

Chương 5:

Cặp biến ngẫu nhiên

Nguyễn Linh Trung
Trần Thị Thúy Quỳnh

Đại học Công nghệ, ĐHQGHN

Nội dung

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

Cặp biến ngẫu nhiên

Rất nhiều thực nghiệm ngẫu nhiên gồm hơn một biến ngẫu nhiên.

Ví dụ:

- 1 Tên của học sinh được chọn ngẫu nhiên từ bình (các thẻ tên được chứa trong bình). ζ là kết quả của thực nghiệm và được định nghĩa thông qua hai hàm:

- $H(\zeta)$ là chiều cao của học sinh ζ
- $W(\zeta)$ là cân nặng của học sinh ζ

$H(\zeta), W(\zeta)$ là cặp số ứng với mỗi ζ thuộc không gian mẫu S .

- 2 ζ là kết quả của thực nghiệm xét ngẫu nhiên một trang Web. Mỗi trang Web cho phép người dùng chọn chức năng xem một đoạn quảng cáo ngắn hoặc không trước khi vào trang Web yêu cầu. Gọi

- $N_1(\zeta)$ là số lần truy cập chọn chức năng xem quảng cáo.
- $N_2(\zeta)$ số lần truy cập chọn chức năng không xem quảng cáo.

$N_1(\zeta), N_2(\zeta)$ là cặp số gắn với mỗi ζ trong không gian mẫu S .

Cặp biến ngẫu nhiên

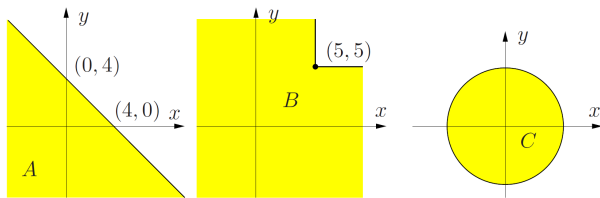
Definition (Cặp biến ngẫu nhiên)

Gọi ζ là kết quả trong không gian mẫu S . Cặp biến ngẫu nhiên $\mathbf{X}(\zeta)$ là một hàm ánh xạ ζ thành cặp số thực:

$$\mathbf{X}(\zeta) = (X(\zeta), Y(\zeta))$$

Cặp biến ngẫu nhiên

Các biến cố mong muốn gồm một cặp biến ngẫu nhiên thỏa mãn các điều kiện mong muốn có thể được biểu diễn bởi một vùng trong mặt phẳng.



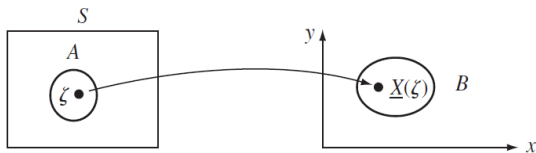
$$A = \{X + Y \leq 4\}, \quad B = \{\min(X, Y) \leq 5\}, \quad C = \{X^2 + Y^2 \leq 25\}$$

Cặp biến ngẫu nhiên

Biến cố và xác suất

Xác suất để biến cố $\mathbf{X} = (X(\zeta), Y(\zeta))$ nằm trong vùng B tương đương với xác suất để ζ nằm trong vùng A của không gian mẫu S . Trong đó,

$$A = \mathbf{X}^{-1}(B) = \{\zeta : (X(\zeta), Y(\zeta)) \in B\}$$



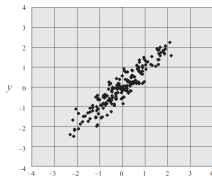
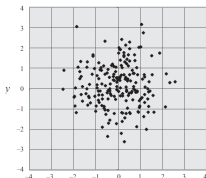
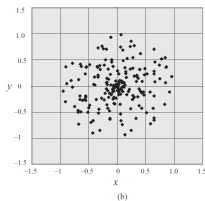
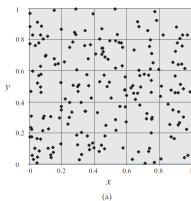
Khi đó,

$$P[\mathbf{X} \in B] = P[A] = P[\{\zeta : (X(\zeta), Y(\zeta)) \in B\}]$$

Cặp biến ngẫu nhiên

Biến cố và xác suất

- Sự khác biệt của cặp biến ngẫu nhiên so với một biến ngẫu nhiên là **biểu hiện đồng thời (kết hợp)** giữa X và Y .
- Biểu hiện đồng thời của cặp biến ngẫu nhiên (X, Y) có thể được quan sát thông qua 200 mẫu của 4 cặp biến ngẫu nhiên:



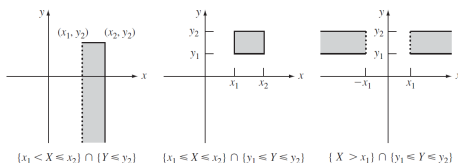
Biểu hiện đồng thời của cặp biến ngẫu nhiên (X, Y) có thể được biểu diễn thông qua các hàm:

- PMF đồng thời
- CDF đồng thời
- PDF đồng thời
- Các moment, kỳ vọng đồng thời

Cặp biến ngẫu nhiên

Biến cố và xác suất

Xét các biến cố tương ứng với các hình chữ nhật trên mặt phẳng:



Xét biến cố có dạng $B = \{X \in A_1\} \cap \{Y \in A_2\}$, với A_k là biến cố một chiều (một tập con của trục thực). Biến cố B xuất hiện khi cả $\{X \in A_1\}$ và $\{Y \in A_2\}$ xuất hiện đồng thời. Xác suất của biến cố được định nghĩa bởi:

$$P[B] = P[\{X \in A_1\} \cap \{Y \in A_2\}] \triangleq P[\{X \in A_1\}, \{Y \in A_2\}]$$

Cặp biến ngẫu nhiên rời rạc

Definition

Cặp biến ngẫu nhiên rời rạc

Vector biến ngẫu nhiên $\mathbf{X} = (X, Y)$ nhận các giá trị trong không gian mẫu $S_{X,Y} = \{(x_j, y_k), j = 1, 2, \dots, k = 1, 2, \dots\}$.

Definition

Xác suất đồng thời

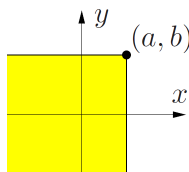
$$P[B] = P[\{X \in A_1\} \cap \{Y \in A_2\}] \triangleq P[\{X \in A_1\}, \{Y \in A_2\}]$$

Hàm phân bố tích lũy đồng thời

Definition

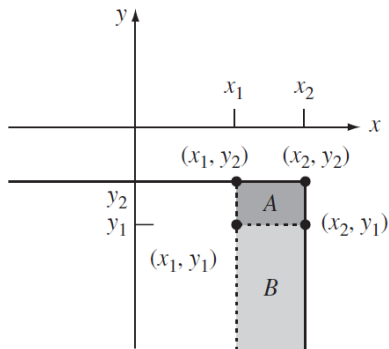
Hàm phân bố tích lũy đồng thời

$$F_{X,Y}(a,b) = P[X \leq a, Y \leq b]$$



Tính chất:

- CDF đồng thời là một hàm không giảm theo x và y
 $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$ nếu $x_1 \leq x_2$ và $y_1 \leq y_2$
- $F_{X,Y}(x_1, -\infty) = 0$, $F_{X,Y}(-\infty, y_1) = 0$, $F_{X,Y}(\infty, \infty) = 1$
- Hàm CDF lẻ có được khi bỏ đi ràng buộc của một biến ngẫu nhiên
 $F_X(x_1) = F_{X,Y}(x_1, \infty)$ và $F_Y(y_1) = F_{X,Y}(\infty, y_1)$



■ $P[x_1 < X \leq x_2, y_1 < Y \leq y_2] =$
 $F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)$

PMF của cặp biến ngẫu nhiên rời rạc

Definition

PMF đồng thời

$$p_{X,Y}(x,y) = P[X = x, Y = y]; \quad (x,y) \in S$$

PMF lề (marginal)

$$p_X(x) = \sum_{y \in S} p_{X,Y}(x,y); \quad p_Y(y) = \sum_{x \in S} p_{X,Y}(x,y)$$

