3.1 KHÁI NIỆM VÉC TƠ NGẪU NHIÊN

3.1.1 Định nghĩa và phân loại

Một véc tơ ngẫu nhiên n chiều là một bộ có thứ tự $(X_1, X_2, ..., X_n)$ với các thành phần $X_1, X_2, ..., X_n$ là các biến ngẫu nhiên xác định trong cùng một phép thử.

Ta ký hiệu véc tơ ngẫu nhiên hai chiều là (X, Y), trong đó X là biến ngẫu nhiên thành phần thứ nhất và Y là biến ngẫu nhiên thành phần thứ hai

Véc tơ ngẫu nhiên n chiều là rời rạc hoặc liên tục nếu tất cả các biến ngẫu nhiên thành phần là rời rạc hoặc liên tục

3.1.2 Hàm phân bố xác suất

Hàm phân bố xác suất của véc tơ ngẫu nhiên $\left(X_1, X_2, \ldots, X_n\right)$ hay còn được gọi là hàm phân bố xác suất đồng thời của các biến ngẫu nhiên $X_1, X_2, \ldots, X_n$

$$F_{X_1 \ldots X_n}(x_1, x_2, \ldots, x_n) = P\left\{X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n\right\}$$

Các tính chất của hàm phân bố xác suất đồng thời

1. $0 \le F_{X_1 \ldots X_n}(x_1, \ldots, x_n) \le 1$
2. $\lim\limits_{x_k \to -\infty} F_{X_1 \ldots X_n}(x_1, \ldots, x_n) = 0,$ với k nào đó thuộc $\{1, \ldots, n\}$
3. $\lim\limits_{(x_1, \ldots, x_n) \to (\infty, \ldots, \infty)} F_{X_1 \ldots X_n}(x_1, \ldots, x_n) = 1$
4. $F_{X_1 \ldots X_n}(x_1, \ldots, x_n)$ không giảm theo từng biến

5. $\lim\limits_{x_1 \to \infty} F_{X_1 \dots X_n}(x_1, \dots, x_n) = F_{X_2 \dots X_n}(x_2, \dots, x_n)$

Tương tự nếu lấy giới hạn của hàm phân bố xác suất đồng thời của $X_1, X_2, \dots, X_n$ khi biến x_k tiến đến vô cùng, với k nào đó thuộc $\{1, \dots, n\}$, thì được hàm phân bố xác suất đồng thời của $n-1$ biến ngẫu nhiên còn lại $X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n$

Đặc biệt nếu $F_{XY}(x, y)$ là hàm phân bố xác suất của véc tơ ngẫu nhiên hai chiều (X,Y) thì

$$\lim_{y \to \infty} F_{XY}(x, y) = P\{X \le x\} = F_X(x)$$

$$\lim_{x \to \infty} F_{XY}(x, y) = P\{Y \le y\} = F_Y(y)$$

$$P\{x_1 < X \le x_2, y_1 < Y \le y_2\} = F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1)$$

Ví dụ 3.3: Cho véc tơ ngẫu nhiên (X,Y) có hàm phân bố xác suất

$$F_{XY}(x,y)=\begin{cases}(1-e^{-\alpha x})(1-e^{-\beta y}) & x\geq 0, y\geq 0;\alpha,\beta>0\\ 0 & \text{neu nguoc lai}\end{cases}$$

Do đó các hàm phân bố xác suất biên

$$F_X(x)=F_{XY}(x,\infty)=\begin{cases}1-e^{-\alpha x} & x\geq 0\\ 0 & x<0\end{cases}$$

$$F_Y(y)=F_{XY}(\infty,y)=\begin{cases}1-e^{-\beta y} & y\geq 0\\ 0 & y<0\end{cases}$$

Các xác suất

$$P\{X \le 1, Y \le 1\} = F_{XY}(1,1) = (1-e^{-\alpha})(1-e^{-\beta})$$

$$P\{X \le 1\} = F_X(1) = 1-e^{-\alpha}$$

$$P\{Y > 1\} = 1 - P\{Y \le 1\} = 1 - F_Y(1) = e^{-\beta}$$

Áp dụng luật De Morgan ta có

$$\overline{\{X > x\}\{Y > y\}} = \overline{\{X > x\}} \cup \overline{\{Y > y\}} = \{X \le x\} \cup \{Y \le y\}$$

$$P\left(\overline{\{X > x\}\{Y > y\}}\right) = P\left(\{X \le x\} \cup \{Y \le y\}\right) = P\{X \le x\} + P\{Y \le y\} - P\{X \le x; Y \le y\}$$

$$= F_X(x) + F_Y(y) - F_{XY}(x,y)$$

$$= (1-e^{-\alpha x}) + (1-e^{-\beta y}) - (1-e^{-\alpha x})(1-e^{-\beta y}) = 1 - e^{-\alpha x}e^{-\beta y}$$

Vậy

$$P\{X > x, Y > y\} = 1 - P\left(\overline{\{X > x\}\{Y > y\}}\right) = e^{-\alpha x}e^{-\beta y}$$

3.2 BẢNG PHÂN BỐ XÁC SUẤT CỦA VÉC TƠ NGẪU NHIÊN RỜI RẠC HAI CHIỀU

Tương tự trường hợp biến ngẫu nhiên rời rạc, quy luật phân bố xác suất của véc tơ ngẫu nhiên rời rạc 2 chiều có thể được xác định thông qua hàm khối lượng xác suất đồng thời hoặc bảng phân bố xác suất đồng thời

3.2.1 Hàm khối lượng xác suất đồng thời

$$p_{XY}(x_i, y_j) = P\left\{X = x_i, Y = y_j\right\}$$

thỏa mãn điều kiện $\begin{cases} p_{X,Y}(x_i, y_j) \geq 0,\ \forall i = 1,\dots,n,\ j = 1,\dots,m \\ \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{m} p_{X,Y}(x_i, y_j) = 1 \end{cases}$

Hàm phân bố xác suất đồng thời $F_{X,Y}(x, y) = \displaystyle\sum_{x_i \leq x}\sum_{y_j \leq y} p_{X,Y}(x_i, y_j)$

