§1. Phân phối xác suất của vector ngẫu nhiên rời rạc
§2. Phân phối xác suất của vector ngẫu nhiên liên tục

Khái niệm vector ngẫu nhiên

- Một bộ có thứ tự n biến ngẫu nhiên $(X_1, \ldots, X_n)$ được gọi là một *vector ngẫu nhiên n chiều*.
- Vector ngẫu nhiên n chiều là liên tục hay rời rạc nếu các biến ngẫu nhiên thành phần là liên tục hay rời rạc. Chẳng hạn, một nhà máy sản xuất một loại sản phẩm, nếu xét đến kích thước của sản phẩm được đo bằng chiều dài X và chiều rộng Y thì ta có vector ngẫu nhiên hai chiều (X, Y). Còn nếu xét thêm cả chiều cao Z nữa thì ta có vector ngẫu nhiên ba chiều (X, Y, Z).
- Trong khuôn khổ của chương trình ta chỉ xét vector ngẫu nhiên hai chiều, thường được ký hiệu là (X, Y).

§1. PHÂN PHỐI XÁC SUẤT CỦA VECTOR NGẪU NHIÊN RỜI RẠC

1.1 Bảng phân phối xác suất đồng thời của (*X*, *Y*)

$X \backslash Y$	y_1	y_2	$\ldots$	y_j	$\ldots$	y_n	Tổng dòng
x_1	p_{11}	p_{12}	$\ldots$	p_{1j}	$\ldots$	p_{1n}	$p_{1\bullet}$
x_2	p_{21}	p_{22}	$\ldots$	p_{2j}	$\ldots$	p_{2n}	$p_{2\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\ldots$	p_{ij}	$\ldots$	p_{in}	$p_{i\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	p_{m1}	p_{m2}	$\ldots$	p_{mj}	$\ldots$	p_{mn}	$p_{m\bullet}$
Tổng cột	$p_{\bullet 1}$	$p_{\bullet 2}$	$\ldots$	$p_{\bullet j}$	$\ldots$	$p_{\bullet n}$	1

Trong đó $P\left(X = x_i; Y = y_j\right) = p_{ij}$ và $\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = 1.$

1.2. Phân phối xác suất thành phần (phân phối lề)

Từ bảng phân phối xác suất đồng thời của (X, Y) ta có:

• **Bảng phân phối xác suất của X**

X	x_1	x_2	$\ldots$	x_m
P	$p_{1\bullet}$	$p_{2\bullet}$	$\cdots$	$p_{m\bullet}$

Trong đó $p_{i\bullet} = p_{i1} + p_{i2} + \cdots + p_{in}$

(tổng dòng i của bảng phân phối xác suất đồng thời).

Kỳ vọng của X là:

$$\boxed{EX = x_1 p_{1\bullet} + x_2 p_{2\bullet} + \cdots + x_m p_{m\bullet}.}$$

- **Bảng phân phối xác suất của *Y***

Y	y_1	y_2	$\cdots$	y_n
P	$p_{\bullet 1}$	$p_{\bullet 2}$	$\cdots$	$p_{\bullet n}$

Trong đó $p_{\bullet j} = p_{1j} + p_{2j} + \cdots + p_{mj}$

(tổng cột j của bảng phân phối xác suất đồng thời).

Kỳ vọng của Y là:

$$\boxed{EY = y_1 p_{\bullet 1} + y_2 p_{\bullet 2} + \cdots + y_n p_{\bullet n}.}$$

VD 1. Phân phối xác suất đồng thời của vector ngẫu nhiên (X,Y) cho bởi bảng:

$X \backslash Y$	1	2	3
6	0,10	0,05	0,15
7	0,05	0,15	0,10
8	0,10	0,20	0,10

1) Tính $P\left(X=6\right)$ và $P\left(X \geq 7,\ Y \geq 2\right)$.

2) Lập bảng phân phối xs thành phần và tính EX, EY.

