CHƯƠNG III
BIẾN NGẪU NHIÊN HAI CHIỀU

Môn học: Xác suất Thống kê

Giảng viên: TS. Nguyễn Kiều Linh

Hà Nội, năm 2020

Nội dung

- Biến ngẫu nhiên hai chiều và luật phân bố xác suất
- Biến ngẫu nhiên rời rạc
- Biến ngẫu nhiên liên tục
- Hàm của biến ngẫu nhiên.

Nội dung

Nội dung

BÀI 1. BIẾN NGẪU NHIÊN HAI CHIỀU VÀ LUẬT PHÂN BỐ XÁC SUẤT

1. Biến ngẫu nhiên n chiều
2. Biến ngẫu nhiên 2 chiều
3. Hàm phân bố xác suất của biến ngẫu nhiên hai chiều

BÀI 2. BIẾN NGẪU NHIÊN HAI CHIỀU RỜI RẠC

1. Bảng phân bố xác suất
2. Hiệp phương sai và hệ số tương quan

BÀI 3. BIẾN NGẪU NHIÊN HAI CHIỀU LIÊN TỤC

1. Hàm mật độ xác suất của biến ngẫu nhiên hai chiều
2. Hàm mật độ xác suất thành phần
3. Hàm mật độ xác suất có điều kiện

BÀI 4. HÀM CỦA BIẾN NGẪU NHIÊN

2. Hàm của hai biến ngẫu nhiên

1. Biến ngẫu nhiên n chiều

Biến ngẫu nhiên n chiều

(i) Trong nhiều bài toán thực tế ta thường phải xét đồng thời nhiều biến ngẫu nhiên $X_1, X_2, \dots, X_n$ có quan hệ với nhau gọi là biến ngẫu nhiên n chiều hay vecto ngẫu nhiên n chiều, ký hiệu là $X = (X_1, X_2, \dots, X_n) \subset \mathbb{R}^n$.

1. Biến ngẫu nhiên n chiều

Biến ngẫu nhiên n chiều

(i) Trong nhiều bài toán thực tế ta thường phải xét đồng thời nhiều biến ngẫu nhiên $X_1, X_2, \ldots, X_n$ có quan hệ với nhau gọi là biến ngẫu nhiên n chiều hay vecto ngẫu nhiên n chiều, ký hiệu là $X = (X_1, X_2, \ldots, X_n) \subset \mathbb{R}^n$.

- Nếu $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên rời rạc thì X được gọi là biến ngẫu nhiên rời rạc n chiều, còn khi $X_1, X_2, \ldots, X_n$ là các biến ngẫu nhiên liên tục thì X là biến ngẫu nhiên liên tục n chiều.

2. Biến ngẫu nhiên 2 chiều

Biến ngẫu nhiên 2 chiều

- Một vectơ ngẫu nhiên hai chiều là một bộ có thứ tự (X, Y) với các thành phần X, Y là các biến ngẫu nhiên.

2. Biến ngẫu nhiên 2 chiều

Biến ngẫu nhiên 2 chiều

- Một vectơ ngẫu nhiên hai chiều là một bộ có thứ tự (X, Y) với các thành phần X, Y là các biến ngẫu nhiên.
- Các kết quả thu được cho biến ngẫu nhiên hai chiều đều có thể mở rộng cho trường hợp biến ngẫu nhiên n chiều.

2. Biến ngẫu nhiên 2 chiều

Biến ngẫu nhiên 2 chiều

- Một vectơ ngẫu nhiên hai chiều là một bộ có thứ tự (X, Y) với các thành phần X, Y là các biến ngẫu nhiên.
- Các kết quả thu được cho biến ngẫu nhiên hai chiều đều có thể mở rộng cho trường hợp biến ngẫu nhiên n chiều.
- Biến ngẫu nhiên hai chiều cũng có 3 phương pháp mô tả quy luật phân bố xác suất: Bảng phân bố xác suất, hàm phân bố xác suất, hàm mật độ xác suất.

3. Hàm phân bố xác suất của biến ngẫu nhiên hai chiều

Định nghĩa

Hàm phân bố xác suất của biến ngẫu nhiên hai chiều (X, Y) được xác định như sau

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

Ta cũng gọi $F(x, y)$ là hàm phân bố xác suất đồng thời của hai biến ngẫu nhiên X và Y.

3. Hàm phân bố xác suất của biến ngẫu nhiên hai chiều

Tính chất

(i) $0 \leq F(x, y) \leq 1 \;\; \forall x, y \in \mathbb{R}$;

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

(iii) $F(-\infty, y) = F(x, -\infty) = 0, \; \forall x, y \in \mathbb{R}$ và $F(+\infty, +\infty) = 1$;

(iv) Với $x_1 < x_2, y_1 < y_2$ ta luôn có

$$\begin{aligned} P(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2) \\ &= F(x_2, y_2) + F(x_1, y_1) - F(x_1, y_2) - F(x_2, y_1). \end{aligned}$$

3. Hàm phân bố xác suất của biến ngẫu nhiên hai chiều

Tính chất (tiếp theo)

(v) Các hàm

$$F_X(x) := F(x, +\infty) = P(X < x, Y < +\infty) = P(X < x)$$

$$F_Y(y) := F(+\infty, y) = P(X < +\infty, Y < y) = P(Y < y)$$

được gọi là các hàm phân phối thành phần của các biến ngẫu nhiên X và Y và còn được gọi là các hàm phân phối biên của biến ngẫu nhiên hai chiều (X, Y).

