\mathbf{D}[Y] + \mathbf{D}[X] + 2k_{XY} = \frac{1}{4} + \frac{31}{64} + 2 \cdot \frac{3}{16} = \frac{71}{64}.$$

Согласно задаче 9.2, $\mathbf{M}[Z] = \dfrac{15}{8}$, $\mathbf{D}[Z] = \dfrac{71}{64}$, где $Z \triangleq X + Y$. Следовательно, справедливы равенства (11.1).

Ответ. Равенства (11.1) верны.

Задача 11.2. Дана ковариационная матрица случайного вектора $Z = \mathrm{col}(X_1, X_2, X_3)$:

$$K_Z = ||k_{ij}|| = \begin{pmatrix} 16 & -14 & 12 \\ -14 & 49 & -21 \\ 12 & -21 & 36 \end{pmatrix}.$$

Составить корреляционную матрицу $R_Z = ||r_{ij}||$.

Решение. Поскольку по определению

$$r_{ij} = \frac{k_{ij}}{\sqrt{k_{ii} \cdot k_{jj}}},$$

то

$$r_{11} = r_{22} = r_{33} = 1; \qquad r_{12} = r_{21} = \frac{-14}{\sqrt{16 \cdot 49}} = -\frac{14}{28} = -\frac{1}{2};$$

$$r_{13} = r_{31} = \frac{12}{\sqrt{16 \cdot 36}} = \frac{12}{24} = -\frac{1}{2}; \quad r_{23} = r_{32} = \frac{-21}{\sqrt{49 \cdot 36}} = -\frac{21}{42} = -\frac{1}{2}.$$

Ответ. $||r_{ij}|| = \begin{pmatrix} 1 & -1/2 & 1/2 \\ -1/2 & 1 & -1/2 \\ 1/2 & -1/2 & 1 \end{pmatrix}$.

Задача 11.3. Известно, что $\mathbf{D}[X] = 1$, $\mathbf{D}[Y] = 2$, $\mathbf{D}[X+Y] = 3$. Найти k_{XY}.

Решение. Используя формулу для дисперсии суммы, получаем

$$\mathbf{D}[X+Y] = \mathbf{D}[X] + \mathbf{D}[Y] + 2 \cdot k_{XY},$$

откуда

$$k_{XY} = \frac{1}{2}\left(\mathbf{D}[X+Y] - \mathbf{D}[X] - \mathbf{D}[Y]\right) = \frac{3-1-2}{2} = 0.$$

Ответ. $k_{XY} = 0$.

Задача 11.4. На столе налогового инспектора лежат три декларации от представителей различных групп населения. Вероятности сокрытия доходов при заполнении декларации для одного представителя каждой группы равны соответственно 0,05, 0,1 и 0,15. Предположим, что сокрытие доходов обнаруживается при проверке в 100% случаев. Найти средний доход государства от проверки этих деклараций, если сумма налагаемого штрафа при обнаружении сокрытия дохода составляет по группам населения 100, 250 и 500 минимальных окладов соответственно.

Р е ш е н и е. Рассмотрим случайную величину X, равную доходу государства от проверки трех деклараций. В задаче требуется найти *средний доход*, т. е. математическое ожидание СВ X. Можно найти $\mathbf{M}[X]$, построив ряд распределения СВ X и воспользовавшись определением математического ожидания. Однако процедура построения ряда распределения СВ X представляется достаточно трудоемкой. Данная задача допускает более простое решение, основанное на использовании распределения Бернулли. Случайная величина X может быть представлена следующим образом:

$$X = 100X_1 + 250X_2 + 500X_3,$$

где $X_i \sim \mathbf{Bi}(1; p_i)$, $i = \overline{1,3}$, $p_1 = 0{,}05$, $p_2 = 0{,}1$, $p_3 = 0{,}15$.

Случайная величина X_i, имеющая распределение Бернулли, принимает значение 1, если i-й подавший декларацию скрывает доход, и 0 — в противном случае, причем $\mathbf{M}[X_i] = p_i$, $i = \overline{1,3}$. Воспользуемся свойствами математического ожидания для вычисления $\mathbf{M}[X]$:

$$\mathbf{M}[X] = 100\,\mathbf{M}[X_1] + 250\,\mathbf{M}[X_2] + 500\,\mathbf{M}[X_3] = \\ = 100p_1 + 250p_2 + 500p_3 = 105.$$

О т в е т. Средний доход государства от проверки поданных трех деклараций составит 105 минимальных окладов.

З а д а ч а 11.5. Известно, что случайный вектор $Z = \mathrm{col}(X, Y)$, где СВ X — рост наугад взятого взрослого мужчины и СВ Y — его вес, удовлетворительно описывается двумерным нормальным законом распределения с математическим ожиданием $m_Z = \mathrm{col}(175, 74)$ и ковариационной матрицей

$$K_Z = ||k_{ij}|| = \begin{pmatrix} 49 & 28 \\ 28 & 36 \end{pmatrix}.$$

Считается, что человек страдает избыточным весом, если выполняется неравенство $X - Y \leqslant 90$. Найти: *а*) математическое ожидание и дисперсию характеристики избыточного веса $X - Y$; *б*) вероятность того, что наугад выбранный мужчина страдает избыточным весом.

