§1 LUẬT PPXS CỦA BIẾN NGẪU NHIÊN NHIỀU CHIỀU

1.1 Khái niệm cơ sở

*** Khái niệm biến ngẫu nhiên nhiều chiều**

Thực tế yêu cầu xét *đồng thời* nhiều thuộc tính có quan hệ với nhau của đối tượng mà ta quan tâm, ký hiệu $X_1, X_2, \ldots, X_n$.

Ta có thể dùng công cụ véc tơ của giải tích toán

$$\boldsymbol{X} = (X_1\ X_2\ \ldots\ X_n)^T \in \mathbb{R}^n$$

và coi $\boldsymbol{X}$ là một *biến ngẫu nhiên n chiều* hay *véc tơ ngẫu nhiên* (n chiều).

Để cho đơn giản, ta chỉ xét trường hợp $n = 2$ và $\boldsymbol{X} = (X\ \ Y)^T$, trong đó X, Y là các biến ngẫu nhiên 1 chiều.

*** Phân loại biến ngẫu nhiên hai chiều**

- Biến $\boldsymbol{X}$ được coi là *rời rạc* nếu cả 2 thành phần là biến rời rạc.
- Biến $\boldsymbol{X}$ được coi là *liên tục* nếu 2 thành phần là biến liên tục.
- Biến $\boldsymbol{X}$ có dạng hỗn hợp nếu 1 thành phần rời rạc và thành phần kia liên tục mà ta không xét ở đây.

Ta mở rộng khái niệm hàm PPXS cho biến ngẫu nhiên 2 chiều. Ký hiệu 2 sự kiện $A = \{X < x\}$ và $B = \{Y < y\}$ }.

*** Định nghĩa.** *Hàm phân phối xác suất của biến 2 chiều* $\boldsymbol{X} =$ $(X \ \ Y)^T$ được xác định như sau

$$F(x, y) = P(AB) = P(X < x, Y < y); x, y \ \in \mathbb{R}.(1)$$

Người ta còn dùng thuật ngữ *hàm phân phối xác suất đồng thời của 2 biến X, Y*.

Mở rộng các tính chất của hàm PPXS của chương 2, ta có

*** Tính chất của $\boldsymbol{F(x, y)}$**

(i) $1 \geq F(x, y) \geq 0$;

(ii) $F(x, y)$ là hàm không giảm theo từng đối số;

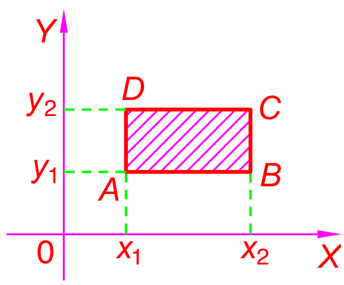
(iii) Với $x_1 \leq x_2, y_1 \leq y_2$

$$P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2)$$

$$= F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1);$$

(iv) $F(+\infty, +\infty) = 1$; $F(-\infty, y_1) = F(x_1, -\infty) = 0$.

Tính chất (iii) có ý nghĩa là XS để điểm ngẫu nhiên (X, Y) rơi vào miền chữ nhật ABCD.



*** Phân phối biên**

Để ý rằng

$$F(x, +\infty) = P(X < x, Y < +\infty) = P(X < x) = F_1(x);$$

$$F(+\infty, y) = P(X < +\infty, Y < y) = P(Y < y) = F_2(y);$$

là các phân phối riêng của từng thành phần X, Y tương ứng; chúng được gọi là các *phân phối biên* của biến hai chiều $\boldsymbol{X}$. Đó cũng chính là các phân phối một chiều đã biết ở chương II.