(vi) Hai biến ngẫu nhiên X, Y được gọi là độc lập nếu

$$F(x, y) = F_X(x)F_Y(y), \ \forall x, y \in \mathbb{R}.$$

Nội dung

1. Bảng phân bố xác suất

Bảng phân bố xác suất của biến ngẫu nhiên hai chiều rời rạc

Cho biến ngẫu nhiên rời rạc hai chiều (X, Y) nhận các giá trị (x_i, y_i) với xác suất

$$p_{ij} = p(x_i, y_j) = P(X = x_i, Y = y_j), \ i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n.$$

$X \backslash Y$	y_1	$\ldots$	y_j	$\ldots$	y_n
x_1	p_{11}	$\ldots$	p_{1j}	$\ldots$	p_{1n}
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_i	p_{i1}	$\ldots$	p_{ij}	$\ldots$	p_{in}
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_m	p_{m1}	$\ldots$	p_{mj}	$\ldots$	p_{mn}

trong đó $0 \leq p_{ij} \leq 1$ và $\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} = 1$.

1.1 Bảng phân bố xác suất thành phần

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

$X \backslash Y$	y_1	$\dots$	y_j	$\dots$	y_n	$\sum_j$
x_1	p_{11}	$\dots$	p_{1j}	$\dots$	p_{1n}	$P(X = x_1)$
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_i	p_{i1}	$\dots$	p_{ij}	$\dots$	p_{in}	$P(X = x_i)$
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_m	p_{m1}	$\dots$	p_{mj}	$\dots$	p_{mn}	$X = x_m$

1.1 Bảng phân bố xác suất thành phần

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

$X \backslash Y$	y_1	$\dots$	y_j	$\dots$	y_n	$\sum_j$
x_1	p_{11}	$\dots$	p_{1j}	$\dots$	p_{1n}	$P(X = x_1)$
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_i	p_{i1}	$\dots$	p_{ij}	$\dots$	p_{in}	$P(X = x_i)$
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_m	p_{m1}	$\dots$	p_{mj}	$\dots$	p_{mn}	$X = x_m$

Khi đó,

X	x_1	$\dots$	x_i	$\dots$	x_m
$P(X = x_i)$	$P(X = x_1)$	$\dots$	$P(X = x_i)$	$\dots$	$P(X = x_m)$

trong đó $P(X = x_i) = \sum_{j=1}^{n} P(X = x_i, Y = y_j) = \sum_{j=1}^{n} p_{ij}, i = 1, 2, \dots, m.$

1.1 Bảng phân bố xác suất thành phần

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

X	x_1	$\ldots$	x_i	$\ldots$	x_m
$P(X = x_i)$	$P(X = x_1)$	$\ldots$	$P(X = x_i)$	$\ldots$	$P(X = x_m)$

trong đó $P(X = x_i) = \sum_{j=1}^{n} P(X = x_i, Y = y_j) = \sum_{j=1}^{n} p_{ij}, i = 1, 2, \ldots, m$. Khi đó kỳ vọng và phương sai của X được tính bởi các công thức sau:

$$EX = \sum_{i=1}^{m} x_i P(X = x_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i p_{ij}$$

$$EX^2 = \sum_{i=1}^{m} x_i^2 P(X = x_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i^2 p_{ij}$$

$$DX = EX^2 - (EX)^2.$$

1.1 Bảng phân bố xác suất thành phần

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

$X \backslash Y$	y_1	$\dots$	y_j	$\dots$	y_n
x_1	p_{11}	$\dots$	p_{1j}	$\dots$	p_{1n}
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_i	p_{i1}	$\dots$	p_{ij}	$\dots$	p_{in}
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_m	p_{m1}	$\dots$	p_{mj}	$\dots$	p_{mn}
$\sum_i$	$P(Y = y_1)$	$\dots$	$P(Y = y_j)$	$\dots$	$P(Y = y_n)$

1.1 Bảng phân bố xác suất thành phần

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

$X \backslash Y$	y_1	$\dots$	y_j	$\dots$	y_n
x_1	p_{11}	$\dots$	p_{1j}	$\dots$	p_{1n}
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_i	p_{i1}	$\dots$	p_{ij}	$\dots$	p_{in}
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
x_m	p_{m1}	$\dots$	p_{mj}	$\dots$	p_{mn}
$\sum_i$	$P(Y = y_1)$	$\dots$	$P(Y = y_j)$	$\dots$	$P(Y = y_n)$