Р е ш е н и е. *а*) Так как вектор $Z = \mathrm{col}(X, Y)$ распределен нормально, то нормально распределены и его координаты: $X \sim \mathbf{N}(175; 49)$, $Y \sim \mathbf{N}(74; 36)$. Теперь, согласно свойству 2) $\mathbf{N}(m; K)$, разность $X - Y$ распределена нормально. Вычислим параметры этого закона распределения

$$\mathbf{M}[X-Y] = 175 - 74 = 101;$$
$$\mathbf{D}[X-Y] = \mathbf{D}[X] + \mathbf{D}[Y] - 2k_{XY} = 49 + 36 - 2 \cdot 28 = 29.$$

б) Таким образом, $X - Y \sim \mathbf{N}(101; 29)$, и следовательно

$$\mathbf{P}\{X - Y \leqslant 90\} = \frac{1}{2} + \Phi_0\left(\frac{90-101}{\sqrt{29}}\right) \approx 0{,}5 - \Phi_0(2{,}04) \approx 0{,}021.$$

Ответ. *а*) $\mathbf{M}[X-Y] = 101$, $\mathbf{D}[X-Y] = 29$; *б*) $\mathbf{P}\{X - Y \leqslant 90\} \approx \approx 0{,}021$.

§ 12. Задачи для самостоятельного решения

1. Пусть $F(x,y)$ — функция распределения случайного вектора $Z = \mathrm{col}(X,Y)$. Будет ли $F(x,0)$ функцией распределения некоторой случайной величины?

2. СВ $X \sim \mathbf{Bi}(1; 0{,}2)$, $Y \sim \mathbf{Bi}(1; 0{,}8)$. Для СВ $Z = X + Y$ вычислить $F_Z(1)$, считая все события, связанные с X и Y, независимыми.

3. Случайные величины $X \sim \mathbf{N}(0;1)$, $Y \sim \mathbf{E}(1)$. Вычислить $\mathbf{M}[X^3 - Y^2]$.

4. Подбрасывают три игральные кости. Рассматриваются случайные величины: X — количество костей, на которых выпало шесть очков, Y — количество костей, на которых выпало пять очков. Найти $\mathbf{M}[X+Y]$ и закон распределения СВ $Z = X + Y$.

5. Задан закон распределения случайного вектора $Z = \mathrm{col}(X,Y)$ (см. табл. 12.1). Требуется:

а) найти закон распределения случайной величины $X + Y$;

б) проверить справедливость равенств:

$$\mathbf{M}[X+Y] = \mathbf{M}[X] + \mathbf{M}[Y],$$

$$\mathbf{D}[X+Y] = \mathbf{D}[X] + \mathbf{D}[Y] + 2k_{XY}.$$

Таблица 12.1

$Y\backslash X$	-1	1
1	1/6	1/3
2	0	1/6
3	1/3	0

6. Найти ковариацию k_{cX}, где X — некоторая СВ, а c — константа.

7. В продукции завода брак вследствие дефекта А составляет 3%, а вследствие дефекта В — $4,5\%$. Годная продукция составляет 95%. Найти коэффициент корреляции дефектов А и В.

8. Случайный вектор $Z = \mathrm{col}(X, Y)$ имеет совместную плотность вероятности $f(x, y) = \dfrac{a}{(1 + x^2 + x^2y^2 + y^2)}$. Требуется:

а) найти неизвестную константу a;

б) исследовать величины X и Y на независимость и некоррелированность.

9. Случайный вектор $Z = \mathrm{col}(X, Y)$ имеет функцию распределения

$$F(x, y) = \begin{cases} 1 + e^{-(\lambda x + \mu y)} - e^{-\lambda x} - e^{-\mu y}, & x \geqslant 0,\ y \geqslant 0, \\ 0 - & \text{в остальных случаях.} \end{cases}$$

Найти частные распределения СВ X и Y. Исследовать эти величины на независимость и некоррелированность.

10. Является ли функция $F(x, y) = x \cdot y$, где $x,\ y \in \mathbb{R}^1$, функцией распределения некоторого случайного вектора?

11. Является ли функция

$$F(x, y) = \frac{1}{\pi^2}\left(\frac{\pi}{2} + \mathrm{arctg}\, x\right)\left(\frac{\pi}{2} + \mathrm{arctg}\, y\right), \quad x,\ y \in \mathbb{R}^1,$$

функцией распределения некоторого случайного вектора?

12. Пусть $f(x, y)$— плотность вероятности случайного вектора $Z = \mathrm{col}(X, Y)$. Будет ли $f(y, x)$ плотностью вероятности некоторого случайного вектора?

