Chương 2: Biến ngẫu nhiên

I. Khái niệm biến ngẫu nhiên

- *Biến ngẫu nhiên* là một biến mà mỗi giá trị của nó được gán tương ứng với mỗi khả năng có thể xảy ra của hiện tượng ngẫu nhiên.
(BNN là cách mô tả kết quả của p.thử dưới dạng số)
- *Biến ngẫu nhiên rời rạc* là BNN mà giá trị của nó là hữu hạn hoặc vô hạn đếm được.
- *Biến ngẫu nhiên liên tục* là BNN mà giá trị của nó có thể lấp đầy một khoảng trên trục số.
- Ký hiệu: X, Y, Z,...: biến ngẫu nhiên.
a, b, c, x, y, ...: giá trị của BNN.
$X(\Omega)$: tập giá trị của BNN X.

Ví dụ 64

Tung một đồng xu 3 lần, quy ước rằng nếu mặt ngửa (N) thì ta mất 2 đồng, sấp thì thắng 1 đồng. Tìm số tiền thu được X sau 3 lần tung.

Phép thử: "tung đồng xu 3 lần"
Không gian mẫu Ω, và giá trị BNN X:

Ω	NNN	NNS	NSN	SNN	NSS	SNS	SSN	SSS
X	-6	-3	-3	-3	0	0	0	3

$X(\Omega) = \{-6, -3, 0, 3\}$.

Ví dụ 65

Một người bắn vào bia cho tới khi nào bắn trúng thì ngừng.

Phép thử: "bắn bia tới khi trúng"
Không gian mẫu: $\Omega = \{T, \overline{T}T, \overline{T}\,\overline{T}T, ...\}$
Gọi Y là số đạn cần dùng.

$$Y(\Omega) = \{1, 2, 3, ..., n, ..\}$$

Ví dụ 66

Rót nước vào một cái can dung tích 12,1 lít.
Gọi X là số lít nước đã rót vào can.
$\Rightarrow 0 \leq X \leq 12,1$

Ví dụ 67

Quan sát các cuộc gọi đến phòng tiếp nhận thông tin một công ty bảo hiểm.
Gọi X là thời gian giữa 2 cuộc gọi liên tiếp.
$\Rightarrow X \in [0, +\infty)$

II. Quy luật phân phối xác suất

Để xác định 1 biến ngẫu nhiên X, ta cần biết
- BNN X nhận những giá trị nào, nghĩa là xác định $X(\Omega)$.
- Nhận giá trị ấy với xác suất bao nhiêu.

Quy luật phân phối xác suất của BNN nhằm chỉ ra mối quan hệ giữa các giá trị có thể có của BNN với xác suất tương ứng.

Ta thường dùng *bảng phân phối xác suất, hàm phân phối xác suất, hoặc hàm mật độ* để biểu diễn quy luật xác suất của BNN.

II.1. BNN rời rạc - bảng phân phối xác suất

Xét BNN rời rạc X có $X(\Omega) = \{x_1, x_2, ..., x_n\}$.
Ta có bảng phân phối xác suất BNN X như sau

X	x_1	x_2	...	...	x_n
P	p_1	p_2	...	...	p_n

Trong đó

i. $p_i = P(X = x_i)$: xác suất để X nhận giá trị x_i.

ii. $0 \le p_i \le 1; i = 1, 2, ..., n$

iii. $\sum_{i=1}^{n} p_i = 1$.

Quy ước

1.
$$P(X \le x_i) = P[(X = x_1)+(X = x_2)+\cdots+(X = x_i)] = P(X = x_1)+P(X = x_2)+\cdots+P(X = x_i)$$

2.
$$P(a < X < b) = \sum_{a<x_i<b} P(X = x_i)$$
$$P(a \le X < b) = \sum_{a\le x_i<b} P(X = x_i)$$
$$P(a < X \le b) = \sum_{a<x_i\le b} P(X = x_i)$$
$$P(a \le X \le b) = \sum_{a\le x_i\le b} P(X = x_i)$$

Ví dụ 68

Trở lại ví dụ 64.

Ω	NNN	NNS	NSN	SNN	NSS	SNS	SSN	SSS
X	-6	-3	-3	-3	0	0	0	3
P	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Ta có
$$P(X = -3) = P(SNN, NSN, NNS) = P(SNN) + P(NSN) + P(NNS) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}.$$

Tương tự, ta có
$$P(X = -6) = \frac{1}{8}; P(X = 0) = \frac{3}{8} \text{ và } P(X = 3) = \frac{1}{8}$$

Bảng phân phối xác suất của BNN X

X	-6	-3	0	3
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Tìm xác suất ta thua trò chơi
$$P(\text{Thua}) = P(X \le -3) = P(X = -6) + P(X = -3) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}.$$

Ví dụ 69

Cho biến ngẫu nhiên X có bảng phân phối xác suất như sau

X	-2	-1	0	1	2
P	0,2	0,4	0,1	0,2	0,1

Tính

1. $P(X \le 2)$
2. $P(X > 2)$
3. $P(-1 < X < 1)$
4. $P(X \le -1 \text{hoặc} X = 2)$

II.2. BNN rời rạc - hàm mật độ BNN rời rạc

Cho bảng phân phối xác suất

X	x_1	x_2	...	...	x_n
P	p_1	p_2	...	...	p_n

Khi đó, hàm mật độ của X xác định như sau

$$f(x) = \begin{cases} p_i, & x = x_i \\ 0, & x \ne x_i, \ \forall i \end{cases}$$

Tính chất 1

1. $f(x_i) = P(X = x_i)$.
2. $f(x) \geq 0, \forall x \in \mathbb{R}$.
3. $f(x_1) + f(x_2) + \cdots + f(x_n) = 1$.