$X \backslash Y$	y_1	y_2	$\ldots$	y_j	$\ldots$	y_m
x_1	$p_{XY}(x_1, y_1)$	$p_{XY}(x_1, y_2)$	$\ldots$	$p_{XY}(x_1, y_j)$	$\ldots$	$p_{XY}(x_1, y_m)$
x_2	$p_{XY}(x_2, y_1)$	$p_{XY}(x_2, y_2)$	$\ldots$	$p_{XY}(x_2, y_j)$	$\ldots$	$p_{XY}(x_2, y_m)$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
x_i	$p_{XY}(x_i, y_1)$	$p_{XY}(x_i, y_2)$	$\ldots$	$p_{XY}(x_i, y_j)$	$\ldots$	$p_{XY}(x_i, y_m)$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\ldots$	$\vdots$
x_n	$p_{XY}(x_n, y_1)$	$p_{XY}(x_n, y_2)$	$\ldots$	$p_{XY}(x_n, y_j)$	$\ldots$	$p_{XY}(x_n, y_m)$

3.2.2 Hàm khối lượng xác suất biên

$$p_Y(y_j) = P\left\{Y = y_j\right\} = \sum_{i=1}^{n} P\left\{X = x_i, Y = y_j\right\} = \sum_{i=1}^{n} p_{XY}(x_i, y_j);\ j = \overline{1,m}$$

$$p_X(x_i) = P\left\{X = x_i\right\} = \sum_{j=1}^{m} P\left\{X = x_i, Y = y_j\right\} = \sum_{j=1}^{m} p_{XY}(x_i, y_j); i = \overline{1,n}$$

Bảng phân bố xác suất biên

Từ bảng phân bố xác suất đồng thời của (X,Y), nếu ta cộng các xác suất theo cột thì ta được các xác suất tương ứng với các giá trị của Y.

Nếu ta cộng các xác suất theo hàng ta được các xác suất tương ứng với giá trị của X .

Ví dụ 3.4: Gieo 3 đồng tiền cân đối **A**, **B**, **C**. Gọi X là số mặt ngửa xuất hiện của 2 đồng tiền **A**, **B** và Y là số mặt ngửa xuất hiện của cả 3 đồng tiền **A**, **B**, **C**. Hãy lập bảng phân bố xác suất đồng thời của X, Y.

Các kết quả đồng khả năng

$\mathscr{A}$	N	N	N	N	S	S	S	S
$\mathscr{B}$	N	N	S	S	N	N	S	S
$\mathscr{C}$	N	S	N	S	N	S	N	S
X	2	2	1	1	1	1	0	0
Y	3	2	2	1	2	1	1	0

Bảng phân bố xác suất đồng thời của X và Y

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

Bảng phân bố xác suất của hai biến ngẫu nhiên thành phần

X	0	1	2
P	2/8	4/8	2/8

Y	0	1	2	3
P	1/8	3/8	3/8	1/8

Ví dụ 3.5: Có hai hộp, mỗi hộp đựng 6 bi.

Hộp I có 1 bi mang số 1, 2 bi mang số 2, 3 bi mang số 3

Hộp II có 2 bi mang số 1, 3 bi mang số 2, 1 bi mang số 3

Lấy ngẫu nhiên từ mỗi hộp 1 bi

Gọi X, Y lần lượt là số ghi trên bi rút được từ hộp I và hộp II

Bảng phân bố xác suất đồng thời của X, Y

X \ Y	1	2	3
1	2/36	3/36	1/36
2	4/36	6/36	2/36
3	6/36	9/36	3/36

3.3 HÀM MẬT ĐỘ XÁC SUẤT CỦA VÉC TƠ NGẪU NHIÊN LIÊN TỤC

Hàm n biến $f_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)\geq 0$ thoả mãn

$$F_{X_1X_2\ldots X_n}(x_1,x_2,\ldots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f_{X_1X_2\ldots X_n}(t_1,t_2,\ldots,t_n)\,dt_1dt_2\ldots dt_n$$

là hàm mật độ xác suất của véc tơ ngẫu nhiên liên tục $(X_1,X_2,\ldots,X_n)$

hoặc hàm mật độ xác suất đồng thời của các biến ngẫu nhiên

$X_1,X_2,\ldots,X_n$

Tính chất của hàm mật độ xác suất

1. $f_{XY}(x,y) \geq 0$ với mọi (x,y) và $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$

2. $P\{(X,Y) \in A\} = \iint_{(x,y)\in A\cap R_{XY}} f_{XY}(x,y)dxdy$ với mọi $A \subset R^2$,

R_{XY} là miền giá trị của (X,Y)

3) $f(x,y) = \begin{cases} \dfrac{\partial^2}{\partial x \partial y} F(x,y) & \text{nếu tồn tại đạo hàm tại } (x,y) \\ 0 & \text{nếu ngược lại} \end{cases}$

3.3.2 Hàm mật độ xác suất biên

Hàm mật độ xác suất của biến ngẫu nhiên thành phần X

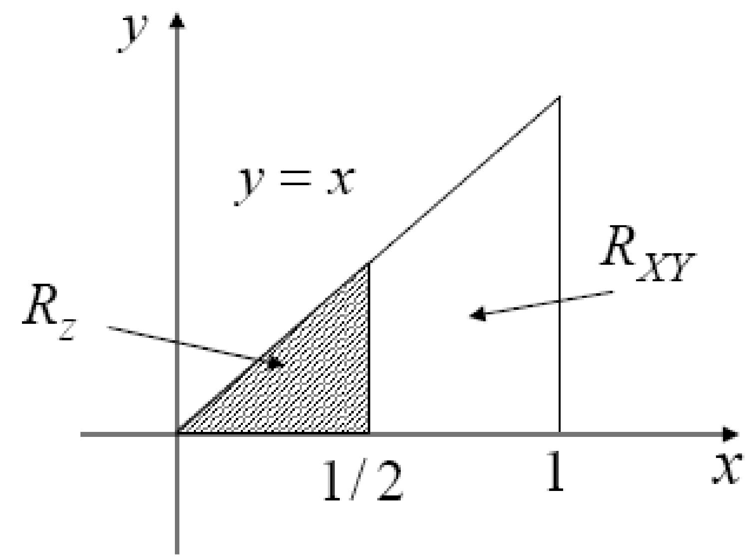
$$\int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = f_X(x)$$

Hàm mật độ xác suất của biến ngẫu nhiên thành phần Y

$$\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = f_Y(y)$$

Ví dụ 3.7: Cho véc tơ ngẫu nhiên (X,Y) có hàm mật độ xác suất xác định như sau

$$f_{XY}(x,y)=\begin{cases} k & \text{neu } 0<y\le x<1 \\ 0 & \text{neu nguoc lai} \end{cases}$$