Giải

1) $P\left(X=6\right)=0{,}1+0{,}05+0{,}15=0{,}3.$

$$P\left(X \ge 7, Y \ge 2\right) = P\{(7,2)\} + P\{(7,3)\} + P\{(8,2)\} + P\{(8,3)\} = 0,15 + 0,1 + 0,2 + 0,1 = 0,55.$$

2) Bảng phân phối của X là:

X	6	7	8
P	0,3	0,3	0,4

$$EX = 6.0,3 + 7.0,3 + 8.0,4 = 7,1.$$

Bảng phân phối của Y là:

Y	1	2	3
P	0,25	0,40	0,35

$$EY = 1.0,25 + 2.0,4 + 3.0,35 = 2,1.$$

1.3. Phân phối xác suất có điều kiện

Từ công thức xác suất có điều kiện, ta có:

$$P\left(X=x_i \middle| Y=y_j\right)=\frac{P(X=x_i,\ Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{p_{\bullet j}},\ i=\overline{1,m}.$$

$$P\left(Y=y_j \middle| X=x_i\right)=\frac{P(X=x_i,\ Y=y_j)}{P(X=x_i)}=\frac{p_{ij}}{p_{i\bullet}},\ j=\overline{1,n}.$$

- Bảng phân phối xác suất của X với điều kiện $Y = y_j$:

X	x_1	x_2	...	x_m
$P\left(X=x_i \middle\vert Y=y_j\right)$	$\dfrac{p_{1j}}{p_{\bullet j}}$	$\dfrac{p_{2j}}{p_{\bullet j}}$	...	$\dfrac{p_{mj}}{p_{\bullet j}}$

Kỳ vọng của X với điều kiện $Y = y_j$ là:

$$\boxed{EX = \frac{1}{p_{\bullet j}}(x_1 p_{1j} + x_2 p_{2j} + \ldots + x_m p_{mj}).}$$

• Bảng phân phối xác suất của Y với điều kiện $X = x_i$:

Y	y_1	y_2	$\ldots$	y_n
$P\left(Y=y_j \middle\vert X=x_i\right)$	$\dfrac{p_{i1}}{p_{i\bullet}}$	$\dfrac{p_{i2}}{p_{i\bullet}}$	$\ldots$	$\dfrac{p_{in}}{p_{i\bullet}}$

Kỳ vọng của Y với điều kiện $X = x_i$ là:

$$\boxed{EY = \frac{1}{p_{i\bullet}}\left(y_1 p_{i1} + y_2 p_{i2} + \ldots + y_n p_{in}\right).}$$

<u>VD 2.</u> Cho bảng phân phối xs đồng thời của (X, Y):

$X \backslash Y$	1	2	3
6	0,10	0,05	0,15
7	0,05	0,15	0,10
8	0,20	0,10	0,10

1) Lập bảng phân phối xác suất của X với điều kiện $Y = 2$ và tính kỳ vọng của X.

2) Lập bảng phân phối xác suất của Y với điều kiện $X = 8$ và tính kỳ vọng của Y.

Giải. 1) Ta có:

$$P\left(X=6\,|\,Y=2\right)=\frac{0,05}{0,05+0,15+0,1}=\frac{1}{6}.$$

$$P\left(X=7\,|\,Y=2\right)=\frac{0,15}{0,05+0,15+0,1}=\frac{1}{2}.$$

$$P\left(X=8\,|\,Y=2\right)=\frac{0,1}{0,05+0,15+0,1}=\frac{1}{3}.$$

Bảng phân phối xác suất của X với điều kiện $Y=2$ là:

X	6	7	8
$P\left(X=x_i\,\|\,Y=2\right)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

$$EX = 6.\frac{1}{6} + 7.\frac{1}{2} + 8.\frac{1}{3} = \frac{43}{6}.$$

2) Bảng phân phối xác suất của Y với điều kiện $X = 8$:

Y	1	2	3
$P\left(Y=y_j \mid X=8\right)$	$0,50$	$0,25$	$0,25$

$$EY = 1.0,5 + 2.0,25 + 3.0,25 = 1,75.$$

VD 3. Cho vector ngẫu nhiên rời rạc (X, Y) có bảng phân phối xác suất đồng thời như sau:

(X,Y)	$(0;0)$	$(0;1)$	$(1;0)$	$(1;1)$	$(2;0)$	$(2;1)$
p_{ij}	$\frac{1}{18}$	$\frac{3}{18}$	$\frac{4}{18}$	$\frac{3}{18}$	$\frac{6}{18}$	$\frac{1}{18}$

1) Tính xác suất $P\left(X - Y = 1\right)$.

2) Tính xác suất $P(X > 0 \mid Y = 1)$.