Khi đó,

Y	y_1	$\dots$	y_j	$\dots$	y_n
$P(Y = y_j)$	$P(Y = y_1)$	$\dots$	$P(Y = y_i)$	$\dots$	$P(Y = y_n)$

1.1 Bảng phân bố xác suất thành phần

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

Y	y_1	$\ldots$	y_j	$\ldots$	y_n
$P(Y=y_j)$	$P(Y=y_1)$	$\ldots$	$P(Y=y_i)$	$\ldots$	$P(Y=y_n)$

trong đó $P(X=x_i) = \sum_{i=1}^{m} P(X=x_i, Y=y_j) = \sum_{i=1}^{m} p_{ij}, j=1,2,\ldots,n$. Khi đó kỳ vọng và phương sai của Y được tính bởi các công thức sau:

$$EY = \sum_{j=1}^{m} y_j P(Y=y_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} y_j p_{ij}$$

$$EY^2 = \sum_{i=1}^{m} y_j^2 P(X=x_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} y_j^2 p_{ij}$$

$$DY = EY^2 - (EY)^2.$$

1.1 Bảng phân bố xác suất thành phần

Ví dụ 1

Từ kết quả phân tích các số liệu thống kê trong tháng về doanh số bán hàng (X) và chi phí cho quảng cáo (Y) (đơn vị triệu đồng) của một công ty, thu được bảng phân bố xác suất đồng thời như sau:

Y \ X	100	200	300
1	0,15	0,2	0,04
1,5	0,05	0,2	0,15
2	0,01	0,05	0,15

(a) Tính giá trị trung bình và phương sai của doanh số bán hàng.

(b) Tính giá trị trung bình và phương sai của chi phí cho quảng cáo.

1.1 Bảng phân bố xác suất thành phần

Lời giải

Từ bảng phân tích ta lập bảng phân bố xác suất thành phần của doanh số bán hàng X và chi phí cho quảng cáo Y

X	100	200	300
P	0,21	0,45	0,34

Y	1	1,5	2
P	0,39	0,4	0,21

1.1 Bảng phân bố xác suất thành phần

Lời giải

Từ bảng phân tích ta lập bảng phân bố xác suất thành phần của doanh số bán hàng X và chi phí cho quảng cáo Y

X	100	200	300
P	0,21	0,45	0,34

Y	1	1,5	2
P	0,39	0,4	0,21

(a) Dựa vào bảng phân bố xác suất của X ta tính được giá trị trung bình và phương sai của doanh số bán hàng là $EX = 213$ và $DX = 5331$.

(b) Dựa vào bảng phân bố xác suất của Y ta tính được giá trị trung bình và phương sai của chi phí quảng cáo là $EY = 1,41$ và $DY = 0,1419$.

Bảng phân bố xác suất có điều kiện

- Bảng phân bố các xuất của X với điều kiện $(Y = y_j)$

$X\|(Y = y_j)$	x_1	x_2	$\dots$	x_m
$P(X\|(Y = y_j))$	$p(x_1\|y_j)$	$p(x_2\|y_j)$	$\dots$	$p(x_m\|y_j)$

trong đó

$$p(x_i|y_j) = P(X = x_i|Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}.$$

- Kỳ vọng của X với điều kiện $(Y = y_j)$ là

$$E(X|(Y = y_j) = \sum_{i=1}^{m} x_i p(x_i|y_j).$$

1.2 Bảng phân bố xác suất có điều kiện

Bảng phân bố xác suất có điều kiện

- Bảng phân bố các xuất của Y với điều kiện $(X = x_i)$

$Y\|(X = x_i)$	y_1	y_2	$\ldots$	y_n
$P(Y\|(X = x_i))$	$p(y_1\|x_i)$	$p(y_2\|x_i)$	$\ldots$	$p(y_n\|x_i)$

trong đó

$$p(y_j|x_i) = P(Y = y_j|X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}.$$

- Kỳ vọng của Y với điều kiện $(X = x_i)$ là

$$E(Y|(X = x_i) = \sum_{j=1}^{n} y_j p(y_j|x_i).$$

1.2 Bảng phân bố xác suất có điều kiện

Bảng phân bố xác suất có điều kiện

- Bảng phân bố các xuất của Y với điều kiện $(X = x_i)$

$Y\|(X = x_i)$	y_1	y_2	$\dots$	y_n
$P(Y\|(X = x_i))$	$p(y_1\|x_i)$	$p(y_2\|x_i)$	$\dots$	$p(y_n\|x_i)$

trong đó

$$p(y_j|x_i) = P(Y = y_j|X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}.$$

- Kỳ vọng của Y với điều kiện $(X = x_i)$ là

$$E(Y|(X = x_i)) = \sum_{j=1}^{n} y_j p(y_j|x_i).$$

- Chú ý: Hai biến ngẫu nhiên X, Y được gọi là độc lập với nhau nếu $P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j),\ \forall i = 1, \dots, m; j = 1, \dots, n.$

Ví dụ 2

Với giả thiết của Ví dụ 1: Từ bảng phân tích ta lập bảng phân bố xác suất thành phần của doanh số bán hàng X và chi phí cho quảng cáo Y

Y \ X	100	200	300
1	0,15	0,2	0,04
1,5	0,05	0,2	0,15
2	0,01	0,05	0,15

(a) Nếu chi phí cho quảng cáo là 1,5 triệu đồng thì doanh số trung bình là bao nhiêu?