*** Khái niệm độc lập**

Hai biến X, Y được gọi là *độc lập* nếu (dùng định nghĩa ở chương I: $P(AB) = P(A)P(B) \Leftrightarrow A, B$ độc lập)

$$F(x, y) = F_1(x)F_2(y).$$

1.2 PPXS của biến ngẫu nhiên hai chiều rời rạc

*** Định nghĩa.** *Bảng phân phối xác suất của biến ngẫu nhiên 2 chiều* $\boldsymbol{X} = (X \;\; Y)^T$ là

$x \setminus y$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_m	$\sum_j$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	p_{1m}	$p_1(x_1)$
x_1	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	p_{2m}	$p_1(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	p_{im}	$p_1(x_i)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nj}	$\cdots$	p_{nm}	$p_1(x_n)$
$\sum_i$	$p_2(y_1)$	$p_2(y_2)$	$\cdots$	$p_2(y_j)$	$\cdots$	$p_2(y_m)$	1

trong đó $p_{ij} = P(X = x_i, Y = y_j)$ là *xác suất đồng thời* để X và Y nhận các giá trị tương ứng.

Tương tự trường hợp 1 chiều ta có thể xác định *hàm xác suất đồng thời* $p(x, y)$ sao cho $p(x_i, y_j) = p_{ij}$; $i = \overline{1, n}, j = \overline{1, m}$.

*** Tính chất của $p(x, y)$**

(i) $p_{ij} \geq 0, i = \overline{1, n}, j = \overline{1, m}$ hay $p(x, y) \geq 0 \ \forall\, x, y$;

(ii) $\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = 1$ hay $\sum_{\forall x, y} p(x, y) = 1$.

*** Hàm phân phối xác suất đồng thời**

$$F(x, y) = \sum_{x_i < x} \sum_{y_j < y} p_{ij}.$$

*** Hàm xác suất biên**

$$P(X = x_i) = p_1(x_i) = \sum_j p_{ij}, i = \overline{1, n};$$

$$P(Y = y_j) = p_2(y_j) = \sum_i p_{ij}, j = \overline{1, m}.$$

*** <u>Thí dụ 1</u>**. Từ bảng PPXS đồng thời của X và Y

$x \backslash y$	1	2	3
1	0,10	0,25	0,15
2	0,15	0,00	0,35

Tìm các PP biên của X và Y, sau đó tính $F(2, 3)$.

Giải. Lấy tổng hàng và tổng cột của bảng PP đồng thời, ta có các bảng PP biên tương ứng:

x	1	2
$p_1(x)$	0,50	0,50

y	0	1	2
$p_2(y)$	0,25	0,25	0,50

Việc tính $F(2, 3)$ dựa vào công thức hàm PP đồng thời

$$F(2,3) = \sum_{x_i<2} \sum_{y_j<3} p_{ij} = p_{11} + p_{12} = 0{,}35.$$

*** Khái niệm độc lập**

Hai biến rời rạc X và Y được gọi là độc lập, nếu với mọi cặp giá trị x_i, y_j, ta luôn có

$$p_{ij} = p_1(x_i)\, p_2(y_j);\ i = \overline{1,n}, j = \overline{1,m}.$$

Trong thí dụ 1 ta thấy $p_{11} = 0{,}1 \neq p_1(1)\, p_2(1) = 0{,}125 \Rightarrow$ hai biến X, Y không độc lập.

*** Phân phối có điều kiện**

Từ chương I ta có

$$P(X = x_i|Y = y_k) = \frac{P(X = x_i; Y = y_k)}{P(Y = y_k)}, i = \overline{1, n}.$$

Công thức này cho phép ta xác định XS của phân phối có điều kiện của X, biết Y nhận một giá trị cụ thể y_k.

