TRƯỜNG ĐẠI HỌC KINH TẾ - LUẬT
KHOA TOÁN KINH TẾ

Chương 2. Đại lượng ngẫu nhiên và quy luật phân phối xác suất

Thành phố Hồ Chí Minh, 2020

Nội dung

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Định nghĩa (Đại lượng ngẫu nhiên)

Một **đại lượng ngẫu nhiên** là mô tả bằng số các kết quả của một phép thử ngẫu nhiên.

- Đại lượng ngẫu nhiên thường được kí hiệu bằng chữ cái in hoa như X, Y, Z.
- Các giá trị của đại lượng ngẫu nhiên được kí hiệu bằng chữ cái in thường như x, y, z.
- Khi đó, xác suất của đại lượng ngẫu nhiên X bằng giá trị x được kí hiệu là:

$$P(X = x)$$

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Phân loại đại lượng ngẫu nhiên

Đại lượng ngẫu nhiên được chia thành hai loại: Đại lượng ngẫu nhiên rời rạc và đại lượng ngẫu nhiên liên tục, phụ thuộc vào giá trị mà nó có thể nhận.

Định nghĩa (Đại lượng ngẫu nhiên rời rạc)

Đại lượng ngẫu nhiên được gọi là **rời rạc** nếu tập hợp các giá trị mà nó có thể nhận được là một tập hữu hạn hoặc vô hạn đếm được, chẳng hạn 0, 1, 2,...

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Ví dụ minh họa

Xác định các giá trị có thể nhận được của các đại lượng ngẫu nhiên rời rạc được cho trong bảng sau:

Phép thử	**Đại lượng ngẫu nhiên (X)**
Bắn 3 viên đạn vào mục tiêu	Số lần bắn trúng mục tiêu
Kiểm tra chất lượng 50 chiếc radio	Số lượng radio kém chất lượng
Mở cửa một nhà hàng trong 1 ngày	Số lượng khách hàng
Bán một chiếc ô tô	Giới tính khách hàng

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Giá trị có thể nhận được của từng biến ngẫu nhiên trong bảng trên là:

- $\{0, 1, 2, 3\}$
- $\{0, 1, 2, \ldots, 49, 50\}$
- $\{0, 1, 2, \ldots\}$
- 0 nếu là nam, 1 nếu là nữ

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Định nghĩa (Đại lượng ngẫu nhiên liên tục)

Một đại lượng ngẫu nhiên có thể nhận bất kỳ giá trị nào trong một khoảng hoặc tập hợp của nhiều khoảng được gọi là đại lượng ngẫu nhiên **liên tục**.

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Ví dụ minh họa

Xác định các giá trị có thể nhận được của các đại lượng ngẫu nhiên liên tục được cho trong bảng sau:

Phép thử	**Đại lượng ngẫu nhiên (X)**
Quan sát một ngân hàng	Thời gian giao dịch giữa 2 khách hàng (phút)
Rót nước vào một cái can (10 lít)	Số lít nước đã rót vào
Xây dựng một thư viện	Phần trăm hoàn thành
Đi xe từ SG đến HN (dài 1700km)	Khoảng cách đã đi được

1. Mô tả khái niệm và phân loại đại lượng ngẫu nhiên

Giá trị có thể nhận được của từng biến ngẫu nhiên trong Bảng **??** là:

- $x \geq 0$
- $0 \leq x \leq 10$
- $0 \leq x \leq 100$
- $0 \leq x \leq 1700$

2. Phân phối xác suất của đại lượng ngẫu nhiên

- Đối với đại lượng ngẫu nhiên rời rạc: Bảng phân phối xác suất.
- Đối với đại lượng ngẫu nhiên liên tục: Hàm mật độ xác suất.
- Trường hợp dùng cho cả đại lượng ngẫu nhiên rời rạc và liên tục: hàm phân phối (tích luỹ) xác suất.

2. Phân phối xác suất của đại lượng ngẫu nhiên

Trường hợp rời rạc

Cho $X = \{x_1, x_2, ..., x_n\}$ là một đại lượng ngẫu nhiên rời rạc với xác suất tương ứng là $p_i = P(X = x_i),\ i = \overline{1, n}$.