13. Известно, что $\mathbf{M}[X] = 1$, $\mathbf{M}[X^2] = 2$. Найти k_{XX}.

14. Известно, что СВ $X \sim \mathbf{E}(1)$, $\mathbf{D}[Y] = 2$, $\mathbf{D}[X - Y] = 3$. Найти k_{XY}.

15. Найти коэффициент корреляции между случайными величинами: *а*) X и $Y \triangleq 18X$; *б*) X и $Y \triangleq 7 - 2X$.

16. Известно, что СВ $X \sim \mathbf{R}(-1; 1)$. Пусть $Y \triangleq \cos X$. Найти k_{XY}.

17. Компания решила осуществить три независимых проекта по созданию уникальной продукции в единственном экземпляре. Затраты на осуществление каждого проекта одинаковы и равны 1000 у.е. Цена экземпляра продукции, полученного в результате осуществления первого проекта, в 3 раза превышает затраты на его создание. Для оставшихся двух проектов соотношение цены к затратам равно соответственно 5 и 7. Вероятности того, что полученные три экземпляра при установленных ценах найдут сбыт, равны соответственно 0,9 , 0,7 , 0,55 . Найти, во сколько раз средний доход от продажи всей продукции превысит суммарные затраты на ее производство (при условии неизменности цен).

18. У каждого из двух основных предвыборных блоков партий и объединений имеется список из трех наиболее вероятных кандидатур на выдвижение в качестве претендента на пост президента. Рассмотрим случайный вектор $Z = \mathrm{col}(X, Y)$, где X — номер претендента, выдвигаемого для финального тура голосования в списке первого блока, а Y — аналогичный номер претендента в списке второго блока. Вероятность появления различных пар в финальном туре голосования, согласно оценкам экспертов-политологов, отражена в табл. 12.2. Определить наиболее вероятных претендентов от каждого блока и коэффициент корреляции между номерами претендентов в финальном туре голосования.

Таблица 12.2

$X\backslash Y$	1	2	3
1	1/3	1/3	1/27
2	1/27	1/9	1/27
3	1/27	1/27	1/27

19. Брак в продукции завода вследствие дефекта А составил 6% , причем среди забракованной по признаку А продукции в 4% случаев встречается дефект В, а в продукции, свободной от дефекта А, дефект В встречается в 1% случаев. Найти вероятность встретить дефект В во всей продукции и коэффициент корреляции между признаками А и В.

20. Распределение случайного вектора задано табл. 12.3. Найти закон распределения случайной величины

$$U = X^2 + Y^2 - 1.$$

Таблица 12.3

$X\backslash Y$	-1	0	1
-1	0,07	0,1	0,13
1	0,2	0,23	0,27

21. Две независимые случайные величины X и Y подчинены показательному закону: $X \sim \mathbf{E}(\lambda)$, $Y \sim \mathbf{E}(\mu)$. Написать выражение совместной плотности вероятности и функции распределения. Подсчитать вероятность попадания в квадрат с вершинами $A(0,0)$, $B(1,0)$, $C(0,1)$, $D(1,1)$.

22. Найти вероятность того, что случайно брошенная точка с координатами (X,Y) попадет на область D, определенную неравенствами $\{1 \leqslant x \leqslant 2, 1 \leqslant y \leqslant 2\}$, если функция распределения координат этой точки равна

$$F(x,y) = \begin{cases} 1 - 2^{-x^2} - 2^{-2y^2} + 2^{-x^2-2y^2}, & x \geqslant 0,\ y \geqslant 0, \\ 0 - & \text{в остальных случаях.} \end{cases}$$

23. Случайный вектор $Z = \mathrm{col}(X, Y)$ имеет совместную плотность вероятности

$$f(x,y) = \begin{cases} 4xye^{-(x^2+y^2)}, & x \geqslant 0,\ y \geqslant 0, \\ 0 - & \text{в остальных случаях.} \end{cases}$$

Найти корреляционную матрицу случайного вектора Z.

24. Плотность распределения вероятностей случайного вектора $Z = \mathrm{col}(X, Y)$ имеет вид

$$f(x,y) = \begin{cases} c(xy + y^2), & 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 2, \\ 0 - & \text{в остальных случаях.} \end{cases}$$

Найти:

а) постоянную c;

б) вероятность $\mathbf{P}\{X + Y < 2\}$;

в) функцию распределения СВ X.

25. Задана функция распределения случайного вектора $Z = \mathrm{col}(X, Y)$:

$$F(x,y) = \begin{cases} 1 - 2^{-x} - 2^{-y} + 2^{-x-y}, & x \geqslant 0,\ y \geqslant 0, \\ 0 - & \text{в остальных случаях.} \end{cases}$$

Найти:

а) двумерную плотность распределения случайного вектора Z;

б) вероятность попадания случайной точки с координатами Z в треугольник с вершинами $A(1,3)$, $B(3,3)$, $C(2,8)$.