Закони розподілу ймовірностей систем дискретних або неперервних випадкових величин.

Якщо на одному й том ж просторі елементарних подій задані випадкові величини $X_1, X_2, ..., X_n$, то кажуть, що заданий ***n*-вимірний випадковий вектор** $\left(X_1, X_2, ..., X_n\right)$.

Функція розподілу випадкового вектора $\left(X_1, X_2, ..., X_n\right)$ визначається співвідношенням

$$F\left(x_1, x_2, ..., x_n\right) = P\left(X_1 < x_1, X_2 < x_2, ..., X_n < x_n\right).$$

Зокрема для двовимірного випадкового вектора $\left(X, Y\right)$ маємо за означенням

$$F\left(x, y\right) = P\left(X < x, Y < y\right).$$

Геометричним тлумаченням цієї функції розподілу є ймовірність попадання випадкової точки $\left(X, Y\right)$ в нескінчений квадрант з вершиною в точці $\left(x, y\right)$, розміщений нижче та лівіше цієї вершини

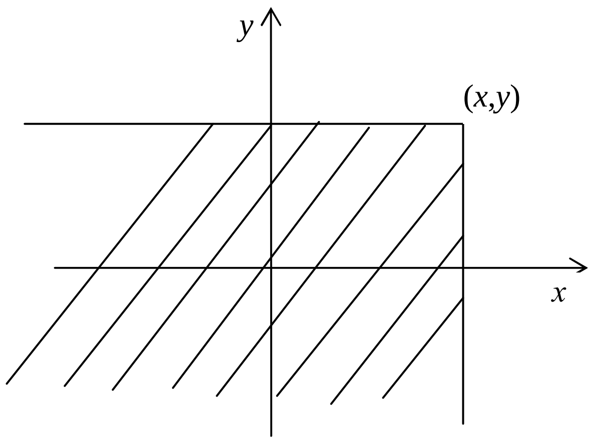

Рис.

Властивості функції розподілу:

1) $F\left(+\infty, +\infty\right) = 1$;

2) $F(x,\ -\infty) = F(-\infty,\ y) = F(-\infty,\ -\infty) = 0$;

3) функція розподілу неперервна зліва по кожному аргументу;

4) умова узгодженості: $F(x,\ +\infty) = F_X(x)$, $F(+\infty,\ y) = F_Y(y)$;

5) $F(x,\ y)$ – монотонно неспадна функція по кожному аргументу;

6) $P\{a \le X < b,\ \ c \le Y < d\} =$

$$= F(b,\ d) - F(b,\ c) - F(a,\ d) + F(a,\ c).$$

Закон розподілу дискретного двовимірного випадкового вектора $(X,\ Y)$ може бути заданим за допомогою таблиці

Y	X				
	x_1	x_2	…	x_i	…
y_1	p_{11}	p_{21}	…	p_{i1}	…
y_2	p_{12}	p_{22}	…	p_{i2}	…
…	…	…	…	…	…
y_j	p_{1j}	p_{2j}	…	p_{ij}	…
…	…	…	…	…	…

де $p_{ij} = P\left(X = x_i,\ Y = y_j\right)$. При цьому $\sum\limits_i \sum\limits_j p_{ij} = 1$.

Додаючи ймовірності стовпців і рядків, одержимо закони розподілу компонент X і Y. Закони розподілу випадкових величин X і Y:

X	x_1 x_2 . . . x_m . . .
p	p_1 p_2 . . . p_m . . .

Y	y_1 y_2 . . . y_n . . .
q	q_1 q_2 . . . q_n . . .

де $p_1 = \sum\limits_j p_{1j}$, $p_2 = \sum\limits_j p_{2j}$, …, $p_m = \sum\limits_j p_{mj}$, …;

$$q_1 = \sum_i p_{i1}, \ q_2 = \sum_i p_{i2}, \ q_n = \sum_i p_{in}, \ \dots$$

Для двовимірного випадкового вектора $(X,\ Y)$ неперервного типу функція розподілу дорівнює

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv,$$

де $f(x, y)$ – **двовимірна щільність** випадкового вектора.

Властивості двовимірної щільності випадкового вектора:

1) $f(x, y) \geq 0$;
2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$;
3) $f(x,y) = F''_{xy}(x,y)$;
4) $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy, \ f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx$;
5) ймовірність попадання випадкової точки в область D дорівнює $P\left((X,\ Y) \in D\right) = \iint_D f(x,\ y)dxdy$.