(b) Nếu muốn doanh số là 300 triệu đồng thì trung bình phải chi phí cho quảng cáo là bao nhiêu?

1.2 Bảng phân bố xác suất có điều kiện

Lời giải

(a) Lập bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$

$$P(X = 100|Y = 1,5) = \frac{P(X = 100, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,05}{0,4} = 0,125$$

1.2 Bảng phân bố xác suất có điều kiện

Lời giải

(a) Lập bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$

$$P(X = 100|Y = 1,5) = \frac{P(X = 100, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,05}{0,4} = 0,125$$

$$P(X = 200|Y = 1,5) = \frac{P(X = 200, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,2}{0,4} = 0,5$$

$$P(X = 300|Y = 1,5) = \frac{P(X = 300, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,15}{0,4} = 0,375$$

1.2 Bảng phân bố xác suất có điều kiện

Lời giải

(a) Lập bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$

$$P(X = 100|Y = 1,5) = \frac{P(X = 100, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,05}{0,4} = 0,125$$

$$P(X = 200|Y = 1,5) = \frac{P(X = 200, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,2}{0,4} = 0,5$$

$$P(X = 300|Y = 1,5) = \frac{P(X = 300, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,15}{0,4} = 0,375$$

- Bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$ là

$X/(Y = 1,5)$	100	200	300
$P(X/(Y = 1,5))$	0,125	0,5	0,375

1.2 Bảng phân bố xác suất có điều kiện

Lời giải

(a) Lập bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$

$$P(X = 100|Y = 1,5) = \frac{P(X = 100, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,05}{0,4} = 0,125$$

$$P(X = 200|Y = 1,5) = \frac{P(X = 200, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,2}{0,4} = 0,5$$

$$P(X = 300|Y = 1,5) = \frac{P(X = 300, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,15}{0,4} = 0,375$$

- Bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$ là

$X/(Y = 1,5)$	100	200	300
$P(X/(Y = 1,5))$	0,125	0,5	0,375

- Từ bảng trên ta tính được doanh số trung bình khi chi phí quảng cáo là 1,5 triệu đồng.

$E(X|(Y = 1,5)) = 100.0,125 + 200.0,5 + 300.0,375 = 225.$

1.2 Bảng phân bố xác suất có điều kiện

Lời giải

(a) Lập bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$

$$P(X = 100|Y = 1,5) = \frac{P(X = 100, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,05}{0,4} = 0,125$$

$$P(X = 200|Y = 1,5) = \frac{P(X = 200, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,2}{0,4} = 0,5$$

$$P(X = 300|Y = 1,5) = \frac{P(X = 300, Y = 1,5)}{P(Y = 1,5)} = .\frac{0,15}{0,4} = 0,375$$

- Bảng phân bố xác suất có điều kiện của X với điều kiện $(Y = 1,5)$ là

$X/(Y = 1,5)$	100	200	300
$P(X/(Y = 1,5))$	0,125	0,5	0,375

- Từ bảng trên ta tính được doanh số trung bình khi chi phí quảng cáo là 1,5 triệu đồng.

$E(X|(Y = 1,5)) = 100.0,125 + 200.0,5 + 300.0,375 = 225.$

(b) tương tự câu (a). $E(Y|(X = 300)) = 1.\frac{4}{34} + 1,5.\frac{15}{34} + 2.\frac{15}{34} \simeq 1,66.$

2.1. Hiệp phương sai

Định nghĩa

Cho biến ngẫu nhiên hai chiều (X, Y), hiệp phương sai của hai thành phần X và Y, ký hiệu là $\text{cov}(X, Y)$, được xác định bởi

$$\mathrm{cov}(X, Y) = E[(X - EX)(Y - EY)] \tag{1}$$

Hay

$$\mathrm{cov}(X, Y) = EXY - EXEY \tag{2}$$

trong đó

$$EXY = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j p_{ij} \tag{3}$$

2.1. Hiệp phương sai

Ý nghĩa của hiệp phương sai

(i) Hiệp phương sai được dùng làm độ đo quan hệ giữa hai biến X và Y:
- Nếu $\mathrm{cov}(X, Y) > 0$ thì cho thấy xu thế Y tăng khi X tăng;
- Nếu $\mathrm{cov}(X, Y) < 0$ thì cho thấy xu thế Y tăng khi X giảm;

(ii) Phương sai là trường hợp riêng của hiệp phương sai khi $X = Y$ tức là $DX = \mathrm{cov}(X, X)$;

(iii) Nếu X và Y độc lập thì $EXY = EXEY$. Như vậy $\mathrm{cov}(X, Y) = 0$. Chiều ngược lại chưa chắc đúng.