Khi đó, **bảng phân phối xác suất** của X như sau:

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

2. Phân phối xác suất của đại lượng ngẫu nhiên

Đối với **bảng phân phối xác suất** của đại lượng ngẫu nhiên rời rạc ta luôn có:

- $0 \le p_i \le 1$
- $\sum_{i=1}^{n} p_i = 1$ (trường hợp đại lượng ngẫu nhiên rời rạc hữu hạn)
- $P(a \le X < b) = \sum_{a \le X < b} p_i$
- $P(a < X < b) = \sum_{a < X < b} p_i$

2. Phân phối xác suất của đại lượng ngẫu nhiên

Ví dụ minh họa

Gọi X là số môn đậu của một sinh viên trong học kỳ phải thi 5 môn. Khi đó $X = \{0, 1, 2, 3, 4, 5\}$.
Giả sử X có bảng phân phối xác suất như sau:

X	0	1	2	3	4	5
P	0.05	0.15	0.3	0.35	0.15	0

Từ bảng phân phối trên, ta có thể đưa ra một vài nhận xét sau:

- $P(X = 5) = 0$: Sinh viên đó không thể đậu 5 môn
- $P(X = 3) = 0.35$: khả năng sinh viên đó đậu 3 môn là nhiều nhất

2. Phân phối xác suất của đại lượng ngẫu nhiên

Định nghĩa (Hàm phân phối xác suất)

Hàm số

$$F(x) = P(X \leq x), \quad \forall x \in \mathbb{R},$$

được gọi là **hàm phân phối xác suất** của đại lượng ngẫu nhiên X.

Nhận xét

Hàm phân phối xác suất định nghĩa ở trên tồn tại đối với cả đại lượng ngẫu nhiên rời rạc và đại lượng ngẫu nhiên liên tục.

2. Phân phối xác suất của đại lượng ngẫu nhiên

Nếu X là đại lượng ngẫu nhiên rời rạc với xác suất tại các giá trị x_i là $p_i = P(X = x_i)$ thì:

$$F(x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{i \in I} p_i,$$

trong đó, $I = \{i | x_i \leq x\}$.

2. Phân phối xác suất của đại lượng ngẫu nhiên

Ví dụ minh họa

Tung hai đồng xu cân đối và đồng chất. Gọi X là số đồng xu xuất hiện mặt ngửa.

a. Tìm bảng phân phối xác suất của X.

b. Tìm và vẽ đồ thị của hàm phân phối xác suất của X.

2. Phân phối xác suất của đại lượng ngẫu nhiên

a. Đặt A là biến cố tung hai đồng xu cân đối và đồng chất. Đặt $(a, b), a, b \in \{S, N\}$, trong đó S tương ứng với mặt sấp, N tương ứng với mặt ngửa, là kết quả của việc tung 2 đồng xu cân đối, đồng chất. Khi đó, không gian mẫu của biến cố A là:

$$\Omega = \begin{bmatrix} (S, S) & (S, N) \\ (N, S) & (N, N) \end{bmatrix}$$

Từ đây, ta lập được bảng phân phối xác suất của X:

X	0	1	2
P	0.25	0.5	0.25

2. Phân phối xác suất của đại lượng ngẫu nhiên

b. Từ định nghĩa hàm phân phối xác suất, ta tìm được hàm phân phối xác suất của X như sau:

$$F(x) = \begin{cases} 0, & \forall x < 0 \\ 0.25, & \forall 0 \le x < 1 \\ 0.25 + 0.5, & \forall 1 \le x < 2 \\ 0.25 + 0.5 + 0.25, & \forall x \ge 2 \end{cases}$$

hay

$$F(x) = \begin{cases} 0, & \forall x < 0 \\ 0.25, & \forall 0 \le x < 1 \\ 0.75, & \forall 1 \le x < 2 \\ 1, & \forall x \ge 2 \end{cases}$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

Định nghĩa (Hàm mật độ xác suất của đại lượng ngẫu nhiên liên tục)

Giả sử X là đại lượng ngẫu nhiên liên tục nhận giá trị trên $\mathbb{R}$.
Hàm mật độ xác suất của đại lượng ngẫu nhiên liên tục X là hàm số $f(x)$ không âm, xác định với mọi giá trị của đại lượng ngẫu nhiên X và thoả mãn tính chất:

$$P(a < x < b) = \int_a^b f(x)\mathrm{d}x, \quad \forall a, b \in \mathbb{R}, a < b.$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

Như vậy, nếu X là **đại lượng ngẫu nhiên liên tục**, thì với số c bất kỳ, ta luôn có $P(X = c) = 0$.
Từ đó, với hai số thực a, b sao cho $a < b$ ta luôn có:

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

Hàm mật độ xác suất của có các tính chất sau:

- $f(x) \geq 0, \ \forall x \in \mathbb{R}$.
- $\int_{-\infty}^{\infty} f(x)dx = 1$.