*** <u>Thí dụ 2</u>**. Tìm phân phối có điều kiện của X, biết $Y = 1$ trong bài toán thí dụ 1.

Giải. Theo công thức XS có điều kiện ở trên

$$P(X = 1|Y = 1) = \frac{P(X=1; Y=1)}{P(Y=1)}, = \frac{p_{11}}{p_2(1)} = \frac{0,1}{0,25} = 0,4;$$

$$P(X = 2|Y = 1) = \frac{p_{21}}{p_2(1)} = \frac{0,15}{0,25} = 0,6.$$

Từ đó bảng PP có điều kiện của X biết $Y = 1$ là

x	1	2
$p_1(x\|Y = 1)$	0,4	0,6

Có thể tổng quát hoá PP có điều kiện với bộ điều kiện C_Y nào đó

$$P(X = x|C_Y) = \frac{P(X = x; C_Y)}{P(C_Y)}.$$

Chẳng hạn nếu ta biết $C_Y = \{ y_1 \leq Y \leq y_2\}$, thì

$$P(X = x| y_1 \leq Y \leq y_2) = \frac{P(X = x;\ y_1 \leq Y \leq y_2)}{P(y_1 \leq Y \leq y_2)}.$$

1.3 PPXS của biến ngẫu nhiên hai chiều liên tục

Ta đưa ra định nghĩa của hàm mật độ hai chiều:

*** Định nghĩa.** Nếu hàm phân phối $F(x,y)$ của biến ngẫu nhiên hai chiều $\boldsymbol{X} = (X \quad Y)^T$ được biểu diễn dưới dạng

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

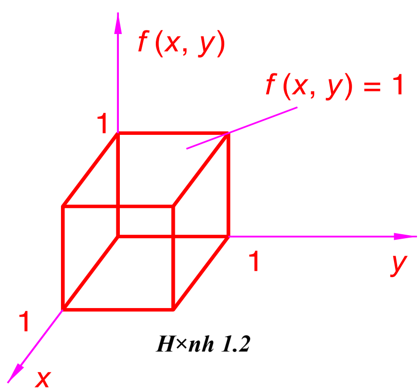
thì hàm $f(x,y)$ được gọi là *hàm mật độ của biến* $\boldsymbol{X}$ hay *hàm mật độ đồng thời của X và Y.*

Về mặt ý nghĩa hình học hàm $z = f(x,y)$ vẽ ra một mặt cong trong $\mathbb{R}^3$ và được gọi là mặt mật độ PPXS. Nếu $F(x,y)$ khả vi theo cả hai biến thì

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

*** <u>Thí dụ 3</u>.** Cho hàm mật độ đồng thời của X và Y là $f(x,y) = 1$, với $0 \leq x, y \leq 1$. Vẽ hàm $f(x,y)$ và tính hàm PP đồng thời $F(x,y)$.

Giải. Mặt cong mật độ ở đây là phần mặt phẳng trong mặt $z = 1$, tức là $f(x,y) \neq 0$ với các (x,y) thuộc hình vuông $[0;1] \times [0;1]$. Hàm PP đồng thời $F(x,y)$ được



H×nh 1.2

tính sau đây:

$$F(x,y)=\begin{cases} 0, & x\le 0 \text{ hoặc } y\le 0;\\ xy, & 0<x\le 1,\ 0<y\le 1;\\ x, & 0<x\le 1,\ y>1;\\ y, & x>1,\ 0<y\le 1;\\ 1, & x>1,\quad y>1. \end{cases}$$

Hàm PP có dạng rất phức tạp, nên người ta hay sử dụng hàm mật độ. Đây là thí dụ về *PP đều hai chiều*, tổng quát hoá PP đều liên tục đã xét ở chương II.

*** Tính chất của $f(x,y)$**

(i) $f(x,y)\ge 0$;

(ii) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dxdy=1$;

(iii) $P(\boldsymbol{X}\in\mathcal{D})=\iint_{\mathcal{D}} f(x,y)\,dxdy$.

Ở tính chất (iii) $\mathcal{D}$ là một miền nào đó trong mặt xOy và tích phân cho thể tích một hộp chữ nhật cong có đáy trên nằm trong mặt $z=f(x,y)$, đáy dưới là miền $\mathcal{D}$ (thuộc mặt $z=0$).

*** Hàm mật độ biên**

$$f_1(x)=\int_{-\infty}^{+\infty} f(x,y)dy,\qquad f_2(y)=\int_{-\infty}^{+\infty} f(x,y)dx.$$

Để ý $f_1(x)=\dfrac{\partial F_1}{\partial x}$ và là mật độ của biến X, tương tự đối với $f_2(y)$.