Miền giá trị của X , Y là tam giác R_{XY}

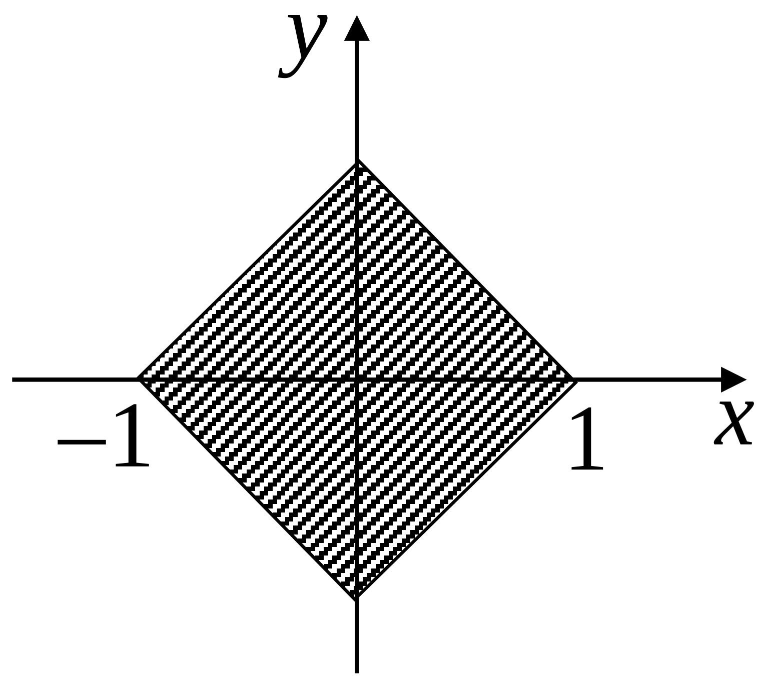
$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy=k\frac{1}{2} \Rightarrow k=2$$

$$f_X(x)=\begin{cases} \int_0^x 2dy=2x & \text{neu } 0<x<1 \\ 0 & \text{neu nguoc lai} \end{cases} \quad f_Y(y)=\begin{cases} \int_y^1 2dx=2(1-y) & \text{neu } 0<y<1 \\ 0 & \text{neu nguoc lai} \end{cases}$$

$$P\{0<X\le 1/2;0<Y\le 1/2\}=\iint_{R_z} f_{XY}(x,y)dxdy=2\left(\frac{1}{8}\right)=\frac{1}{4}$$

Ví dụ 3.8: Cho véc tơ ngẫu nhiên (X,Y) có hàm mật độ xác suất xác định như sau

$$f_{XY}(x,y)=\begin{cases} k & \text{neu } |x|+|y|\le 1 \\ 0 & \text{neu nguoc lai} \end{cases}$$



Miền D: $|x|+|y|\le 1$ đối xứng qua hai trục toạ độ Ox, Oy. Phần của D nằm trong góc phần tư thứ nhất là tam giác vuông cân $0\le x, 0\le y;\ x+y\le 1$

Vậy D là hình vuông có độ dài cạnh bằng $\sqrt{2}$, do đó

$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = k\,\mathrm{dt}\,\mathrm{D} = 2k \Rightarrow k=\frac{1}{2}$$

$$f_X(x) = \begin{cases} \dfrac{1}{2} \displaystyle\int_{-(1-|x|)}^{1-|x|} dy & \text{neu } |x| \le 1 \\ 0 & \text{neu nguoc lai} \end{cases} = \begin{cases} 1-|x| & \text{khi } |x| \le 1 \\ 0 & \text{khi } |x| > 1 \end{cases}$$

Do tính chất đối xứng của X và Y nên ta cũng có

$$f_Y(y) = \begin{cases} 1-|y| & \text{khi } |y| \le 1 \\ 0 & \text{khi } |y| > 1 \end{cases}$$

3.4 TÍNH ĐỘC LẬP CỦA CÁC BIẾN NGẪU NHIÊN

Hai biến ngẫu nhiên X, Y là độc lập nếu X nhận các giá trị nào đó không phụ thuộc Y và ngược lại. Nói cách khác với mọi số thực x, y hai biến cố $\{X \le x\}$, $\{Y \le y\}$ là độc lập.

Các dấu hiệu để nhận biết tính độc lập của hai biến ngẫu nhiên

Giả sử $F_{XY}(x,y)$ là hàm phân bố xác suất của véc tơ ngẫu nhiên (X,Y). Khi đó X, Y độc lập khi và chỉ khi

$$F_{XY}(x,y) = F_X(x)F_Y(y)$$

Giả sử $f_{XY}(x,y)$ là hàm mật độ xác suất của véc tơ ngẫu nhiên liên tục (X,Y). Khi đó X, Y độc lập khi và chỉ khi

$$f_{XY}(x,y) = f_X(x)f_Y(y)$$

Véc tơ ngẫu nhiên liên tục (X,Y) có hàm mật độ

$$f_{XY}(x,y)=\begin{cases}4xy & \text{neu } 0<x<1,\ 0<y<1\\ 0 & \text{neu nguoc lai}\end{cases}$$

Có hai hàm mật độ thành phần

$$f_X(x)=\begin{cases}2x & \text{neu } 0<x<1\\ 0 & \text{neu nguoc lai}\end{cases}$$

$$f_Y(y)=\begin{cases}2y & \text{neu } 0<y<1\\ 0 & \text{neu nguoc lai}\end{cases}$$

$$f_{XY}(x,y)=f_X(x)\,f_Y(y)$$

Vậy X và Y độc lập

Véc tơ ngẫu nhiên rời rạc hai chiều (X,Y).

X nhận các giá trị $x_1, \ldots ,x_n$, Y nhận các giá trị $y_1, \ldots, y_m$ là độc lập khi và chỉ khi

$$p_{XY}(x_i, y_j) = p_X(x_i)p_Y(y_j) \quad \forall i = 1,\ldots,n\ ;\ \ j = 1,\ldots,m$$

Một dấu hiệu để nhận biết hai biến ngẫu nhiên rời rạc độc lập là bảng phân bố xác suất đồng thời có tính chất:

- Hai hàng bất kỳ tỉ lệ với nhau.
- Hai cột bất kỳ tỉ lệ với nhau.