PDF của cặp biến ngẫu nhiên liên tục

Definition

PDF đồng thời

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

PDF lề (marginal)

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy; \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

Tính độc lập của hai biến ngẫu nhiên

Definition

X và Y được gọi là độc lập khi

$$P[X \in A_1, Y \in A_2] = P[X \in A_1]P[Y \in A_2]$$

Tương đương,

- $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ với mọi x và y
- $F_{X,Y}(x, y) = F_X(x)F_Y(y)$ với mọi x và y
- $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ với mọi x và y

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên**
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

Kỳ vọng đồng thời hàm của hai biến ngẫu nhiên

Definition

Kỳ vọng của $Z = g(X, Y)$ được định nghĩa bởi:

- $E[Z] = \sum_x \sum_y g_{X,Y}(x, y)p_{X,Y}(x, y)$ với biến ngẫu nhiên rời rạc
- $E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_{X,Y}(x, y)f_{X,Y}(x, y)dxdy$ với biến ngẫu nhiên liên tục

Tính chất:

- $E[X + Y] = E[X] + E[Y]$
- $E[XY] = E[X]E[Y]$ nếu X và Y độc lập

Moment đồng thời

Definition

Moment đồng thời của X và Y được định nghĩa bởi:

- $E[X^j Y^k] = \sum_x \sum_y x^j y^k p_{X,Y}(x, y)$ với biến ngẫu nhiên rời rạc
- $E[X^j Y^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^j y^k f_{X,Y}(x, y) dx dy$ với biến ngẫu nhiên liên tục

Hiệp phương sai - Covariance

Definition

Hiệp phương sai của X và Y được định nghĩa bởi:

$$COV(X, Y) = E[(X - E[X])(Y - E[Y])]$$

Tính chất:

- $COV(X, Y) = E[XY] - E[X]E[Y]$
- $COV(X, Y) = 0$ nếu X và Y độc lập

Hệ số tương quan

Definition

Hệ số tương quan của X và Y được định nghĩa bởi:

$$\rho_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y}$$

với $\sigma_X = \sqrt{VAR(X)}$, $\sigma_Y = \sqrt{VAR(Y)}$ lần lượt là độ lệch chuẩn của biến ngẫu nhiên X và Y .

Tính chất:

- $-1 \leq \rho_{X,Y} \leq 1$
- X và Y gọi là phụ thuộc tuyến tính $Y = aX + b$ nếu: $\rho_{X,Y} = 1$ ứng với $a > 0$ và $\rho_{X,Y} = -1$ ứng với $a < 0$
- X và Y được gọi là không tương quan nếu $\rho_{X,Y} = 0$
- Nếu X và Y là **độc lập** thì chúng **không tương quan**

Hàm đặc trưng

Definition

Hàm đặc trưng đồng thời của X và Y được định nghĩa bởi:

$$\Phi_{X,Y}(\lambda, \omega) = E[e^{j(\lambda X + \omega Y)}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j(\lambda x + \omega y)} f_{X,Y}(x, y) dx dy$$

Tính chất:

- Nếu X và Y độc lập thì

$$\Phi_{X,Y}(\lambda, \omega) = E[e^{j\lambda X}]E[e^{j\omega Y}] = \Phi_X(\lambda)\Phi_Y(\omega)$$

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

Xác suất có điều kiện

■ Biến ngẫu nhiên rời rạc

$$p_Y(y|x) = P[Y = y|X = x] = \frac{P[Y = y, X = x]}{P[X = x]} = \frac{p_{X,Y}(x, y)}{p_X(x)}$$