Chú ý

Ta có công thức tính phương sai tổng quát cho hai biến X, Y

$$D(aX \pm bY) = a^2 DX + b^2 DY \pm 2ab\,\mathrm{cov}(X, Y).$$

2.2. Hệ số tương quan

Định nghĩa

Hệ số tương quan của hai biến ngẫu nhiên X và Y, ký hiệu ρ_{XY}, được xác định như sau

$$\rho_{XY} = \frac{\operatorname{cov}(X, Y)}{\sqrt{DXDY}} = \frac{EXY - EXEY}{\rho_X \rho_Y}. \tag{4}$$

Nhận xét

(i) Nếu $\rho_{XY} \leq 1$ với mọi (X, Y) ta nói X, Y không tương quan, ngược lại thì gọi X, Y là tương quan. Dễ thấy rằng X, Y độc lập thì X, Y không tương quan.

(ii) Nếu $\rho_{XY} = \pm 1$ ta nói hai biến ngẫu nhiên X và Y có quan hệ tuyến tính, tức là tồn tại a và b sao cho $Y = aX + b$.

Nội dung

BÀI 1. BIẾN NGẪU NHIÊN HAI CHIỀU VÀ LUẬT PHÂN BỐ XÁC SUẤT

1. Biến ngẫu nhiên n chiều
2. Biến ngẫu nhiên 2 chiều
3. Hàm phân bố xác suất của biến ngẫu nhiên hai chiều

BÀI 2. BIẾN NGẪU NHIÊN HAI CHIỀU RỜI RẠC

1. Bảng phân bố xác suất
2. Hiệp phương sai và hệ số tương quan

BÀI 3. BIẾN NGẪU NHIÊN HAI CHIỀU LIÊN TỤC

1. Hàm mật độ xác suất của biến ngẫu nhiên hai chiều
2. Hàm mật độ xác suất thành phần
3. Hàm mật độ xác suất có điều kiện

BÀI 4. HÀM CỦA BIẾN NGẪU NHIÊN

2. Hàm của hai biến ngẫu nhiên

1. Hàm mật độ xác suất của biến ngẫu nhiên hai chiều

Định nghĩa

Nếu hàm phân bố $F(x, y)$ của biến ngẫu nhiên hai chiều X, Y có dạng

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv \tag{5}$$

trong đó $f(x, y) \geq 0$, thì hàm $f(x, y)$ được gọi là hàm mật độ đồng thời của hai biến ngẫu nhiên (X, Y).

2. Hàm mật độ xác suất thành phần

Hàm mật độ xác suất thành phần

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

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy, \quad x \in \mathbb{R}.$$

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

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx, \quad y \in \mathbb{R}.$$

3. Hàm mật độ xác suất có điều kiện

Hàm mật độ xác suất có điều kiện

- Các hàm mật độ xác suất có điều kiện của biến ngẫu nhiên hai chiều X, Y có hàm mật độ xác suất $f(x, y)$ là

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)};$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}.$$

- Chú ý: Hai biến X, Y là độc lập khi và chỉ khi

$$f(x, y) = f_X(x) f_Y(y) \quad \forall (x, y) \in \mathbb{R}^2.$$

4. Ví dụ

Ví dụ

Cho biến ngẫu nhiên hai chiều (X, Y) có hàm mật độ xác suất

$$f(x,y) = \begin{cases} \frac{1}{2}\sin(x+y) & \text{nếu } 0 \le x \le \frac{\pi}{2}, 0 \le y \le \frac{\pi}{2} \\ 0 & \text{nếu trái lại.} \end{cases} \quad (6)$$

- Tính $\mathbb{P}\left(0 < Y \le X \le \frac{\pi}{2}\right)$.
- Tìm hàm mật độ xác suất của biến ngẫu nhiên X.

4. Ví dụ

Lời giải

a) Đặt

$$D = \left\{(x, y) : 0 < y \leq x \leq \frac{\pi}{2}\right\}.$$

Khi đó

$$P\left(0 < Y \leq X \leq \frac{\pi}{2}\right)$$

4. Ví dụ

Lời giải

a) Đặt

$$D = \left\{(x, y) : 0 < y \leq x \leq \frac{\pi}{2}\right\}.$$

Khi đó

$$P\left(0 < Y \leq X \leq \frac{\pi}{2}\right) = \mathbb{P}((X, Y) \in D)$$

4. Ví dụ

Lời giải

a) Đặt

$$D = \left\{(x, y) : 0 < y \leq x \leq \frac{\pi}{2}\right\}.$$

Khi đó

$$P\left(0 < Y \leq X \leq \frac{\pi}{2}\right) = \mathbb{P}((X, Y) \in D)$$

$$= \iint_D f(x, y) dxdy$$

4. Ví dụ

Lời giải

a) Đặt

$$D = \left\{(x, y) : 0 < y \le x \le \frac{\pi}{2}\right\}.$$

Khi đó

$$P\left(0 < Y \le X \le \frac{\pi}{2}\right) = \mathbb{P}((X, Y) \in D)$$

$$= \iint_D f(x, y)dxdy = \int_0^{\frac{\pi}{2}} \left(\int_0^x f(x, y)dy\right) dx$$