X \ Y	1	2	3
1	2/36	3/36	1/36
2	4/36	6/36	2/36
3	6/36	9/36	3/36

X \ Y	0	1	2	3
0	1/8	1/8	0	0
1	0	2/8	2/8	0
2	0	0	1/8	1/8

3.5 CÁC THAM SỐ ĐẶC TRƯNG CỦA VÉC TƠ NGẪU NHIÊN

3.5.1 Kỳ vọng và phương sai của các biến ngẫu nhiên thành phần

a. Trường hợp X, Y rời rạc

$$\mathrm{E}X = \sum_{i=1}^{n} x_i p_X(x_i) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i p_{XY}(x_i, y_j)$$

$$\mathrm{E}Y = \sum_{j=1}^{m} y_j p_Y(y_j) = \sum_{j=1}^{m}\sum_{i=1}^{n} y_j p_{XY}(x_i, y_j)$$

$$\mathrm{D}X = \mathrm{E}X^2 - (\mathrm{E}X)^2;\ \mathrm{E}X^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^2 p_{XY}(x_i, y_j)$$

$$\mathrm{D}Y = \mathrm{E}Y^2 - (\mathrm{E}Y)^2;\ \mathrm{E}Y^2 = \sum_{j=1}^{m}\sum_{i=1}^{n} y_j^2 p_{XY}(x_i, y_j)$$

b. Trường hợp X, Y liên tục

$$\mathrm{E}X = \int_{-\infty}^{\infty} x f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)dxdy$$

$$\mathrm{E}Y = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{X,Y}(x,y)dxdy$$

$$\mathrm{D}X = \mathrm{E}X^2 - (\mathrm{E}X)^2;\ \mathrm{E}X^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f_{X,Y}(x,y)dxdy$$

$$\mathrm{D}Y = \mathrm{E}Y^2 - (\mathrm{E}Y)^2;\ \mathrm{E}Y^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2 f_{X,Y}(x,y)dxdy$$

3.5.2 Hiệp phương sai

Hiệp phương sai (hay còn gọi là Covariance) của hai biến ngẫu nhiên X, Y, ký hiệu $\operatorname{cov}(X,Y)$, là kỳ vọng toán của tích các sai lệch của hai biến ngẫu nhiên đó với kỳ vọng toán của chúng

$$\operatorname{cov}(X,Y) = \mathrm{E}\left(X - \mathrm{E}X\right)\left(Y - \mathrm{E}Y\right)$$

Khai triển vế phải và áp dụng tính chất của kỳ vọng ta được

$$\operatorname{cov}(X,Y) = \mathrm{E}\,XY - (\mathrm{E}\,X)(\mathrm{E}\,Y)$$

Nếu X, Y rời rạc thì $$\mathrm{E}\,XY = \sum_{j=1}^{m}\sum_{i=1}^{n} x_i y_j p_{XY}(x_i, y_j)$$

Nếu X, Y liên tục thì $$\mathrm{E}\,XY = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\,dx\,dy$$

Tính chất của hiệp phương sai

1. $\operatorname{cov}(X,Y)=\operatorname{cov}(Y,X)$

2. $\operatorname{cov}(X,X)=\mathrm{D}X$

3. $\operatorname{cov}(aX+c,bY+d)=ab\operatorname{cov}(Y,X)$ với mọi hằng số a,b,c,d

4. Nếu X,Y độc lập thì $\operatorname{cov}(X,Y)=0$

   nhưng ngược lại chưa chắc đúng

3.5.3 Ma trận hiệp phương sai

Ma trận $M = \left[C_{ij} \right]_{n\times n}$

với $C_{ij} = \text{cov}(X_i, X_j);\ i, j = 1,\dots,n$

được gọi là ma trận hiệp phương sai (ma trận covariance) của véc tơ ngẫu nhiên $X=(X_1, X_2, \dots, X_n)$

Tính chất của ma trận hiệp phương sai

1. Ma trận hiệp phương sai là ma trận đối xứng
2. Với mọi $t_1, t_2, \dots, t_n \in 3^n$ luôn có $\sum_j \sum_i C_{ij} t_i t_j \geq 0$
3. Các định thức con chính của M không âm

3.5.4 Hệ số tương quan

Hệ số tương quan của hai biến ngẫu nhiên X, Y ký hiệu và định nghĩa bởi công thức

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}$$

Tính chất của hệ số tương quan

1. $-1 \le \rho_{X,Y} \le 1$ với mọi X, Y

2. Nếu X, Y độc lập thì $\rho_{X,Y} = 0$ điều ngược lại chưa chắc đúng

3) Với mọi hằng số a, b, c, d

$$\rho_{aX+c,bY+d} = \begin{cases} \rho_{X,Y} & \text{nếu } ab > 0 \\ -\rho_{X,Y} & \text{nếu } ab < 0 \end{cases}$$

4) $Y = aX + b$, $a \neq 0$ khi và chỉ khi

$$\rho_{X,Y} = \begin{cases} 1 & \text{nếu } a > 0 \\ -1 & \text{nếu } a < 0 \end{cases}$$

Xét véc tơ ngẫu nhiên (X,Y) có bảng phân bố xác suất

X \ Y	0	1	2	3
0	1/8	1/8	0	0
1	0	2/8	2/8	0
2	0	0	1/8	1/8

Bảng phân bố xác suất biên

X	0	1	2
P	2/8	4/8	2/8

Y	0	1	2	3
P	1/8	3/8	3/8	1/8

$$\mathrm{E}\,X = 0.\frac{2}{8} + 1.\frac{4}{8} + 2.\frac{2}{8} = 1;\ \mathrm{E}\,X^2 = 0^2.\frac{2}{8} + 1^2.\frac{4}{8} + 2^2.\frac{2}{8} = \frac{3}{2}$$

$$\Rightarrow \mathrm{D}\,X = \mathrm{E}\,X^2 - (\mathrm{E}\,X)^2 = \frac{1}{2}$$

$$\mathrm{E}Y = 0.\frac{1}{8} + 1.\frac{3}{8} + 2.\frac{3}{8} + 3.\frac{1}{8} = \frac{3}{2};\mathrm{E}Y^2 = 3 \Rightarrow \mathrm{D}Y = \frac{3}{4}$$

$$\mathrm{E}\,XY = 0.0.\frac{1}{8} + 0.1.\frac{1}{8} + 0.2.0 + 0.3.0 + 1.0.0 + 1.1.\frac{2}{8} + 1.2.\frac{2}{8} + 1.3.0 + 2.0.0$$

$$+ 2.1.0 + 2.2.\frac{1}{8} + 2.3.\frac{1}{8} = 2$$

$$\operatorname{cov}(X,Y) = \mathrm{E}\,XY - \mathrm{E}\,X\,\mathrm{E}\,Y = 2 - 1.\frac{3}{2} = \frac{1}{2}$$