Ngược lại, một hàm số $f(x)$ thoả mãn 2 tính chất trên được gọi là hàm mật độ xác suất của một biến ngẫu nhiên nào đó.

2. Phân phối xác suất của đại lượng ngẫu nhiên

Ví dụ minh họa

Cho biến ngẫu nhiên liên tục có hàm mật độ xác suất:

$$f(x) = \begin{cases} kx^2, & \text{khi } x \in [0,2] \\ 0, & \text{khi } x \notin [0,2] \end{cases}$$

a. Tìm hằng số k.

b. Tính $P(0.5 \leq X \leq 1)$.

2. Phân phối xác suất của đại lượng ngẫu nhiên

a. Do $f(x)$ là hàm mật độ nên theo Chú ý **??** ta có:

$$\begin{cases} f(x) \geq 0, \quad x \in \mathbb{R} \\ \int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1 \end{cases}$$

Do đó, $k \geq 0$ và:

$$\int_0^2 kx^2\mathrm{d}x = 1$$

$$\Leftrightarrow k\frac{8}{3} = 1$$

$$\Leftrightarrow k = \frac{3}{8}.$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

b. Từ kết quả câu a ta tính được:

$$P(0.5 \leq X \leq 1) = \int_{0.5}^{1} \frac{3}{8} x^2 \mathrm{d}x = 0.1093.$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

Định nghĩa (Hàm phân phối xác suất cho đại lượng ngẫu nhiên liên tục)

Cho X là đại lượng ngẫu nhiên liên tục có hàm mật độ xác suất $f(x)$.
Khi đó, **hàm phân phối xác suất** của X được biểu diễn dưới dạng:

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

Tính chất của hàm phân phối (cho cả đại lượng ngẫu nhiên rời rạc và liên tục):

- $0 \leq F(x) \leq 1$,
- $P(a < x < b) = F(b) - F(a)$,
- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$,
- $F(x)$ là hàm không giảm, nghĩa là nếu $a < b$ thì $F(a) \leq F(b)$,
- $\lim_{x \to -\infty} F(x) = 0$ và $\lim_{x \to +\infty} F(x) = 1$,
- $f(x) = F'(x)$ tại x là điểm liên tục của $f(x)$.

2. Phân phối xác suất của đại lượng ngẫu nhiên

Ví dụ minh họa

Cho hàm số

$$f(x) = \begin{cases} 2x, & \text{khi } x \in [0,1] \\ 0, & \text{khi } x \notin [0,1] \end{cases}$$

a. Chứng tỏ $f(x)$ là hàm mật độ xác suất của một đại lượng ngẫu nhiên X.

b. Tìm hàm phân phối xác suất của X.

c. Tính xác suất $P(0 < X < \frac{1}{2})$.

2. Phân phối xác suất của đại lượng ngẫu nhiên

a. Từ đề bài, ta suy ra $f(X) \geq 0, \forall x \in \mathbb{R}$. Mặt khác, ta lại có:

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_0^1 2x\mathrm{d}x = 1.$$

Do đó, ta có $f(x)$ là hàm mật độ xác suất của một đại lượng ngẫu nhiên X.

2. Phân phối xác suất của đại lượng ngẫu nhiên

b.

- Với $x \le 0$ ta có

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{-\infty}^{x} 0\mathrm{d}t = 0.$$

- Với $x \in (0, 1)$ ta có

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{0}^{x} 2t\mathrm{d}t = x^2.$$

2. Phân phối xác suất của đại lượng ngẫu nhiên

- Với $x > 1$ ta có:

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \int_{-\infty}^{0} 0\mathrm{d}t + \int_{0}^{1} 2t\mathrm{d}t + \int_{0}^{x} 0\mathrm{d}t = 1.$$

- Vậy hàm phân phối $F(x)$ của X có dạng:

$$F(x) = \begin{cases} 0, & \text{khi} \quad x < 0 \\ x^2, & \text{khi} \quad x \in [0,1] \\ 1, & \text{khi} \quad x \geq 1 \end{cases}$$

c. $P(0 < X < 0.5) = F(0.5) - F(0) = \dfrac{1}{4}$.