4. Ví dụ

Lời giải

$$= \int_0^{\frac{\pi}{2}} \left(\int_0^x \frac{1}{2} \sin(x+y) dy \right) dx$$

4. Ví dụ

Lời giải

$$= \int_0^{\frac{\pi}{2}} \left(\int_0^x \frac{1}{2} \sin(x+y) dy \right) dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} \left(-\cos(x+y)\Big|_{y=0}^{y=x} \right) dx$$

4. Ví dụ

Lời giải

$$= \int_0^{\frac{\pi}{2}} \left(\int_0^x \frac{1}{2} \sin(x+y) dy \right) dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} \left(-\cos(x+y)\Big|_{y=0}^{y=x} \right) dx$$

$$= \frac{1}{2} \int_0^{\frac{\pi}{2}} (\cos x - \cos 2x) dx$$

4. Ví dụ

Lời giải

$$= \int_0^{\frac{\pi}{2}} \left(\int_0^x \frac{1}{2} \sin(x+y)dy \right) dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} \left(-\cos(x+y)\big|_{y=0}^{y=x} \right) dx$$

$$= \frac{1}{2} \int_0^{\frac{\pi}{2}} (\cos x - \cos 2x) dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} \left(\sin x - \frac{\sin 2x}{2} \right) \Bigg|_0^{\frac{\pi}{2}}$$

4. Ví dụ

Lời giải

$$= \int_0^{\frac{\pi}{2}} \left(\int_0^{x} \frac{1}{2} \sin(x+y) dy \right) dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} \left(-\cos(x+y)\Big|_{y=0}^{y=x} \right) dx$$

$$= \frac{1}{2} \int_0^{\frac{\pi}{2}} (\cos x - \cos 2x) dx = \frac{1}{2} \int_0^{\frac{\pi}{2}} \left(\sin x - \frac{\sin 2x}{2} \right) \Bigg|_0^{\frac{\pi}{2}}$$

$$= \frac{1}{2}.$$

4. Ví dụ

Lời giải

b) Tìm hàm mật độ của biến ngẫu nhiên X.

Nếu $x < 0$ hoặc $x > \frac{\pi}{2}$ thì $f(x, y) = 0$ với mọi $y \in \mathbb{R}$, do đó

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

4. Ví dụ

Lời giải

b) Tìm hàm mật độ của biến ngẫu nhiên X.

Nếu $x < 0$ hoặc $x > \frac{\pi}{2}$ thì $f(x, y) = 0$ với mọi $y \in \mathbb{R}$, do đó

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy = \int_{-\infty}^{+\infty} 0dy$$

4. Ví dụ

Lời giải

b) Tìm hàm mật độ của biến ngẫu nhiên X.

Nếu $x < 0$ hoặc $x > \dfrac{\pi}{2}$ thì $f(x,y) = 0$ với mọi $y \in \mathbb{R}$, do đó

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\infty}^{+\infty} 0dy = 0.$$

Nếu $0 \leq x \leq \dfrac{\pi}{2}$, ta có

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

4. Ví dụ

Lời giải

b) Tìm hàm mật độ của biến ngẫu nhiên X.

Nếu $x < 0$ hoặc $x > \dfrac{\pi}{2}$ thì $f(x,y) = 0$ với mọi $y \in \mathbb{R}$, do đó

$$f_X(x) = \int\limits_{-\infty}^{+\infty} f(x,y)dy = \int\limits_{-\infty}^{+\infty} 0dy = 0.$$

Nếu $0 \leq x \leq \dfrac{\pi}{2}$, ta có

$$f_X(x) = \int\limits_{-\infty}^{+\infty} f(x,y)dy = \int\limits_{-\infty}^{0} 0dy + \int\limits_{0}^{\frac{\pi}{2}} \frac{1}{2}\sin(x+y)dy + \int\limits_{0}^{+\infty} 0dy$$

4. Ví dụ

Lời giải

$$= \frac{1}{2}\left(-\cos\left(x+y\right)\right)\Big|_{y=0}^{y=\frac{\pi}{2}}$$

4. Ví dụ

Lời giải

$$= \frac{1}{2}\left(-\cos(x+y)\right)\Big|_{y=0}^{y=\frac{\pi}{2}} = \frac{1}{2}\left(\cos x - \cos\left(x+\frac{\pi}{2}\right)\right)$$

4. Ví dụ

Lời giải

$$= \frac{1}{2}\left(-\cos(x+y)\right)\Big|_{y=0}^{y=\frac{\pi}{2}} = \frac{1}{2}\left(\cos x - \cos\left(x+\frac{\pi}{2}\right)\right) = \frac{1}{2}(\cos x + \sin x).$$

Vậy hàm mật độ xác xuất của X là

$$f_X(x) = \begin{cases} \frac{1}{2}(\cos x + \sin x) & \text{nếu} \quad 0 \le x \le \frac{\pi}{2} \\ 0 & \text{nếu ngược lại} \end{cases}$$

Nội dung

1. Hàm của một biến ngẫu nhiên

Nếu ta xác định $Z = g(X)$ là một hàm của biến ngẫu nhiên X thì Z trở thành một biến ngẫu nhiên mới. Ta sẽ tìm hàm phân phối xác suất cho Z trong một số trường hợp đơn giản.