Випадкові величини X і Y називаються **незалежними**, якщо $F(x,\ y) = F_X(x)F_Y(y)$.

Для дискретних незалежних випадкових величин X і Y

$$p_{ij} = P\left(X = x_i\right) \cdot P\left(\ Y = y_j\right).$$

Для неперервних незалежних випадкових величин X і Y

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

Точка $(MX,\ MY)$ називається **центром розсіювання** випадкового вектора $(X,\ Y)$.

Кореляційним моментом випадкових величин X і Y називається число $K(X,\ Y) = M\left((X - MX)(Y - MY)\right)$.

Кореляційний момент можна знаходити за формулою

$$K\left(X,Y\right)=M\left(XY\right)-MX\cdot MY.$$

Для дискретного двовимірного випадкового вектора $\left(X,\ Y\right)$

$$M\left(XY\right)=\sum_{i}\sum_{j}x_i y_j p_{ij}\ ,$$

для неперервного двовимірного випадкового вектора $\left(X,\ Y\right)$

$$M\left(XY\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf\left(x,\ y\right)dxdy.$$

Якщо випадкові величини X і Y незалежні, то $K\left(X,\ Y\right)=0$.

Нехай задано *n*-вимірний випадковий вектор $\left(X_1,\ X_2,\ ...,\ X_n\right)$. Матриця, складена з кореляційних моментів, називається **кореляційною матрицею**:

$$\begin{pmatrix} K\left(X_1,\ X_1\right) & K\left(X_1,\ X_2\right) & ... & K\left(X_1,\ X_n\right) \\ K\left(X_2,\ X_1\right) & K\left(X_2,\ X_2\right) & ... & K\left(X_2,\ X_n\right) \\ & & ... & \\ K\left(X_n,\ X_1\right) & K\left(X_n,\ X_2\right) & ... & K\left(X_n,\ X_n\right) \end{pmatrix}.$$

Це симетрична матриця є *n*-го порядку, елементами головної діагоналі якої є дисперсії відповідних випадкових величин.

Для характеристики зв’язку між випадковими величинами X і Y служить **коефіцієнт кореляції**

$$r\left(X,\ Y\right)=\frac{K\left(X,\ Y\right)}{у\left(X\right)\cdot у\left(Y\right)}.$$

Властивості коефіцієнта кореляції:

1) для довільних двох випадкових величин $\left|r\left(X,\ Y\right)\right|\le 1$;

2) $\left|r\left(X,\ Y\right)\right|=1$ тоді і тільки тоді, коли випадкові величини X и Y зв'язані лінійною залежністю $Y = aX + b$, причому $r\left(X,\ Y\right)=1$ при $a > 0$, $r\left(X,\ Y\right)=-1$ при $a < 0$;

3) якщо випадкові величини X і Y незалежні, то

$$r\left(X,\ Y\right)=0.$$

Випадкові величини X та Y називаються **некорельованими**, якщо $r\left(X,\ Y\right)=0$.

З некорельованості двох випадкових величин, взагалі кажучи, не випливає їх незалежність.

Коефіцієнт кореляції служить для оцінки тісноти лінійного зв'язку між випадковими величинами X і Y: чим ближче модуль коефіцієнта кореляції до одиниці, тим зв'язок сильніший, чим ближче модуль коефіцієнта кореляції до нуля, тим зв'язок слабший.

Двовимірний випадковий вектор (X, Y) має **нормальний закон розподілу**, якщо його щільність розподілу має вигляд

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\,e^{-\frac{1}{2(1-r^2)}\left(\frac{(x-a_1)^2}{\sigma_1^2}-2r\frac{(x-a_1)(y-a_2)}{\sigma_1\sigma_2}+\frac{(y-a_2)^2}{\sigma_2^2}\right)}.$$

Нормальний закон на площині визначається п'ятьма параметрами. Можна довести, що a_1, a_2 – математичні сподівання, σ_1, σ_2 – середні квадратичні відхилення, r – коефіцієнт кореляції випадкових величин X і Y.