2. Phân phối xác suất của đại lượng ngẫu nhiên

Ý nghĩa của hàm phân phối xác suất:

- Từ định nghĩa của hàm phân phối xác suất $F(x) = P(X \leq x)$, ta thấy hàm $F(x)$ phản ánh mức độ tập trung xác suất về phía bên trái của điểm x.
- Giá trị của hàm $F(x)$ cho biết có bao nhiêu phần của một đơn vị xác suất phân phối trong khoảng $(-\infty, x)$.
- Hàm $F(x)$ là một hàm không giảm nên đạo hàm $f(x)$ của nó là một hàm không âm. Về mặt hình học điều này có nghĩa là đồ thị của hàm $f(x)$ không nằm dưới trục Ox.
- Giá trị của hàm $F(x)$ tại điểm a bằng diện tích hình phẳng giới hạn bởi trục Ox, đường cong $y = f(x)$ và đường thẳng $x = a$.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên Kỳ vọng

Định nghĩa (Kỳ vọng cho đại lượng ngẫu nhiên rời rạc)

Giả sử đại lượng ngẫu nhiên rời rạc X có bảng phân phối xác suất như sau:

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

Khi đó, **kỳ vọng của X**, ký hiệu là $\mathbb{E}(X)$ được tính bởi công thức:

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i p_i.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên

Kỳ vọng

Định nghĩa (Kỳ vọng cho đại lương ngẫu nhiên liên tục)

Cho X là đại lượng ngẫu nhiên liên tục với hàm mật độ $f(x)$. Nếu $\int\limits_{-\infty}^{+\infty} xf(x)\mathrm{d}x$ hội tụ tuyệt đối thì giá trị tích phân đó được gọi là kỳ vọng của X:

$$\mathbb{E}(X) = \int\limits_{-\infty}^{+\infty} xf(x)dx.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Kỳ vọng

Tính chất của kỳ vọng:
Kỳ vọng của một đại lượng ngẫu nhiên (liên tục hoặc rời rạc) có những tính chất sau:

- $\mathbb{E}(c) = c$ với c là hằng số,
- $\mathbb{E}(cX) = c\mathbb{E}(X)$,
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$,
- $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ nếu X, Y độc lập.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Định nghĩa (Phương sai của đại lượng ngẫu nhiên rời rạc)

Phương sai của đại lượng ngẫu nhiên rời rạc X được biểu diễn dưới dạng

$$Var(X) = \sum_{i=1}^{n} (x_i - \mathbb{E}(X))^2 p_i.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Định nghĩa (Phương sai của đại lượng ngẫu nhiên liên tục)

Phương sai của đại lượng ngẫu nhiên liên tục X được biểu diễn dưới dạng:

$$Var(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x)\mathrm{d}x,$$

trong đó, $f(x)$ là hàm mật độ xác suất của X.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên Phương sai

Định nghĩa (Độ lệch chuẩn của đại lượng ngẫu nhiên)

Độ lệch chuẩn của đại lượng ngẫu nhiên X, ký hiệu là $\sigma(X)$, được tính bởi công thức:

$$\sigma(X) = \sqrt{Var(X)}.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Tính chất của phương sai

Phương sai của một đại lượng ngẫu nhiên (rời rạc hoặc liên tục) có những tính chất dưới đây:

- $Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$,
- $Var(c) = 0$ với c là hằng số,
- $Var(cX) = c^2 Var(X)$,
- $Var(X + Y) = Var(X) + Var(Y)$ nếu X, Y độc lập.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Trong tính toán người ta thường sử dụng công thức sau đây để tính **phương sai của đại lượng ngẫu nhiên rời rạc**:

$$\mathrm{Var}(X) = \sigma^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \sum_{i=1}^{n} x_i^2 p_i - [\mathbb{E}(X)]^2.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Ví dụ minh họa

Cho X là đại lượng ngẫu nhiên rời rạc có bảng phân phối xác suất:

X	1	2	3
P	0.3	0.4	0.3

Tính **kỳ vọng và phương sai** của X.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Ta có kỳ vọng của X là:

$$\mathbb{E}(X) = 1 \times 0.3 + 2 \times 0.4 + 3 \times 0.3 = 2,$$

$$\mathbb{E}(X^2) = 1^2 \times 0.3 + 2^2 \times 0.4 + 3^2 \times 0.3 = 4.6,$$

từ đó suy ra, phương sai của X là:

$$Var(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 4.6 - 4 = 0.6.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Ví dụ minh họa

Cho đại lượng ngẫu nhiên X có hàm mật độ xác suất:

$$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1], \end{cases}$$

Tính **kỳ vọng và phương sai** của X.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Phương sai

Ta có $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_0^1 x \times 1\mathrm{d}x = 0.5$.
Từ đây ta tính được:

$$Var(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 f(x)\mathrm{d}x = \int_0^1 (x - 0.5)^2\mathrm{d}x = \frac{1}{12}.$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Số yếu vị và trung vị

Định nghĩa (Số yếu vị (Mode))

Số yếu vị (hay Mode) của đại lượng ngẫu nhiên X là giá trị xuất hiện nhiều nhất trong dữ liệu.

- Nếu X rời rạc thì mode là giá trị X có xác suất cực đại.
- Nếu X liên tục, thì mode là giá trị X mà tại đó hàm mật độ xác suất nhận giá trị lớn nhất, ký hiệu là $mode(X)$.

Đại lượng ngẫu nhiên X có thể có một hay nhiều *mode*.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Số yếu vị và trung vị

Định nghĩa (Trung vị - median)

Trung vị (hay median) của đại lượng ngẫu nhiên X là trị số m thỏa điều kiện

$$P(X < m) \leq 0.5,$$

và

$$P(X > m) \leq 0.5,$$

được ký hiệu là $med(X)$.

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Số yếu vị và trung vị

Ví dụ minh họa

Tính **số yếu vị, trung vị** của đại lượng ngẫu nhiên X có bảng phân phối xác suất như sau:

X	1	2	3
P	0.3	0.4	0.3

Theo định nghĩa, ta có thể tính được

$$mode(X) = 2, med(X) = 2$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Số yếu vị và trung vị

Ví dụ minh họa

Tính **trung vị** của đại lượng ngẫu nhiên X, biết hàm mật độ xác suất như sau:

$$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1]. \end{cases}$$

3. Các tham số đặc trưng của đại lượng ngẫu nhiên
Số yếu vị và trung vị

Nếu m là trung vị của X, tức là $F(m) = 0.5$. Từ đây, ta suy ra được:

$$F(m) = 0.5 \quad \Leftrightarrow \int_{-\infty}^{m} f(x)\mathrm{d}x = 0.5$$

$$\Leftrightarrow \int_{0}^{m} 1\mathrm{d}x = 0.5 \quad \Leftrightarrow m = 0.5$$

Vậy, trung vị của X là 0.5.

4. Vector ngẫu nhiên

Định nghĩa (Vector ngẫu nhiên)

Tổ hợp của hai hay nhiều đại lượng ngẫu nhiên $\mathbf{X} = (X_1, X_2, \cdots, X_n)$, $\forall n \geq 2$ được xét đồng thời được gọi là vector ngẫu nhiên. Khi đó ta nói, $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ là vector ngẫu nhiên n chiều.

Phân loại vector ngẫu nhiên

Vector ngẫu nhiên n chiều $X = (X_1, X_2, \cdots, X_n)$ là **liên tục hay rời rạc** nếu tất cả các biến ngẫu nhiên thành phần $X_1, X_2, \cdots, X_n$ là liên tục hay rời rạc.

Trong chương này chỉ xét **vector ngẫu nhiên hai chiều rời rạc**.

4. Vector ngẫu nhiên
Bảng phân phối xác suất đồng thời

Định nghĩa (Xác suất đồng thời của hai đại lượng ngẫu nhiên)

Xác suất xảy ra đồng thời hai biến cố $X = x_i$ và $Y = y_j$ được gọi là xác suất đồng thời, ký hiệu là:

$$p_{ij} = P[(X = x_i)(Y = y_j)] = P(X = x_i; Y = y_j).$$

4. Vector ngẫu nhiên
Bảng phân phối xác suất đồng thời

Định nghĩa (Bảng phân phối xác suất đồng thời)