Định nghĩa

Cho biến ngẫu nhiên X có hàm phân bố xác suất $F_X(x)$. Khi đó hàm phân phối xác suất của Z được xác định theo cách sau:

$$F_Z(z) = P(Z < z) = P(g(X) < z) = P(x \in D), \tag{7}$$

trong đó $D = \{x | g(x) < z\}$.

1. Hàm của một biến ngẫu nhiên

Ví dụ

Cho biến ngẫu nhiên X có bảng phân phối xác suất

X	-1	0	1	2	3
$P(X = x_i)$	0,1	0,2	0,3	0,2	0,2

Xác định luật phân bố xác suất của $Z = X^2$ và tìm kỳ vọng của Z.

1. Hàm của một biến ngẫu nhiên

Lời giải

Ta có $X \in \{-1, 0, 1, 2, 3\}$, ta có bảng

X	-1	0	1	2	3
Z	1	0	1	4	9

suy ra $Z \in \{0, 1, 4, 9\}$. Tính các xác suất

$P(Z = 0) = P(X = 0) = 0,2;$

1. Hàm của một biến ngẫu nhiên

Lời giải

Ta có $X \in \{-1, 0, 1, 2, 3\}$, ta có bảng

X	-1	0	1	2	3
Z	1	0	1	4	9

suy ra $Z \in \{0, 1, 4, 9\}$. Tính các xác suất

$P(Z = 0) = P(X = 0) = 0,2;$

$P(Z = 1) = P(X = 1) + P(X = -1) = 0,4$

1. Hàm của một biến ngẫu nhiên

Lời giải

Ta có $X \in \{-1, 0, 1, 2, 3\}$, ta có bảng

X	-1	0	1	2	3
Z	1	0	1	4	9

suy ra $Z \in \{0, 1, 4, 9\}$. Tính các xác suất

$P(Z = 0) = P(X = 0) = 0,2;$

$P(Z = 1) = P(X = 1) + P(X = -1) = 0,4$

$P(Z = 4) = P(X = 2) = 0,2;$

1. Hàm của một biến ngẫu nhiên

Lời giải

Ta có $X \in \{-1, 0, 1, 2, 3\}$, ta có bảng

X	-1	0	1	2	3
Z	1	0	1	4	9

suy ra $Z \in \{0, 1, 4, 9\}$. Tính các xác suất

$P(Z = 0) = P(X = 0) = 0,2;$

$P(Z = 1) = P(X = 1) + P(X = -1) = 0,4$

$P(Z = 4) = P(X = 2) = 0,2;$

$P(Z = 9) = P(X = 3) = 0,2.$

Z	0	1	4	9
$P(Z = z_i)$	0,2	0,4	0,2	0,2

Kỳ vọng của Z là $EZ = 0.0,2 + 1.0,4 + 4.0,2 + 9.0,2 = 3.$

1. Hàm của một biến ngẫu nhiên

Ví dụ

Thanh AB dài 10 cm bỗng nhiên bị gãy ở một điểm C bất kỳ. Hai đoạn AC và BC được dùng làm hai cạnh của một hình chữ nhật. Tìm hàm phân phối xác suất của biến ngẫu nhiên chỉ diện tích hình chữ nhật đó.

1. Hàm của một biến ngẫu nhiên

Ví dụ

Thanh AB dài 10 cm bỗng nhiên bị gãy ở một điểm C bất kỳ. Hai đoạn AC và BC được dùng làm hai cạnh của một hình chữ nhật. Tìm hàm phân phối xác suất của biến ngẫu nhiên chỉ diện tích hình chữ nhật đó.

Lời giải

Gọi X là biến ngẫu nhiên chỉ độ dài đoạn AC, ta có $X \sim U(1; 10)$.

1. Hàm của một biến ngẫu nhiên

Ví dụ

Thanh AB dài 10 cm bỗng nhiên bị gãy ở một điểm C bất kỳ. Hai đoạn AC và BC được dùng làm hai cạnh của một hình chữ nhật. Tìm hàm phân phối xác suất của biến ngẫu nhiên chỉ diện tích hình chữ nhật đó.

Lời giải

Gọi X là biến ngẫu nhiên chỉ độ dài đoạn AC, ta có $X \sim U(1;10)$.
Gọi Y là biến ngẫu nhiên chỉ diện tích hình chữ nhật, ta có

$$Y = X(10 - X).$$

Do $X \in (0;10)$ nên $Y \in (0;25)$.