* **<u>Thí dụ 4</u>**. Tìm các hàm mật độ biên của biến $\boldsymbol{X} = (X,\ Y)^T$ có hàm mật độ của X và Y là

$$f(x,y) = \frac{1}{\pi^2(1+x^2)(1+y^2)}\,, x, y \in \mathbb{R}.$$

Giải. Từ công thức mật độ biên

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\infty}^{+\infty} \frac{1}{\pi^2(1+x^2)(1+y^2)}\ dy = \frac{1}{\pi(1+x^2)}\,.$$

Do tính đối xứng ta có $f_2(y) = \dfrac{1}{\pi(1+y^2)}$.

*** Khái niệm độc lập**

Hai biến ngẫu nhiên X và Y được coi là *độc lập*, nếu

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

*** Hàm mật độ có điều kiện**

$$\varphi(x|y) = \frac{f(x,y)}{f_2(y)} = \frac{f(x,y)}{\int_{-\infty}^{+\infty} f(x,y)dx}\,.$$

Tương tự $\psi(y|x)$ hàm mật độ có điều kiện của Y đối với $X = x$ và nó sẽ bằng $f(x,y)/\ f_1(x)$. Các mật độ có điều kiện có tính chất như các mật độ không điều kiện.

* **<u>Thí dụ 5</u>**. Cho hàm mật độ đồng thời của X và Y

$$f(x,y) = \begin{cases} x+y, & 0 \leq x, y \leq 1; \\ 0, & \text{với } x, y \text{ khác.} \end{cases}$$

Xác định các hàm mật độ có điều kiện.

Giải. Để dùng được công thức định nghĩa, ta phải tính các hàm mật độ biên $f_1(x), f_2(y)$:

$$f_1(x) = \int_0^1 (x+y)\, dy = x + 0{,}5,\ 0 \leq x \leq 1;$$

tương tự $f_2(y) = y + 0{,}5,\ 0 \leq y \leq 1$. Từ đó với $0 \leq y \leq 1$

$$\varphi(x|y) = \frac{f(x,y)}{f_2(y)} = \begin{cases} \dfrac{x+y}{y+0{,}5}, 0 \leq x \leq 1; \\ 0, x \notin [0,1]; \end{cases}$$

và với $0 \leq x \leq 1$

$$\psi(y|x) = \begin{cases} \dfrac{x+y}{x+0{,}5}, 0 \leq y \leq 1 \\ 0, y \notin [0,1]. \end{cases}$$

Để ý $\varphi(x|y)$ là hàm của cả x và y. Ngoài ra

$$f(x,y) = f_1(x)\, \psi(y|x) = f_2(y)\, \varphi(x|y);$$

và điều kiện X, Y độc lập cho thấy

$$\varphi(x|y) = f_1(x); \psi(y|x) = f_2(y).$$

Có thể viết công thức tổng quát hơn

$$\varphi(x|y_1 \leq Y \leq y_2) = \frac{\int_{y_1}^{y_2} f(x,y) dy}{\int_{-\infty}^{+\infty} dx \int_{y_1}^{y_2} f(x,y) dy}.$$

Cuối cùng ta viết các công thức cho *hàm phân phối có điều kiện*

$$\phi(x|y) = \frac{\int_{-\infty}^{x} f(u,y)du}{\int_{-\infty}^{+\infty} f(x,y)dx};$$

$$\phi(x|y_1 \le Y \le y_2) = \frac{\int_{-\infty}^{x} du \int_{y_1}^{y_2} f(u,y)dy}{\int_{-\infty}^{+\infty} dx \int_{y_1}^{y_2} f(x,y)dy}.$$

* **<u>Thí dụ 6</u>**. Lấy hàm mật độ của thí dụ 5, hãy tính các hàm mật độ có điều kiện của X, biết a) $Y = 0,5$; b) $Y \in [0,5; 0,75]$.