Bảng phân phối xác suất đồng thời của vector ngẫu nhiên hai chiều rời rạc (X, Y), trong đó X, Y lần lượt là các đại lượng ngẫu nhiên nhận các giá trị $\{x_1, x_2, \cdots, x_n\}$ và $\{y_1, y_2, \cdots, y_m\}$, trong đó

$$x_1 < x_2 < \cdots < x_n; \quad y_1 < y_2 < \cdots < y_m,$$

được biểu diễn dưới dạng sau đây:

4. Vector ngẫu nhiên
Bảng phân phối xác suất đồng thời

Định nghĩa (Bảng phân phối xác suất đồng thời (tiếp))

X╲Y	y_1	y_2	$\cdots$	y_m	$P(X) = x_i$
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}	p_{1*}
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}	p_{2*}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nm}	p_{n*}
$P(Y = y_j)$	p_{*1}	p_{*2}	$\cdots$	p_{*m}	$\mathbf{1}$

4. Vector ngẫu nhiên
Bảng phân phối xác suất đồng thời

Ví dụ minh họa

Giả sử, đại lượng ngẫu nhiên X, Y có bảng phân phối xác suất đồng thời như sau:

X\Y	1	2	3	4	$P(X = x_i)$
1	0.1	0	0.1	0	**0.2**
2	0.3	0	0.1	0.2	**0.6**
3	0	0.2	0	0	**0.2**
$P(Y = y_j)$	**0.4**	**0.2**	**0.2**	**0.2**	**1**

4. Vector ngẫu nhiên
Bảng phân phối xác suất đồng thời

Lựa chọn ngẫu nhiên một hộ gia đình tham gia khảo sát. Tính **xác suất gia đình đó có đúng 2 chiếc xe máy và 3 chiếc smartphone**?
Ta có, xác suất để gia đình đó có đúng 2 chiếc xe máy và 3 chiếc smartphone là:

$$P(X = 2; Y = 3) = p_{23} = 0.1$$

4. Vector ngẫu nhiên
Phân phối xác suất biên của đại lượng ngẫu nhiên thành phần

Định nghĩa (Xác suất biên)

Xác suất của các đại lượng ngẫu nhiên thành phần, được gọi là **xác suất biên**:

$P(X = x_i) = \sum_{j=1}^{m} P(X = x_i; Y = y_j) = \sum_{j=1}^{m} p_{ij} = p_{i*}$,

$P(Y = y_j) = \sum_{i=1}^{n} P(X = x_i; Y = y_j) = \sum_{i=1}^{n} p_{ij} = p_{*j}$

Hai đại lượng ngẫu nhiên X, Y được gọi là **độc lập với nhau** nếu:

$$P(X = x_i;\ Y = y_j) = P(X = x_i)P(Y = y_j), \quad \forall i, j$$

4. Vector ngẫu nhiên
Phân phối xác suất biên của đại lượng ngẫu nhiên thành phần

Ví dụ minh họa

Giả sử có bảng phân phối xác suất đồng thời như sau:

X \ Y	1	2	3	4	$P(X) = x_i$
1	0.1	0	0.1	0	**0.2**
2	0.3	0	0.1	0.2	**0.6**
3	0	0.2	0	0	**0.2**
$P(Y = y_j)$	**0.4**	**0.2**	**0.2**	**0.2**	**1**

4. Vector ngẫu nhiên
Phân phối xác suất biên của đại lượng ngẫu nhiên thành phần

Ví dụ minh họa

a. Lập bảng phân phối xác suất của X.
b. Lập bảng phân phối xác suất của Y.
c. Hai đại lượng X và Y có độc lập với nhau hay không? Vì sao?

4. Vector ngẫu nhiên
Phân phối xác suất biên của đại lượng ngẫu nhiên thành phần

a. Bảng phân phối xác suất của X

X	1	2	3
P	0.2	0.6	0.2

b. Bảng phân phối xác suất của Y

Y	1	2	3	4
P	0.4	0.2	0.2	0.2

4. Vector ngẫu nhiên
Phân phối xác suất biên của đại lượng ngẫu nhiên thành phần

c. Từ bảng phân phối xác suất đồng thời ta có:

$$P(X=1;Y=3)=0.1, \quad P(X=1)P(Y=3)=0.2\times 0.2=0.04$$

Do đó, hai đại lượng ngẫu nhiên X, Y không độc lập, vì $P(X=1;Y=3)\neq P(X=1)P(Y=3)$