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\mathrm{D}\,X}\sqrt{\mathrm{D}Y}} = \frac{1/2}{\sqrt{(1/2)(3/4)}} = \sqrt{\frac{2}{3}} \qquad M = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 3/4 \end{bmatrix}$$

(X,Y) là véc tơ ngẫu nhiên có bảng phân bố xác suất đồng thời

$X \backslash Y$	0	1
0	0,45	0,05
1	0,10	0,40

Bảng phân bố xác suất thành phần X và Y

X	0	1
P	0,5	0,5

Y	0	1
P	0,55	0,45

X và Y không độc lập

$$\mathrm{E}\,X = 0{,}5,\ \mathrm{E}\,X^2 = 0{,}5 \Rightarrow \mathrm{D}\,X = 0{,}25$$

$$\mathrm{E}Y = 0{,}45,\ \mathrm{E}Y^2 = 0{,}45 \Rightarrow \mathrm{D}Y = 0{,}2475$$

Hiệp phương sai

$$\mathrm{E}\,XY = 0{,}4 \Rightarrow \operatorname{cov}(X,Y) = 0{,}4 - 0{,}5 \times 0{,}45 = 0{,}175$$

Hệ số tương quan

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\mathrm{D}\,X}\sqrt{\mathrm{D}Y}} = \frac{0{,}175}{\sqrt{0{,}25 \times 0{,}2475}} = 0{,}704$$

Ma trận hiệp phương sai

$$M = \begin{bmatrix} 0{,}25 & 0{,}175 \\ 0{,}175 & 0{,}2475 \end{bmatrix}$$

Ta thấy giá trị $\rho_{X,Y} = 0{,}704$ khá xa 1, do đó Y không phụ thuộc tuyến tính đối với X

3.6 HÀM CỦA CÁC BIẾN NGẪU NHIÊN

3.6.1 Hàm của một biến ngẫu nhiên

Giả sử X là một biến ngẫu nhiên

$g(x)$ là hàm liên tục một biến số.

Khi đó $Y=g(X)$ cũng là một biến ngẫu nhiên và gọi là hàm của biến ngẫu nhiên X .

Điều kiện liên tục của hàm $g(x)$ có thể thay bởi điều kiện đo được. Tuy nhiên vần đề này vượt khỏi nội dung của chương trình.

1. Trường hợp X rời rạc

Giả sử X là biến ngẫu nhiên rời rạc có miền giá trị R_X và hàm khối lượng xác suất $p_X(x)$ thì Y cũng rời rạc

có miền giá trị $R_Y = \{g(x);\, x \in R_X\}$

hàm khối lượng xác suất

$$p_Y(y) = \sum_{x_k \in g^{-1}(y)} p_X(x_k)$$

2. Trường hợp X liên tục

Giả sử $g(x)$ là một song ánh, khi đó $g(x)$ hoặc đơn điệu tăng hoặc đơn điệu giảm

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

Nếu $g(x)$ là một hàm liên tục bất kỳ (không đơn điệu) ta chia miền xác định của $g(x)$ thành các miền mà $g(x)$ đơn điệu và áp dụng công thức trên trong các miền này.

Ngoài ra cũng có thể tính trực tiếp xác suất $P\{Y=g(X) \le y\}$, từ đó suy ra hàm phân bố xác suất và hàm mật độ xác suất của Y

Ví dụ: Giả sử $g_1(x) = x^3,\ g_2(x) = x^2,\ g_3(x) = e^{-x}$.

X là một biến ngẫu nhiên liên tục có hàm phân bố xác suất $F_X(x)$ và hàm mật độ xác suất $f_X(x)$

Hàm mật độ xác suất của $Y = g_1(X)$

$$f_Y(y) = f_X(x)\frac{dx}{dy} = \frac{1}{3}\frac{f_X(\sqrt[3]{y})}{\sqrt[3]{y^2}}$$

Hàm mật độ xác suất của $Z = g_2(X)$

Ta tính trực tiếp

$$F_Z(z) = P\{Z \le z\} = P\{X^2 \le z\} = P\{-\sqrt{z} \le X \le \sqrt{z}\} = F_X(\sqrt{z}) - F_X(-\sqrt{z})$$

$$f_Z(z) = \begin{cases} \dfrac{1}{2\sqrt{z}}\left(f_X(\sqrt{z}) + f_X(-\sqrt{z})\right) & \text{khi} \quad z > 0 \\ 0 & \text{khi} \quad z \le 0 \end{cases}$$

Hàm mật độ xác suất của $T = g_3(X)$

$g_3(x) = e^{-x}$ đơn điệu giảm. $t = e^{-x} \Leftrightarrow x = -\ln t;\ \dfrac{dx}{dt} = -\dfrac{1}{t}$

$$f_T(t) = f_X(x)\left|\frac{dx}{dt}\right| = f_X(-\ln t)\frac{1}{t}$$

Nếu hàm mật độ xác suất $f_X(x)$ của X xác định bởi

$$f_X(x) = \begin{cases} \dfrac{2}{x^2} & \text{neu } 1 \le x \le 2 \\ 0 & \text{neu nguoc lai} \end{cases}$$

Ta có hàm mật độ xác suất $f_Y(y)$

$$f_Y(y) = \frac{1}{3}\frac{f_X(\sqrt[3]{y})}{\sqrt[3]{y^2}} = \begin{cases} \dfrac{2}{3\sqrt[3]{y^4}} & \text{neu } 1 \le y \le 8 \\ 0 & \text{neu nguoc lai} \end{cases}$$

Cho biến ngẫu nhiên X có phân bố chuẩn $\mathbf{N}(\mu;\sigma^2)$

Xét hàm số $y = g(x) = ax + b$

Áp dụng công thức mật độ của hàm đơn điệu của biến ngẫu nhiên ta được

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) = \frac{1}{\sqrt{2\pi}|a|\sigma}\exp\left\{-\frac{1}{2\sigma^2}\left[\left(\frac{y-b}{a}\right)-\mu\right]^2\right\}$$

$$= \frac{1}{\sqrt{2\pi}|a|\sigma}\exp\left\{-\frac{1}{2a^2\sigma^2}\left[y-(a\mu+b)\right]^2\right\}$$