Giải. a) Trong trường hợp biết $Y = 0,5$

$$\varphi(x|Y = 0,5) = x + 0,5,\ 0 \le x \le 1.$$

b) Trường hợp biết $Y \in [0,5; 0,75]$, với $0 \le x \le 1$

$$\varphi(x|0,5 \le Y \le 0,75) = \frac{\int_{0,5}^{0,75}(x+y)dy}{\int_0^1 dx \int_{0,5}^{0,75}(x+y)dy} = \frac{8x+5}{9}.$$

Để ý nếu X $\notin$ [0;1], cả hai mật độ có điều kiện đều bằng 0.

* **<u>Thí dụ 7</u>**. Cho $X,\ Y$ có hàm mật độ đồng thời

$$f(x,y) = \begin{cases} \frac{3}{\pi}\left(1 - \sqrt{x^2+y^2}\right), & x^2+y^2 \le R^2; \\ 0, & x^2+y^2 > R^2. \end{cases}$$

Tính $P(X^2 + Y^2 \le 1/4)$.

Giải. Gọi $\mathcal{D}$ là mặt tròn $x^2 + y^2 \le 1/4$. Khi đó

$$P(X^2 + Y^2 \le 1/4) = \frac{3}{\pi}\iint_{\mathcal{D}}\left(1 - \sqrt{x^2 + y^2}\right)dxdy$$

$$= \frac{3}{\pi}\int_0^{2\pi}\int_0^{1/2}(1 - r)\, rdrd\varphi = 1/2.$$

*** <u>Thí dụ 8</u>**. Cho biến 2 chiều $\boldsymbol{X} = (X,\ Y)^T$ có PP đều trên một mặt tròn $\mathcal{D}$: $x^2 + y^2 \le R^2$. Tìm các hàm mật độ (đồng thời, biên, có điều kiện).

Giải. Biến 2 chiều có PP đều trên $\mathcal{D}$: $f(x, y) = c$, $(x, y) \in \mathcal{D}$.

Từ đó $$c\iint_{\mathcal{D}} dxdy = 1 \Rightarrow c = \frac{1}{\pi R^2}$$

và $$f(x, y) = \begin{cases} \dfrac{1}{\pi R^2},\ x^2 + y^2 \le R^2; \\ 0,\ \ x^2 + y^2 > R^2. \end{cases}$$

Do tính đối xứng, ta chỉ tìm $f_2(y)$ và $\varphi(x|y)$:

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y)dx$$

$$= \begin{cases} \dfrac{1}{\pi R^2}\displaystyle\int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} dx,\ |y| \le R; \\ 0,\qquad |y| > R; \end{cases}$$

$$\Rightarrow\quad f_2(y) = \begin{cases} \dfrac{2\sqrt{R^2-y^2}}{\pi R^2},\ |y| \le R; \\ 0,\qquad |y| > R; \end{cases}$$

$$\varphi(x|y) = \frac{f(x,y)}{f_2(y)} = \begin{cases} \dfrac{1}{2\sqrt{R^2 - y^2}}, & x^2 + y^2 \le R^2; \\ 0, & x^2 + y^2 > R^2. \end{cases}$$

$$\Longrightarrow \qquad f_1(x) = \begin{cases} \dfrac{2\sqrt{R^2 - x^2}}{\pi R^2}, & |x| \le R; \\ 0, & |x| > R; \end{cases}$$

$$\psi(y|x) = \frac{f(x,y)}{f_1(x)} = \begin{cases} \dfrac{1}{2\sqrt{R^2 - x^2}}, & x^2 + y^2 \le R^2; \\ 0, & x^2 + y^2 > R^2. \end{cases}$$

BÀI TẬP

1. Cho X và Y là hai biến ngẫu nhiên độc lập có PPXS:

x	0	1	2	3
$p(x)$	0,3	0,1	0,4	0,2

y	0	1	2
$p(y)$	0,2	0,5	0,3

a) Tìm PPXS đồng thời của X và Y.
b) Tính $P(X > Y)$.