3) Tính trung bình của X và Y.

4) Tính trung bình của Y khi $X = 1$.

Giải. 1) Ta có:

$$P(X - Y = 1) = P\{(1,0)\} + P\{(2,1)\} = \frac{4}{18} + \frac{1}{18} = \frac{5}{18}.$$

2) $P(X > 0 \mid Y{=}1) = P(X{=}1 \mid Y{=}1) + P(X{=}2 \mid Y{=}1)$

$$= \frac{P\{(1,1)\}}{P(Y=1)} + \frac{P\{(2,1)\}}{P(Y=1)} = \frac{4}{7}.$$

3) Bảng phân phối thành phần của X và Y là:

X	0	1	2
P	$\frac{4}{18}$	$\frac{7}{18}$	$\frac{7}{18}$

Y	0	1
P	$\frac{11}{18}$	$\frac{7}{18}$

Vậy $EX = 0.\frac{4}{18} + 1.\frac{7}{18} + 2.\frac{7}{18} = \frac{21}{18}$ và $EY = \frac{7}{18}$.

4) Bảng phân phối xác suất của Y khi $X = 1$ là:

Y	0	1
$P(Y=y_j \mid X=1)$	$\frac{4}{7}$	$\frac{3}{7}$

Vậy $EY = \dfrac{3}{7}$.

VD 4. Chi phí quảng cáo X (triệu đồng) và doanh thu Y (triệu đồng) của một công ty có bảng phân phối xác suất đồng thời như sau:

X \ Y	500 (400 – 600)	700 (600 – 800)	900 (800 – 1000)
30	$0,10$	$0,05$	0
50	$0,15$	$0,20$	$0,05$
80	$0,05$	$0,05$	$0,35$

Nếu doanh thu là 700 triệu đồng thì chi phí quảng cáo trung bình là:

A. 60,5 triệu đồng; B. 48,3333 triệu đồng;

C. 51,6667 triệu đồng; D. 76,25 triệu đồng.

§2. PHÂN PHỐI XÁC SUẤT CỦA VECTOR NGẪU NHIÊN LIÊN TỤC

2.1. Hàm mật độ đồng thời của (*X*, *Y*)

- Hàm hai biến $f(x,y) \ge 0$ xác định trên $\mathbb{R}^2$ được gọi là ***hàm mật độ*** của vector ngẫu nhiên (X,Y) nếu:

$$\boxed{\iint\limits_{\mathbb{R}^2} f(x,y)dxdy = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} f(x,y)dxdy = 1.}$$

- Xác suất của vector (X,Y) trên tập $D \subset \mathbb{R}^2$ là:

$$\boxed{P\{(X,Y) \in D\} = \iint\limits_{D} f(x,y)dxdy.}$$

2.2. Hàm mật độ thành phần

- Hàm mật độ của X là:

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

- Hàm mật độ của Y là:

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

Chú ý

Khi tìm hàm $f_X(x)$, ta lấy tích phân hàm $f(x,y)$ theo biến y và điều kiện x phải độc lập đối với y.

Tìm hàm $f_Y(y)$, ta làm tương tự.

Trung bình thành phần

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

2.3. Hàm mật độ có điều kiện

- Hàm mật độ có điều kiện của X khi biết $Y = y$ là:

$$\boxed{f_X\left(x\middle|y\right) = \frac{f(x,y)}{f_Y(y)}.}$$

- Hàm mật độ có điều kiện của Y khi biết $X = x$ là:

$$\boxed{f_Y\left(y\middle|x\right) = \frac{f(x,y)}{f_X(x)}.}$$

<u>VD 1.</u> Cho hàm $f(x,y)=\begin{cases}10x^2y, \text{ khi } 0\le y\le x\le 1,\\ 0, \text{ nơi khác.}\end{cases}$

1) Chứng tỏ vector (X,Y) có hàm mật độ là $f(x,y)$.

2) Tính xác suất $P\left(Y\ge \frac{1}{2}X\right)$.

3) Tìm hàm mật độ thành phần của X, Y.

4) Tìm hàm mật độ có điều kiện $f_X(x\mid y)$, $f_Y(y\mid x)$.