Chọn $a = 1/\sigma$, $b = -\mu/\sigma$ thì Y có phân bố chuẩn tắc

Vậy: nếu $X \sim \mathbf{N}(\mu;\sigma^2)$ thì $Y = \dfrac{X-\mu}{\sigma} \sim \mathbf{N}(0;1)$

3.6.2 Hàm của hai biến ngẫu nhiên

Giả sử X, Y là hai biến ngẫu nhiên và $g(x,y)$ là một hàm hai biến liên tục cho trước, khi đó $Z = g(X,Y)$ là một biến ngẫu nhiên được gọi là hàm của hai biến ngẫu nhiên X và Y

a. Trường hợp rời rạc

Nếu X, Y là hai biến ngẫu nhiên rời rạc có miền giá trị R_X, R_Y và hàm khối lượng xác suất $p_X(x)$, $p_Y(y)$ thì miền giá trị và hàm khối lượng xác suất của Z xác định như sau

$$R_Z = \left\{ g(x,y) \middle| x \in R_X, y \in R_Y \right\}$$

$$p_Z(z) = \sum_{(i,j);g(x_i,y_j)=z} P\left\{ X = x_i; Y = y_j \right\}$$

b. Trường hợp liên tục

Giả sử véc tơ ngẫu nhiên (X,Y) liên tục có miền giá trị R_{XY} và hàm mật độ xác suất đồng thời $f_{XY}(x,y)$

$Z=g(X,Y)$ cũng là biến ngẫu nhiên liên tục

Với mỗi $z\in \mathrm{R}$, ký hiệu $D_z=\{(x,y)\in R_{XY} \mid g(x,y)\le z\}$

Hàm phân bố xác suất của biến ngẫu nhiên $Z=g(X,Y)$ được xác định như sau

$$F_Z(z)=\iint_{D_z} f_{XY}(x,y)dxdy$$

3.6.3 Hàm phân bố của tổng hai biến ngẫu nhiên

Nếu X, Y rời rạc thì hàm khối lượng xác suất của Z xác định như sau

$$p_Z(z) = \sum_{x \in R_X} P\{X = x; Y = z - x\} = \sum_{y \in R_Y} P\{Y = y; X = z - y\}$$

Nếu X, Y liên tục có hàm mật độ xác suất đồng thời $f_{XY}(x,y)$ thì hàm mật độ xác suất của Z

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z-x)dx$$

Nếu X, Y độc lập thì hàm mật độ xác suất của Z có dạng

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx = f_X(z) * f_Y(z)$$

Ví dụ: Giả sử X, Y là hai biến ngẫu nhiên độc lập có phân bố Poisson với các tham số λ_1, λ_2. Xét biến ngẫu nhiên $Z=X+Y$

$$P\{Z=n\}=\sum_{k=0}^{n}P\{X=k\}P\{Y=n-k\}=\sum_{k=0}^{n}e^{-\lambda_1}\frac{\lambda_1^k}{k!}e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n$$

Ví dụ: Giả sử X, Y là hai biến ngẫu nhiên độc lập có phân bố đều trong khoản $(0;1)$. Xét biến ngẫu nhiên $Z=X+Y$

$$f_Z(z)=\int_{-\infty}^{\infty}f_X(x)f_Y(z-x)dx=\begin{cases} z & \text{neu } 0<z<1 \\ 2-z & \text{neu } 1<z<2 \\ 0 & \text{neu nguoc lai} \end{cases}$$

3.6.4 Hai hàm của hai biến ngẫu nhiên liên tục

Giả sử X, Y là hai biến ngẫu nhiên và $z=g(x,y)$, $w=h(x,y)$ là hai hàm hai biến liên tục cho trước. Khi đó ta có thể thiết lập một cặp biến ngẫu nhiên mới Z, W bởi công thức

$$\begin{cases} Z = g(X,Y) \\ W = h(X,Y) \end{cases}$$

Với mỗi $(z,w)\in \mathrm{R}^2$, ký hiệu

$$D_{zw} = \left\{(x,y)\in R_{XY} \middle| g(x,y)\le z; h(x,y)\le w\right\}$$

$$\left\{Z\le z; W\le w\right\} = \left\{g(X,Y)\le z; h(X,Y)\le w\right\} = \left\{(X,Y)\in D_{zw}\right\}$$

$$F_{ZW}(z,w) = \iint\limits_{D_{zw}} f_{XY}(x,y)\,dxdy$$

Công thức xác định hàm mật độ $f_{ZW}(z,w)$ từ hàm mật độ $f_{XY}(x,y)$

$$f_{ZW}(z,w)=\begin{cases} f_{XY}\left(x(z,w),y(z,w)\right)\left|\dfrac{D(x,y)}{D(z,w)}\right| & \text{neu } (z,w)\in\Delta \\ 0 & \text{neu nguoc lai}\end{cases}$$

trong đó $\Delta = T^{-1}(R_{XY})$

T là ánh xạ từ R^2 vào R^2 xác định bởi:

$$T(x,y)=\left(g(x,y),h(x,y)\right)$$

R_{XY} là miền giá trị của (X,Y)

3.6.5 Kỳ vọng của hàm các biến ngẫu nhiên

A. Kỳ vọng của hàm một biến ngẫu nhiên $Y = g(X)$

$$\mathrm{E}Y = \begin{cases} \sum_{i} g(x_i)\,p_X(x_i) & \text{neu } X \text{ roi rac} \\ \int_{-\infty}^{\infty} g(x)\,f_X(x)\,dx & \text{neu } X \text{ liên tuc} \end{cases}$$

B. Kỳ vọng của hàm n biến ngẫu nhiên $Y = g(X_1, \ldots, X_n)$

$$\mathrm{E}Y = \begin{cases} \sum_{i_1} \ldots \sum_{i_n} g(x_{i_1}, \ldots, x_{i_n})\,p_{X_1 \ldots X_n}(x_{i_1}, \ldots, x_{i_n}) & \text{neu } X \text{ roi rac} \\ \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n)\,f_{X_1 \ldots X_n}(x_1, \ldots, x_n)\,dx_1 \ldots dx_n & \text{neu } X \text{ liên tuc} \end{cases}$$