1. Hàm của một biến ngẫu nhiên

Ví dụ

Thanh AB dài 10 cm bỗng nhiên bị gãy ở một điểm C bất kỳ. Hai đoạn AC và BC được dùng làm hai cạnh của một hình chữ nhật. Tìm hàm phân phối xác suất của biến ngẫu nhiên chỉ diện tích hình chữ nhật đó.

Lời giải

Gọi X là biến ngẫu nhiên chỉ độ dài đoạn AC, ta có $X \sim U(1;10)$.
Gọi Y là biến ngẫu nhiên chỉ diện tích hình chữ nhật, ta có

$$Y = X(10 - X).$$

Do $X \in (0;10)$ nên $Y \in (0;25)$. Vậy ta có hàm phân bố xác suất của Y là

1. Hàm của một biến ngẫu nhiên

Ví dụ

Thanh AB dài 10 cm bỗng nhiên bị gãy ở một điểm C bất kỳ. Hai đoạn AC và BC được dùng làm hai cạnh của một hình chữ nhật. Tìm hàm phân phối xác suất của biến ngẫu nhiên chỉ diện tích hình chữ nhật đó.

Lời giải

Gọi X là biến ngẫu nhiên chỉ độ dài đoạn AC, ta có $X \sim U(1; 10)$.
Gọi Y là biến ngẫu nhiên chỉ diện tích hình chữ nhật, ta có

$$Y = X(10 - X).$$

Do $X \in (0; 10)$ nên $Y \in (0; 25)$. Vậy ta có hàm phân bố xác suất của Y là
- Nếu $y \leq 0$ thì $F_Y(y) = 0$.

1. Hàm của một biến ngẫu nhiên

Lời giải

- Nếu $0 < y \leq 25$ thì

$$F_Y(y) =$$

1. Hàm của một biến ngẫu nhiên

Lời giải

- Nếu $0 < y \leq 25$ thì

$$F_Y(y) = P(Y < y) =$$

1. Hàm của một biến ngẫu nhiên

Lời giải

- Nếu $0 < y \leq 25$ thì

$$F_Y(y) = P(Y < y) = P(X(10 - X) < y) =$$

1. Hàm của một biến ngẫu nhiên

Lời giải

- Nếu $0 < y \leq 25$ thì

$$F_Y(y) = P(Y < y) = P(X(10 - X) < y) = P(X^2 - 10X + y > 0)$$

1. Hàm của một biến ngẫu nhiên

Lời giải

- Nếu $0 < y \leq 25$ thì

$$F_Y(y) = P(Y < y) = P(X(10 - X) < y) = P(X^2 - 10X + y > 0)$$
$$= P(0 < X < 5 - \sqrt{25 - y}) + P(10 > X > 5 + \sqrt{25 - y})$$

1. Hàm của một biến ngẫu nhiên

Lời giải

- Nếu $0 < y \leq 25$ thì

$$\begin{aligned} F_Y(y) &= P(Y < y) = P(X(10 - X) < y) = P(X^2 - 10X + y > 0) \\ &= P(0 < X < 5 - \sqrt{25 - y}) + P(10 > X > 5 + \sqrt{25 - y}) \\ &= \frac{5 - \sqrt{25 - y}}{5}. \end{aligned}$$

- Nếu $y > 25$ thì $F_Y(y) = 1$. Vậy hàm phân bố xác suất của Y có thể viết thành

$$F_Y(y) = \begin{cases} 0 & \text{nếu } y \leq 0 \\ \dfrac{5 - \sqrt{25 - y}}{5} & \text{nếu } 0 < y \leq 25 \\ 1 & \text{nếu } y > 25 \end{cases}$$

2. Hàm của hai biến ngẫu nhiên

Định nghĩa

Xét biến ngẫu nhiên $Z = g(X, Y)$, trong đó (X, Y) là biến ngẫu nhiên hai chiều đã biết luật phân phối xác suất. Ta sẽ xét luật phân bố xác suất của Z trong một số trường hợp đơn giản theo cách sau:

$$F_Z(z) = P(Z < z) = P(g(X, Y) < z) = P((X, Y) \in D), \tag{8}$$

trong đó $D = \{(x, y) | g(x, y) < z\}$.

2. Hàm của hai biến ngẫu nhiên

Chú ý

Đối với biên ngẫu nhiên liên tục có hàm mật độ đồng thời $f(x,y)$ thì

$$P((X,Y)\in D)=\iint\limits_{D} f(x,y)dxdy, \tag{9}$$

trong đó $D=\{(x,y)|g(x,y)<z\}$. Kỳ vọng

$$EZ=Eg(X,Y)=\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy. \tag{10}$$

2. Hàm của hai biến ngẫu nhiên

Ví dụ

Cho hai biến ngẫu nhiên X và Y độc lập có bảng phân bố xác suất:

X	-1	0	1	2
P	0,2	0,3	0,3	0,2

Y	-1	0	1
P	0,3	0,4	0,3

1. Lập bảng phân bố xác suất của các biến ngẫu nhiên $X^2, X+Y, X.Y$
2. Tính các kỳ vọng $EX, EY, E(X+Y), E(XY)$
3. Tính $P(X>Y)$.