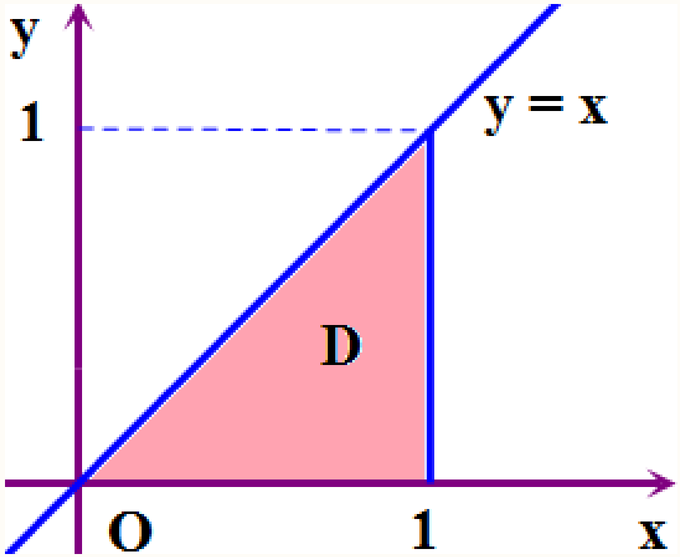
5) Tính xác xuất $P\left(Y<\frac{1}{8}\middle| X=\frac{1}{4}\right)$.

Giải

1) Đặt $D = \left\{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq x \leq 1\right\}.$

Chiếu D lên Ox, ta được:

$D = \left\{0 \leq x \leq 1,\ 0 \leq y \leq x\right\}.$



Suy ra:

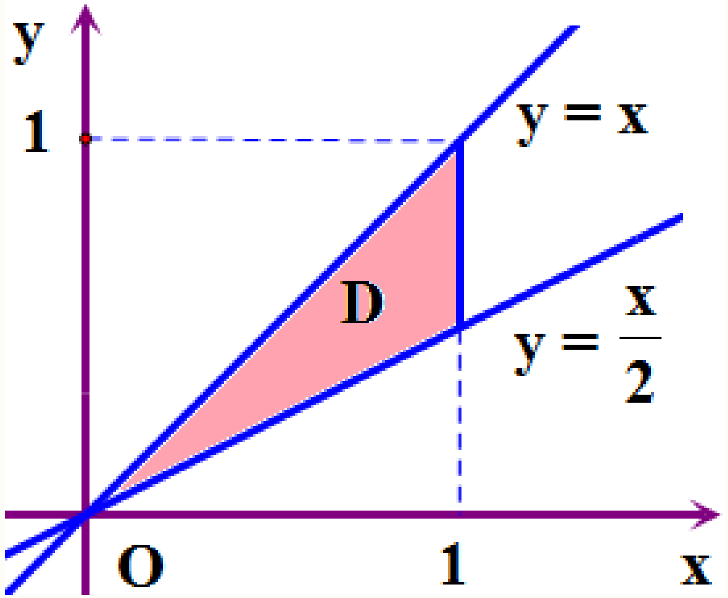
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)dxdy = \iint_D f(x, y)dxdy$$

$$= \int_0^1 5x^2 dx \int_0^x 2y dy = \int_0^1 5x^4 dx = 1 \blacksquare$$

2) Đặt $D = \left\{(x, y) : 0 \le y \le x \le 1,\ y \ge \dfrac{x}{2}\right\}.$

Chiếu D lên Ox, ta được:

$$D = \left\{0 \le x \le 1,\ \frac{x}{2} \le y \le x\right\}.$$



$$P\left(Y \ge \frac{1}{2}X\right) = \iint_D f(x, y)\,dxdy$$

$$= \int_0^1 5x^2dx \int_{\frac{x}{2}}^{x} 2ydy = \frac{3}{4}.$$

3) Khi $0 \le x \le 1$, ta có: $D = \left\{ 0 \le x \le 1,\ 0 \le y \le x \right\}$.