C. Tính chất tuyến tính của kỳ vọng

Với mọi biến ngẫu nhiên $X_1, \ldots, X_n$ và với mọi hằng số $a_1, \ldots, a_n$

$$\mathrm{E}\left(\sum_{k=1}^{n} a_k X_k\right) = \sum_{k=1}^{n} a_k \mathrm{E}(X_k)$$

d. Kỳ vọng của các biến ngẫu nhiên độc lập

Nếu X, Y là hai biến ngẫu nhiên độc lập thì

$$\mathrm{E}[g(X)h(Y)] = \mathrm{E}[g(X)]\mathrm{E}[h(Y)]$$

Nếu các biến ngẫu nhiên $X_1, \ldots, X_n$ độc lập thì

$$\mathrm{E}\left[\prod_{k=1}^{n} g_k(X_k)\right] = \prod_{k=1}^{n} \mathrm{E}[g_k(X_k)]$$

3.7. PHÂN BỐ CÓ ĐIỀU KIỆN VÀ KỲ VỌNG CÓ ĐIỀU KIỆN

3.7.1. Phân bố và kỳ vọng có điều kiện của biến ngẫu nhiên rời rạc

Cho biến ngẫu nhiên X rời rạc và B là một biến cố trong cùng phép thử với X có xác suất $P(B) > 0$

Biến ngẫu nhiên X được xét trong điều kiện biết rằng B đã xảy ra được gọi là biến ngẫu nhiên với điều kiện B, ký hiệu $X|B$

Hàm khối lượng xác suất của $X|B$

$$p_{X|B}(x_i|B) = \frac{P(\{X = x_i\}B)}{P(B)}$$

Kỳ vọng của X với điều kiện B

$$\mathrm{E}\left[X|B\right] = \sum_{i=1}^{n} x_i p_{X|B}(x_i|B)$$

Giả sử X, Y là hai biến ngẫu nhiên rời rạc có tập các giá trị $\{x_1, x_2, \dots, x_n\}$ và $\{y_1, y_2, \dots, y_n\}$

Với mỗi $y_j \in R_Y$, biến ngẫu nhiên X với điều kiện biến cố $\{Y=y_j\}$ có hàm khối lượng xác suất có điều kiện

$$p_{X|Y}(x_i \mid y_j) = \frac{p_{XY}(x_i, y_j)}{p_Y(y_j)}\ ;\ i = \overline{1,n},\ \ j = \overline{1,m}$$

Tương tự, hàm khối lượng xác suất có điều kiện của biến ngẫu nhiên Y với điều kiện $\{X=x_i\}$

$$p_{Y|X}(y_j \mid x_i) = \frac{p_{XY}(x_i, y_j)}{p_X(x_i)}\ ;\ i = \overline{1,n},\ \ j = \overline{1,m}$$

Tính chất của hàm khối lượng xác suất có điều kiện

1. $0 \le p_{X|Y}(x_i \mid y_j) \le 1;\ \forall i = \overline{1,n},\ \forall j = \overline{1,m}$
2. Với mỗi j: $\sum_{i=1}^{n} p_{X|Y}(x_i \mid y_j) = 1$
3. Nếu X, Y độc lập thì

$p_{X|Y}(x_i \mid y_j) = p_X(x_i)$ và $p_{Y|X}(y_j \mid x_i) = p_Y(y_j)$

Kỳ vọng có điều kiện

$$\mathrm{E}\left[X \middle| Y = y_j\right] = \sum_{i=1}^{n} x_i p_{X|Y}(x_i \mid y_j)$$

$$\mathrm{E}\left[Y \middle| X = x_i\right] = \sum_{j=1}^{m} y_j p_{Y|X}(y_j \mid x_i)$$

Thống kê dân cư của một thành phố nọ ở độ tuổi trưởng thành về thu nhập hàng tháng X và lứa tuổi Y, thu được kết quả trong bảng sau

X \ Y	30	45	70	$\sum$
1	0,01	0,02	0,05	0,08
2	0,03	0,06	0,10	0,19
3	0,18	0,21	0,15	0,54
4	0,07	0,08	0,04	0,19
$\sum$	0,29	0,37	0,34	

Tìm thu nhập trung bình theo lứa tuổi

Với $Y=30$ bảng phân bố xác suất điều kiện tương ứng

$X \backslash Y$	30	45	70	$\sum$
1	0,01	0,02	0,05	0,08
2	0,03	0,06	0,10	0,19
3	0,18	0,21	0,15	0,54
4	0,07	0,08	0,04	0,19
$\sum$	0,29	0,37	0,34	

$X \mid Y=30$	1	2	3	4
P	$\frac{0,01}{0,29}$	$\frac{0,03}{0,29}$	$\frac{0,18}{0,29}$	$\frac{0,07}{0,29}$

$$\mathrm{E}\left[X\middle|Y=30\right]=1\cdot\frac{1}{29}+2\cdot\frac{3}{29}+3\cdot\frac{18}{29}+4\cdot\frac{7}{29}=\frac{89}{29}=3,069$$

$$\mathrm{E}\left[X\middle|Y=45\right]=\frac{109}{37}=2,9459$$

$$\mathrm{E}\left[X\middle|Y=70\right]=\frac{86}{34}=2,5294$$

Vậy thu nhập trung bình ở độ tuổi 30 là 3.069.000đ/tháng,
độ tuổi 45 là 2.945.900đ/tháng
và độ tuổi 70 là 2.529.400 đ/tháng

3.7.2 Phân bố và kỳ vọng có điều kiện của biến ngẫu nhiên liên tục

Xét biến ngẫu nhiên X và biến cố B trong cùng một phép thử thỏa mãn $P(B)>0$.

Hàm phân bố xác suất, hàm mật độ xác suất của X với điều kiện B được định nghĩa và ký hiệu như sau

$$F_{X|B}(x|B)=P\left\{X\le x\middle|B\right\}=\frac{P\left(\left\{X\le x\right\}B\right)}{P(B)}$$

$$f_{X|B}(x|B)=\frac{dF_{X|B}(x|B)}{dx}$$

Giả sử $f_{XY}(x,y)$ là hàm mật độ xác suất đồng thời của hai biến ngẫu nhiên liên tục X, Y và $f_X(x)$ là hàm mật độ của biến ngẫu nhiên thành phần X.