Нехай X і Y некорельовані, тобто $r = 0$. Тоді

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2}\,e^{-\left(\frac{(x-a_1)^2}{2\sigma_1^2}+\frac{(y-a_2)^2}{2\sigma_2^2}\right)}=\frac{1}{\sigma_1\sqrt{2\pi}}\,e^{-\frac{(x-a_1)^2}{2\sigma_1^2}}\ \frac{1}{\sigma_2\sqrt{2\pi}}\,e^{-\frac{(y-a_2)^2}{2\sigma_2^2}},$$

$$f(x,y)=f_1(x)\cdot f_2(y).$$

Отже, X і Y незалежні. Для нормально розподілених випадкових величин некорельованість рівносильна незалежності.

Двовимірний випадковий вектор (X, Y) має **рівномірний закон розподілу** в області D, якщо його щільність розподілу має вигляд

$$f(x, y) = \begin{cases} \dfrac{1}{S(D)}, \ (x, y) \in D \\ 0, \ (x, y) \notin D \end{cases},$$

де $S(D)$ – площа області D.

Нехай (X, Y) дискретний двовимірний випадковий вектор. **Умовним розподілом** компоненти X при $Y = y_j$ називається розподіл:

X	x_1	x_2	$\dots$	x_i	$\dots$
$p(X / y_j)$	$p(x_1 / y_j)$	$p(x_2 / y_j)$	$\dots$	$p(x_i / y_j)$	$\dots$

Умовні ймовірності $p(x_i / y_j)$ знаходять за формулою:

$$p(x_i / y_j) = \frac{p_{ij}}{\sum_i p_{ij}} = \frac{p_{ij}}{p(Y = y_j)}.$$

Аналогічно визначається умовний закон розподілу Y при $X = x_i$.

Нехай (X, Y) – неперервний двовимірний випадковий вектор. **Умовною щільністю** $\varphi(x/y)$ розподілу компоненти X при даному значенні $Y = y$ називається відношення щільності розподілу двовимірної випадкової величини $f(x, y)$ до щільності розподілу $f_2(y)$ компоненти Y:

$$\text{ц}(x / y) = \frac{f(x, y)}{f_2(y)}.$$

Аналогічно визначається умовна щільність компоненти Y при $X = x$:

$$\text{ш}(y / x) = \frac{f(x, y)}{f_1(x)}.$$

Розв’язування типових задач

Задача 1. Двовимірний дискретний випадковий вектор $(X,\ Y)$ заданий таблицею розподілу

	X			
Y	1	2	4	$\sum$
0	0,1	0	0,1	0,2
2	0	0,3	0,3	0,6
5	0,2	0	0	0,2
$\sum$	0,3	0,3	0,4	1

Знайдіть коефіцієнт кореляції.

Розв’язання. Додаючи стовбці та рядки таблиці розподілу, знайдемо закони розподілу X і Y відповідно:

X	1	2	4
p	0,3	0,3	0,4

Y	0	2	5
p	0,2	0,6	0,2

Знаходимо:

$M(X) = 1\cdot 0,3 + 2\cdot 0,3 + 4\cdot 0,4 = 2,5;$

$M(X^2) = 1\cdot 0,3 + 4\cdot 0,3 + 16\cdot 0,4 = 7,9;$

$D(X) = M(X^2) - (M(X))^2 = 7,9 - 2,5^2 = 1,65;$

$\sigma(X) = \sqrt{D(X)} = 1,2845;$

$M(Y) = 0\cdot 0,2 + 2\cdot 0,6 + 5\cdot 0,2 = 2,2;$

$M(Y^2) = 0\cdot 0,2 + 4\cdot 0,6 + 25\cdot 0,2 = 7,4;$

$D(Y) = M(Y^2) - (M(Y))^2 = 7,4 - 2,2^2 = 2,56;$

$\sigma(Y) = \sqrt{D(Y)} = 1,6;$

$$M(X\cdot Y) = 4\cdot 0,3 + 8\cdot 0,3 + 5\cdot 0,2 = 4,6; \quad K(X,Y) = M(X\cdot Y) - $$
$$= M(X)M(Y) = 4,6 - 2,5\cdot 2,2 = -0,9;$$

$$r(X,\ Y) = \frac{K(X,\ Y)}{\text{у}(X)\text{у}(Y)} = \frac{-0,9}{1,2845\cdot 1,6} = -0,438.$$