Suy ra:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_0^x 10x^2 y dy = 5x^4.$$

$$\text{Vậy } f_X(x) = \begin{cases} 5x^4, & \text{khi } 0 \le x \le 1, \\ 0, & \text{nôi khaùc.} \end{cases}$$

$$\text{Tương tự, } f_Y(y) = \begin{cases} \dfrac{10}{3} y(1 - y^3), & \text{khi } 0 \le y \le 1, \\ 0, & \text{nôi khaùc.} \end{cases}$$

4) Trên miền $D = \left\{(x, y) \in \mathbb{R}^2 : 0 \le y \le x \le 1\right\}$, ta có:

- $f_X(x \mid y) = \dfrac{f(x, y)}{f_Y(y)} = \dfrac{3x^2}{1 - y^3}.$

- $f_Y(y \mid x) = \dfrac{f(x, y)}{f_X(x)} = \dfrac{2y}{x^2}.$

Vậy: $f_X(x \mid y) = \begin{cases} \dfrac{3x^2}{1 - y^3}, & \text{khi } 0 \le y \le x \le 1, \\ 0, & \text{nôi khaù.} \end{cases}$

$$f_Y(y \mid x) = \begin{cases} \dfrac{2y}{x^2}, \text{ khi } (x,y) \in D, \\ 0, \text{ nôi khaù.} \end{cases}$$

5) Từ câu 4, ta có:

$$f_Y\left(y \middle| x = \frac{1}{4}\right) = \begin{cases} 32y, \text{ khi } 0 \le y \le \dfrac{1}{4}, \\ 0, \text{ nôi khaù.} \end{cases}$$

$$\text{Vậy } P\left(Y < \frac{1}{8} \middle| X = \frac{1}{4}\right) = \int_0^{\frac{1}{8}} 32y\,dy = \frac{1}{4}.$$

<u>VD 2.</u> Cho hàm mật độ đồng thời của vector (X, Y) là:

$$f(x,y) = \begin{cases} 6x, \text{ khi } 0 < x < 1;\ 0 < y < 1 - x, \\ 0, \text{ nôi khaùc.} \end{cases}$$

1) Tính trung bình thành phần của X, Y.

2) Tính xác suất $P\left(X > 0,3 \middle| Y = 0,5\right)$.

<u>Giải</u>

1) Tính $E\left\{f_X(x)\right\}$, $E\left\{f_Y(y)\right\}$.

- Đặt $D = \{0 < x < 1,\ 0 < y < 1 - x\}$, ta có:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_0^{1-x} 6x dy$$

$$= 6x(1-x),\ 0 < x < 1$$

$$\Rightarrow E\left\{f_X(x)\right\} = \int_{-\infty}^{+\infty} x f_X(x)dx = \int_0^1 x.6x(1-x)dx = \frac{1}{2}.$$

- Đặt $D = \left\{0 < y < 1,\ 0 < x < 1-y\right\}$, ta có:

$$f_Y(y) = \int_0^{1-y} 6x dx = 3(1-y)^2,\ 0 < y < 1$$

$$\Rightarrow E\left\{f_Y(x)\right\} = \int_{-\infty}^{+\infty} y f_Y(y)dy = \int_0^1 y.3(1-y)^2 dy = \frac{1}{4}.$$

2) Trên $D = \{0 < x < 1,\ 0 < y < 1 - x\}$, ta có:

$$f_X\left(x \middle| y\right) = \frac{f(x,y)}{f_Y(y)} = \frac{2x}{(1-y)^2}$$

$$\Rightarrow f_X\left(x \middle| y = 0,5\right) = \begin{cases} 8x, \text{ khi } 0 < x < 0,5 \\ 0, \text{ nôi khaù.} \end{cases}$$

Vậy $P\left(X > 0,3 \middle| Y = 0,5\right) = \int\limits_{0,3}^{0,5} 8x\,dx = 0,64.$

<u>VD 3.</u> Tuổi thọ X (năm) và thời gian chơi thể thao Y (giờ) có hàm mật độ đồng thời được cho như sau:

$$f(x,y) = \begin{cases} \dfrac{15}{4}x(1-y^2), & \text{khi } 0 \le y < x \le 1, \\ 0, & \text{nôi khaù.} \end{cases}$$

Thời gian chơi thể thao trung bình là:

A. 0,3125 giờ; B. 0,5214 giờ;

C. 0,1432 giờ; D. 0,4132 giờ.

Hướng dẫn

$$f_Y(y) = \frac{15}{4}\int_y^1 x(1-y^2)dx = \frac{15}{8}(1-y^2)^2,\ 0 \le y < 1$$

$$\Rightarrow E\left[f_Y(y)\right] = \frac{15}{8}\int_0^1 (1-y^2)^2.ydy = 0,3125 \Rightarrow A.$$