Hàm phân bố xác suất của Y với điều kiện $\{X = x\}$ được định nghĩa và ký hiệu như sau

$$F_{Y|X}(y\,|\,x) = \int_{-\infty}^{y} \frac{f_{XY}(x,v)}{f_X(x)}\,dv, \text{ nếu } \quad f_X(x) > 0$$

Đạo hàm của hàm phân bố xác suất điều kiện được hàm mật độ xác suất có điều kiện

$$f_{Y|X}(y\,|\,x) = \frac{f_{XY}(x,y)}{f_X(x)} \text{ với mọi } y \in 3 \text{ và } x \text{ thoả mãn } f_X(x) > 0$$

Tương tự ta có hàm phân bố xác suất và hàm mật độ xác suất có điều kiện của X với $\{Y = y\}$

$$F_{X|Y}(x \mid y) = \int_{-\infty}^{x} \frac{f_{XY}(u, y)}{f_Y(y)} du \text{, nếu } \quad f_Y(y) > 0$$

$$f_{X|Y}(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)} \text{ với mọi } x \in 3 \text{ và } y \text{ thoả mãn } f_Y(y) > 0$$

Tính chất của mật độ có điều kiện

1. $f_{X|Y}(x \mid y) \geq 0$
2. Với mỗi y thỏa mãn $f_Y(y) > 0$ thì có $\displaystyle\int_{-\infty}^{\infty} f_{X|Y}(x \mid y) dx = 1$
3. Nếu X, Y độc lập thì

$$f_{X|Y}(x \mid y) = f_X(x) \text{ và } f_{Y|X}(y \mid x) = f_Y(y)$$

Kỳ vọng với điều kiện

Kỳ vọng của Y với điều kiện $\{X = x\}$ được ký hiệu và định nghĩa theo công thức sau

$$\mathrm{E}\left[Y \middle| X = x\right] = \int_{-\infty}^{\infty} y f_{Y|X}(y \mid x) dy$$

$\mathrm{E}[Y|X = x]$ xác định là một hàm của biến x, được gọi là hàm hồi qui của Y đối với X

$\mathrm{E}[X|Y = y]$ xác định là một hàm của biến y, được gọi là hàm hồi qui của X đối với Y

3.7.3 Biến ngẫu nhiên kỳ vọng có điều kiện

Hàm hồi qui của Y đối với X là một hàm phụ thuộc biến x:

$$H(x) = \mathrm{E}[Y|X = x]$$

Xét hàm $H(X)$ của biến ngẫu nhiên X được xác định bởi $H(x)$ gọi là biến ngẫu nhiên kỳ vọng có điều kiện

Ta ký hiệu là $\mathrm{E}[Y|X]$

Tương tự $\mathrm{E}[X|Y] = G(Y)$, trong đó $G(y) = \mathrm{E}[X|Y=y]$

X rời rạc

$$\mathrm{E}(Y) = \sum_i \mathrm{E}\left[Y|X = x_i\right] p_X(x_i) = \mathrm{E}\left(\mathrm{E}\left[Y|X\right]\right)$$

X liên tục

$$\mathrm{E}(Y) = \int_{-\infty}^{\infty} \mathrm{E}\left[Y|X = x\right] f_X(x)dx = \mathrm{E}\left(\mathrm{E}\left[Y|X = x\right]\right)$$

3.8 PHÂN BỐ CHUẨN NHIỀU CHIỀU

3.8.1 Khái niệm phân bố chuẩn *n* chiều

Hàm mật độ đồng thời của phân bố chuẩn n chiều

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det M}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\mathbf{a})M^{-1}(\mathbf{x}-\mathbf{a})^t \right\}$$

trong đó $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$; $\mathbf{a} = (\mu_1, \mu_2, ..., \mu_n) \in \mathbb{R}^n$

$M = \left[C_{ij} \right]_{n\times n}$ là ma trận đối xứng xác định dương

ma trận cột $(\mathbf{x}-\mathbf{a})^t$ là chuyển vị của ma trận hàng $(\mathbf{x}-\mathbf{a})$

Ký hiệu $\left(X_1, X_2, ..., X_n \right) \sim \mathbf{N}(\mathbf{a}, M)$

Tính chất của véc tơ ngẫu nhiên nhiều chiều

1. Với mọi $i = 1,\dots,n$: $\mu_i = \mathrm{E}X_i$

2. M là ma trận hiệp phương sai $C_{ij} = \operatorname{cov}(X_i, X_j)$

3. Các biến ngẫu nhiên $X_1, X_2, \dots, X_n$ độc lập khi và chỉ khi chúng không tương quan, nghĩa là M là ma trận chéo

4. Biến ngẫu nhiên thành phần X_i có phân bố chuẩn $\mathbf{N}(\mu_i; C_{ii})$

5. $(Y_1, Y_2, \dots, Y_n) = H(X_1, X_2, \dots, X_n)^t$ cũng có phân bố chuẩn, trong đó H là ma trận vuông cấp n bất kỳ có $\det H \neq 0$.

3.8.2 Phân bố chuẩn hai chiều

Hàm mật độ xác suất đồng thời của véc tơ ngẫu nhiên (X,Y) phân bố chuẩn hai chiều

$$f_{XY}(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}Q(x,y)\right)$$

trong đó

$$Q(x,y)=\frac{1}{(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}$$

Giả sử véc tơ ngẫu nhiên (X,Y) có phân bố chuẩn hai chiều, khi đó

1. X có phân bố chuẩn $\mathbf{N}(\mu_1;\sigma_1^2)$, Y có phân bố chuẩn $\mathbf{N}(\mu_2;\sigma_2^2)$
2. Hệ số tương quan $\rho(X,Y)=\rho$
3. X , Y không tương quan khi và chỉ khi X , Y độc lập
4. Hàm mật độ xác suất của X với điều kiện $Y=y$

$$f_{X|Y}(x\,|\,y)=\frac{f_{XY}(x,y)}{f_Y(y)}=\frac{1}{\sigma_1\sqrt{1-\rho^2}}\varphi\left(\frac{x-\mu_1-(\sigma_1/\sigma_2)\rho(y-\mu_2)}{\sigma_1\sqrt{1-\rho^2}}\right)$$

5. Ma trận tương quan của (X,Y) là

$$M=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}\Rightarrow M^{-1}=\frac{1}{1-\rho^2}\begin{bmatrix}1/\sigma_1^2 & -\rho/(\sigma_1\sigma_2)\\ -\rho/(\sigma_1\sigma_2) & 1/\sigma_2^2\end{bmatrix}$$

6. Kỳ vọng có điều kiện $\mathrm{E}\left[X\middle|Y=y\right]=\mu_1+\rho\dfrac{\sigma_1}{\sigma_2}(y-\mu_2)$