2. Cho X và Y có bảng PPXS đồng thời

x \ y	1	2	3
1	0,12	0,15	0,03
2	0,28	0,35	0,07

a) X và Y có độc lập không?
b) Tìm luật PPXS của $Z = XY$, sau đó tính EZ.

3. Hai biến ngẫu nhiên X và Y có bảng PPXS đồng thời

x \ y	1	2	3
1	0,20	0,05	3λ
2	λ	0,20	0,10
3	0,05	0,00	0,25

a) Xác định tham số λ.
b) Lập các bảng phân phối biên.
c) Lập bảng PP có điều kiện $P(X = x|\, Y = 3)$.

3. Bảng PPXS đồng thời của số lỗi vẽ màu X và số lỗi đúc Y của một loại sản phẩm nhựa ở một công ty được cho bởi

x \ y	0	1	2
0	0,59	0,06	0,03
1	0,10	0,05	0,01
2	0,06	0,05	0,01
3	0,02	0,02	0,00

Hai biến có độc lập không? Tính XS để tổng số lỗi vẽ màu và lỗi đúc lớn hơn 4. Nếu ta biết trên sản phẩm có 2 lỗi vẽ màu thì XS để không có lỗi đúc bằng bao nhiêu?

4. Cho luật PP của một biến 2 chiều như sau (X, Y)

x \ y	2	3	5
1	0,1	0	0,1
4	0,2	0,5	0,1

Tìm luật PPXS của các hàm $X+Y$ và XY, sau đó tính các kỳ vọng và phương sai.

5. Hai máy tự động hoạt động độc lập, với XS để từng máy sản xuất ra sản phẩm tốt tương ứng là p_1 và p_2. Giả sử mỗi máy làm được 2 sản phẩm và gọi X và Y là số sản phẩm tốt của các máy tương ứng. Tìm PPXS đồng thời của hai biến X, Y.

6. Biến ngẫu nhiên 2 chiều có hàm mật độ đồng thời

$$f(x,y) = a(x^2 + y^2), \text{ nếu } x^2 + y^2 \leq 4.$$

Tìm hệ số a, sau đó tìm các mật độ biên và mật độ có điều kiện.

7. Cho hàm mật độ đồng thời của X và Y

$$f(x,y) = a\, e^{-\frac{1}{2}(x^2+2xy+5y^2)}.$$

Xác định hằng số a, sau đó tìm các mật độ biên và mật độ có điều kiện.

8. Cho X và Y là hai biến ngẫu nhiên có phân phối đều trên $(0; 2)$. Tính XS đồng thời $P(XY \leq 1, Y \leq 2X, X \leq 2Y)$.

9. Cho X và Y là hai biến ngẫu nhiên có hàm mật độ đồng thời $f(x,y) = a\left(1 - \sqrt{x^2 + y^2}\right)$, nếu $x^2 + y^2 \leq 1$. Xác định a, sau đó tính $P(X^2 + Y^2 \leq 1/4)$.

10. Một điểm A rơi ngẫu nhiên vào một hình vuông $\mathcal{D}$ có độ dài cạnh bằng 1 và giả sử (X, Y) là toạ độ của A. Biết rằng hàm mật độ độ đồng thời của X và Y là $f(x, y) \ = 1$, nếu $(x, y) \in \mathcal{D}$. Tính XS để khoảng cách từ A đến cạnh gần nhất của hình vuông không vượt quá 0,3.

11. Lấy ngẫu nhiên một điểm trong tam giác 3 đỉnh O(0,0), A(0,1) và B(4,0). Gọi X và Y là toạ độ của điểm đó.
    a) Tìm hàm mật độ đồng thời của X và Y.
    b) Tìm các mật độ biên.
    c) X và Y có độc lập không?

12. A và B đến cửa hàng cùng một thời điểm. Giả sử thời gian mua hàng của A có phân phối đều trong (10;20) và thời gian mua hàng của B có phân phối đều trong (15;25). Tính XS để A kết thúc mua trước B.