■ Biến ngẫu nhiên liên tục

$$F_Y(y|x) = \frac{P[Y \leq y, X = x]}{P[X = x]}$$

$$f_Y(y|x) = \frac{d}{dy} F_Y(y|x)$$

Nếu X và Y độc lập thì $P[Y \leq y, X = x] = P[Y \leq y]P[X = x]$,
dẫn đến

$$F_Y(y|x) = F_Y(y) \text{ và } f_Y(y|x) = f_Y(y)$$

Kì vọng có điều kiện

- Biến ngẫu nhiên rời rạc

$$E[Y|x] = E[Y|X = x] = \sum_y y p_Y(y|x)$$

- Biến ngẫu nhiên liên tục

$$E[Y|x] = E[Y|X = x] = \int_{-\infty}^{\infty} y f_Y(y|x) dy$$

Tính chất:

$$E[E[Y|X]] = E[Y]$$

Nội dung

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

Tổng của hai biến ngẫu nhiên

Cho $Z = X + Y$ thì

$$P[Z \leq z] = P[X + Y \leq z]$$

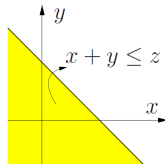
■ **CDF** của Z :

$$F_Z(z) = P[Z \leq z] = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_{X,Y}(x,y) dx dy$$

■ **PDF** của Z :

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

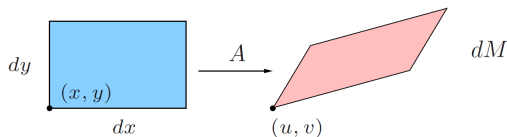
$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$ khi X và Y độc lập.



Phép biến đổi tuyến tính

Cho ma trận A là phép biến đổi tuyến tính

$$\begin{bmatrix} U \\ V \end{bmatrix} = A \begin{bmatrix} X \\ Y \end{bmatrix} \iff \begin{bmatrix} X \\ Y \end{bmatrix} = A^{-1} \begin{bmatrix} U \\ V \end{bmatrix}$$



$$P(X \in dx, Y \in dy) = f_{XY}(x, y) dx dy, \quad P(U \in du, V \in dv) = f_{UV}(u, v) dM$$

$$dM = |\det A| dx dy,$$

$$f_{UV}(u, v) = \frac{1}{|\det A|} f_{XY}(x, y)$$

Nội dung

- 1 Khái niệm và xác suất của cặp biến ngẫu nhiên
- 2 Tính độc lập của hai biến ngẫu nhiên
- 3 Kỳ vọng, Moment, Hiệp phương sai, Hệ số tương quan, và Hàm đặc trưng của hai biến ngẫu nhiên
- 4 Xác suất và kỳ vọng có điều kiện
- 5 Hàm của hai biến ngẫu nhiên
- 6 Cặp biến ngẫu nhiên phân bố Gauss đồng thời

Định nghĩa

Definition

Cặp biến ngẫu nhiên (X, Y) phân bố Gauss đồng thời được định nghĩa bởi:

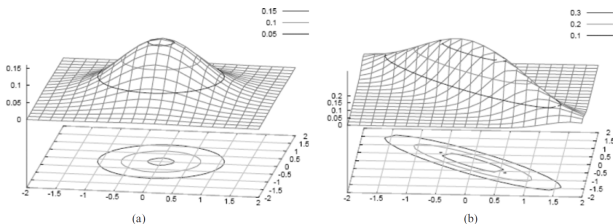
$f_{X,Y}(x, y)$

$$= \frac{\exp \left\{ \frac{-1}{2(1-\rho_{X,Y}^2)} \left[\left(\frac{x-m_1}{\sigma_1} \right)^2 - 2\rho_{X,Y} \left(\frac{x-m_1}{\sigma_1} \right) \left(\frac{y-m_2}{\sigma_2} \right) + \left(\frac{y-m_2}{\sigma_2} \right)^2 \right] \right\}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{X,Y}^2}}$$

với $-\infty \leq x \leq +\infty$ và $-\infty \leq y \leq +\infty$

Hàm mật độ xác suất PDF

PDF phân bố Gauss đồng thời của X và Y được biểu diễn như sau:



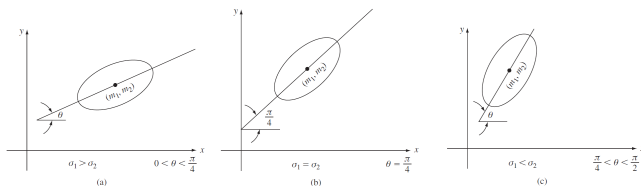
Jointly Gaussian pdf (a) $\rho = 0$ (b) $\rho = -0.9$.