4. Vector ngẫu nhiên
Phân phối xác suất có điều kiện

Định nghĩa (Phân phối xác suất có điều kiện)

Giả sử (X, Y) là một vector ngẫu nhiên hai chiều rời rạc. Khi đó, **xác suất có điều kiện** được xác định như sau:

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i; Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{*j}},$$

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i; Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i*}}.$$

4. Vector ngẫu nhiên
Phân phối xác suất có điều kiện

Bảng phân phối xác suất có điều kiện được biểu diễn dưới dạng:

$(X\|Y = y_j)$	x_1	x_2	$\cdots$	x_n
P	$\frac{p_{1j}}{p_{*j}}$	$\frac{p_{2j}}{p_{*j}}$	$\cdots$	$\frac{p_{nj}}{p_{*j}}$

$(Y\|X = x_i)$	y_1	y_2	$\cdots$	y_m
P	$\frac{p_{i1}}{p_{i*}}$	$\frac{p_{i2}}{p_{i*}}$	$\cdots$	$\frac{p_{im}}{p_{i*}}$

4. Vector ngẫu nhiên
Phân phối xác suất có điều kiện

Ví dụ minh họa

Cho bảng phân phối xác suất đồng thời của vector ngẫu nhiên hai chiều (X, Y) như sau:

X ＼Y	1	2	3	4	$P(X) = x_i$
1	0.1	0	0.1	0	**0.2**
2	0.3	0	0.1	0.2	**0.6**
3	0	0.2	0	0	**0.2**
$P(Y = y_j)$	**0.4**	**0.2**	**0.2**	**0.2**	**1**

4. Vector ngẫu nhiên
Phân phối xác suất có điều kiện

Ví dụ minh họa

- Lập bảng phân phối xác suất của X với điều kiện $Y = 1$.
- Lập bảng phân phối xác suất của Y với điều kiện $X = 2$.

4. Vector ngẫu nhiên
Phân phối xác suất có điều kiện

a. Từ bảng phân phối xác suất đồng thời của vector ngẫu nhiên hai chiều (X, Y) ở trên, ta suy ra bảng phân phối xác suất của X với điều kiện $Y = 1$ dưới đây:

X	1	2	3
$P(X\|Y = 1)$	$\frac{0.1}{0.4} = 0.25$	$\frac{0.3}{0.4} = 0.75$	$\frac{0}{0.4} = 0$

4. Vector ngẫu nhiên
Phân phối xác suất có điều kiện

b. Từ bảng phân phối xác suất đồng thời của vector ngẫu nhiên hai chiều (X, Y) ở trên, ta suy ra bảng phân phối xác suất của Y với điều kiện $X = 2$ dưới đây:

Y	1	2	3	4
$P(Y\|X=2)$	$\frac{0.3}{0.6} = 0.5$	$\frac{0}{0.6} = 0$	$\frac{0.1}{0.6} \approx 0.17$	$\frac{0.2}{0.6} \approx 0.33$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

Kỳ vọng:

$$\mathbb{E}(X) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i p_{ij}, \mathbb{E}(Y) = \sum_{j=1}^{m} \sum_{i=1}^{n} y_j p_{ij}$$

$$\mathbb{E}(XY) = \sum_{j=1}^{m} \sum_{i=1}^{n} x_i y_j p_{ij}.$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

Hiệp phương sai

Định nghĩa (Hiệp phương sai)

Hiệp phương sai (hay Covariance) của vector ngẫu nhiên (X, Y) được định nghĩa như sau:

$$\begin{aligned} Cov(X, Y) &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

- Nếu X, Y là hai đại lượng ngẫu nhiên độc lập thì $\mathbb{E}(XY) = \mathbb{E}(X) \times \mathbb{E}(Y)$, do đó $Cov(X, Y) = 0$. Khi đó ta nói X và Y không tương quan.
- Nếu $Cov(X, Y) \neq 0$ thì X và Y tương quan. Khi đó X, Y là hai đại lượng ngẫu nhiên không độc lập.
- Nếu $X = Y$ thì

$$Cov(X, X) = Cov(Y, Y) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = Var(X) = Var(Y).$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

Định nghĩa (Ma trận hiệp phương sai)

Ma trận tương quan hay còn gọi là **ma trận hiệp phương sai** được biểu diễn dưới dạng

$$Var(X,Y) = \begin{bmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{bmatrix} = \begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(Y,X) & Var(Y) \end{bmatrix}$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

Định nghĩa (Hệ số tương quan)

Hệ số tương quan giữa X và Y là R_{XY}, được tính bởi công thức:

$$R_{XY} = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\sigma(X)\sigma(Y)}$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

Tính chất của hệ số tương quan:

- $|R_{XY}| \leq 1$
- $|R_{XY}| = 1$ khi và chỉ khi X và Y phụ thuộc tuyến tính. Nếu $R_{XY} = 1$ ta nói X và Y có quan hệ tuyến tính với hệ số dương và ngược lại.
- Nếu X và Y độc lập với nhau thì $R_{XY} = 0$. Tuy nhiên, nếu $R_{XY} = 0$ thì chưa chắc X và Y đã độc lập, mà ta chỉ có thể nói chúng không tương quan với nhau.

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

Ví dụ minh họa

Hãy tính các đặc trưng của vector $Z = (X, Y)$, biết bảng phân phối xác suất của X và Y như sau:

X \ Y	1	2	3	$P(X = x_i)$
2	0.10	0.30	0.15	0.55
4	0.15	0.25	0.05	0.45
$P(Y = y_j)$	0.25	0.55	0.20	1

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

a. Kỳ vọng:

$E(X) = 2 \times 0.55 + 4 \times 0.45 = 2.9;$

$E(Y) = 1 \times 0.25 + 2 \times 0.55 + 3 \times 0.20 = 1.95;$

$E(XY) = 2 \times 1 \times 0.1 + 2 \times 2 \times 0.3 + 2 \times 3 \times 0.15$

$4 \times 1 \times 0.15 + 4 \times 2 \times 0.25 + 4 \times 3 \times 0.05 = 5.5.$

Vậy $E(Z) = (2.9; 1.95)$.

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

b. Phương sai:
$E(X^2) = 2^2 \times 0.55 + 4^2 \times 0.45 = 9.4;$
$E(Y^2) = 1^2 \times 0.25 + 2^2 \times 0.55 + 3^2 \times 0.2 = 4.25.$
Suy ra:

$$Var(X) = 9.4 - 2.9^2 = 0.99;$$

$$Var(Y) = 4.25 - 1.95^2 = 0.4475.$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

c. Hiệp phương sai:

$$Cov(X, Y) = 5.5 - 2.9 \times 1.95 = -0.155.$$

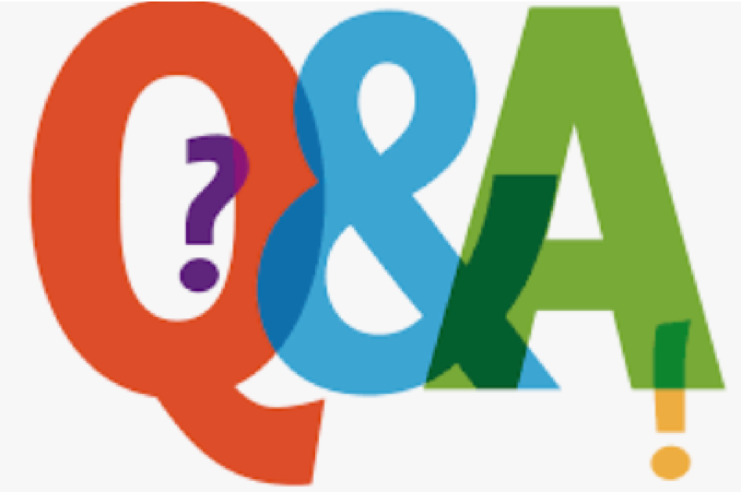
d. Hệ số tương quan:

$$R_{XY} = \frac{Cov(X, Y)}{\sqrt{\mathrm{Var(X)Var(Y)}}} = \frac{-0.155}{\sqrt{0.99 \times 0.4475}} \approx -0.233.$$

4. Vector ngẫu nhiên
Các tham số đặc trưng của đại lượng ngẫu nhiên hai chiều

e. Ma trận hiệp phương sai:

$$Var(X, Y) = \begin{bmatrix} Cov(X, X) & Cov(X, Y) \\ Cov(Y, X) & Cov(Y, Y) \end{bmatrix} = \begin{bmatrix} 0.99 & -0.155 \\ -0.155 & 0.4475 \end{bmatrix}.$$


Q&A
?
!