- PDF có tâm tại điểm (m_1, m_2) .
- Dạng chuông của PDF phụ thuộc vào σ_1 , σ_2 , và $\rho_{X,Y}$.
- PDF là hằng số ứng với các giá trị x, y để

$$\left[\left(\frac{x-m_1}{\sigma_1} \right)^2 - 2\rho_{X,Y} \left(\frac{x-m_1}{\sigma_1} \right) \left(\frac{y-m_2}{\sigma_2} \right) + \left(\frac{y-m_2}{\sigma_2} \right)^2 \right] = \text{const}$$

Hàm PDF

Hướng của các đường cong ellip phụ thuộc vào các giá trị σ_1 , σ_2 , và $\rho_{X,Y}$ được biểu diễn như sau:



Orientation of contours of equal value of joint Gaussian pdf for $\rho_{X,Y} > 0$.

- Khi $\rho_{X,Y} = 0$, (X và Y độc lập), đường cong PDF là một hình ellip có bán kính 2 trục bằng nhau (hình tròn), trục cơ sở song song với trục x và y .
- Khi $\rho_{X,Y} \neq 0$, trục lớn của ellip tạo với trục x một góc:

$$\theta = \frac{1}{2} \arctan \left(\frac{2\rho_{X,Y}\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \right)$$

$$\theta = 45^\circ \text{ khi } \sigma_1^2 = \sigma_2^2.$$

Hàm margin-PDF

$$f_X(x) = \int_y f_{X,Y}(x,y) = \frac{e^{-(x-m_1)^2/2\sigma_1^2}}{\sqrt{2\pi}\sigma_1}$$

Như vậy, X là biến ngẫu nhiên phân bố Gauss trung bình m_1 , phương sai σ_1^2 .

$$f_Y(y) = \int_x f_{X,Y}(x,y) = \frac{e^{-(y-m_2)^2/2\sigma_2^2}}{\sqrt{2\pi}\sigma_2}$$

Như vậy, Y là biến ngẫu nhiên phân bố Gauss trung bình m_2 , phương sai σ_2^2 .

Hàm PDF có điều kiện

Hàm PDF điều kiện của X với $Y = y$:

$$f_X(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$
$$= \frac{\exp \left\{ \frac{-1}{2(1-\rho_{X,Y}^2)\sigma_1^2} \left[x - \rho_{X,Y} \frac{\sigma_1}{\sigma_2} (y - m_2) - m_1 \right]^2 \right\}}{\sqrt{2\pi\sigma_1^2(1-\rho_{X,Y}^2)}}$$

Như vậy, PDF có điều kiện cũng là Gauss với giá trị trung bình điều kiện $m_1 + \rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)$ và phương sai điều kiện $\sigma_1^2(1 - \rho_{X,Y}^2)$.

Tương tự với $f_Y(y|x)$.

Hệ số tương quan

$$COV(X, Y) = E[(X - m_1)(Y - m_2)] = E[E[(X - m_1)(Y - m_2)|Y]]$$

$$\begin{aligned} E[(X - m_1)(Y - m_2)|Y = y] &= (y - m_2)E[(X - m_1)|Y = y] \\ &= (y - m_2)(E[X|Y = y] - m_1) \end{aligned}$$

Do $E[X|Y = y] = m_1 + \rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)$ nên:

$$\begin{aligned} COV(X, Y) &= E[(X - m_1)(Y - m_2)] = (y - m_2)[\rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)] \\ &= \rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)^2 \end{aligned}$$

Như vậy,

$$\begin{aligned} COV(X, Y) &= E[E[(X - m_1)(Y - m_2)|Y]] = E[\rho_{X,Y}(\sigma_1/\sigma_2)(Y - m_2)^2] \\ \Rightarrow \rho_{X,Y} &= \frac{COV(X, Y)}{\sigma_1\sigma_2} \end{aligned}$$