Chú ý: đối với BNN rời rạc bảng phân phối xác suất và hàm mật độ có giá trị như nhau!

Ví dụ 70

Với bảng phân phối xác suất

X	-6	-3	0	3
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Ta có hàm mật độ

$$f(x) = \begin{cases} \frac{1}{8}, & x = -6 \\ \frac{3}{8}, & x = -3 \\ \frac{3}{8}, & x = 0 \\ \frac{1}{8}, & x = 3 \\ 0, & x \notin \{-6, -3, 0, 3\} \end{cases}$$

Ví dụ 71

Thời gian lưu trú thực tế tại khoa cấp cứu bệnh viện A năm 2009 được thể hiện trong bảng sau:

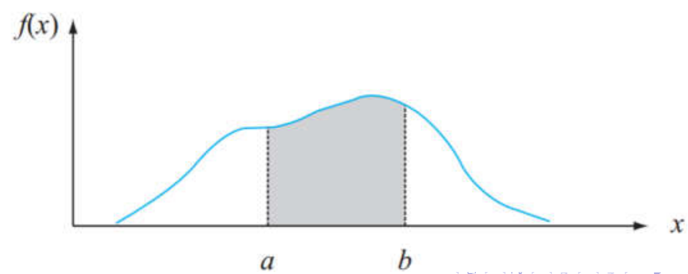
Hours	Count	Percent	Hours	Count	Percent
1	19	3.80	7	40	8.00
2	51	10.20	8	18	3.60
3	86	17.20	9	14	2.80
4	102	20.40	10	11	2.20
5	87	17.40	15	10	2.00
6	62	12.40			

Viết hàm mật độ.

II.3. BNN liên tục - hàm mật độ

Hàm số $f(x)$ được gọi là hàm mật độ của BNN *liên tục* X nếu thỏa

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



Tính chất 2

1. $P(a \leq X \leq b) = \int_a^b f(x)\, dx$
2. $P(X = a) = \int_a^a f(x)\, dx = 0$,
3. $P(a < X < b) = P(a \leq X < b)$
$= P(a < X \leq b)$
$= P(a \leq X \leq b)$

Ví dụ 72

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

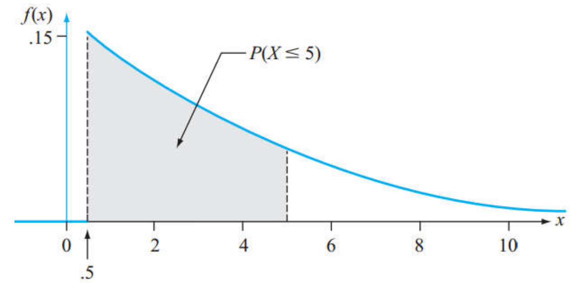
1. Chứng minh f là hàm mật độ của một đại lượng ngẫu nhiên X.
2. Tính $P(0 < X < \frac{1}{2})$

Ví dụ 73

Thời gian giữa 2 xe được chọn ngẫu nhiên trên cao tốc là một biến ngẫu nhiên có

$$f(x) = \begin{cases} 0,15e^{-0,15(x-0,5)} & x \geq 0,5 \\ 0, & \text{chỗ khác.} \end{cases}$$

1. Kiểm tra f là hàm mật độ.
2. Tính xác suất thời gian giữa 2 xe nhỏ hơn 5.

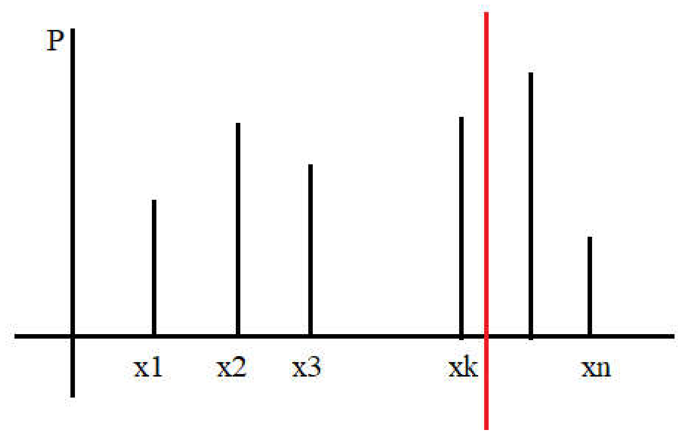


III. Hàm phân phối

1. BNN rời rạc

X	x_1	x_2	...	...	x_n
P	p_1	p_2	...	...	p_n

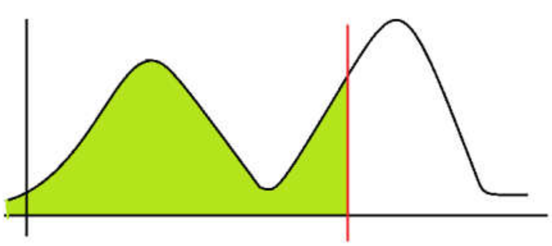
$$F(x) = \begin{cases} 0, & x \leq x_1 \\ p_1, & x_1 < x \leq x_2 \\ p_1 + p_2, & x_2 < x \leq x_3 \\ \ldots \\ p_1 + p_2 + p_3 + \cdots + p_k, & x_k < x \leq x_{k+1} \\ 1, & x > x_n \end{cases}$$



2. BNN liên tục

Cho X là BNN liên tục, ta có hàm phân phối của X là

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$



Tính chất

1. $0 \leq F(x) \leq 1, \forall x \in \mathbb{R}$,
2. $\lim\limits_{x \to -\infty} F(x) = 0$; $\lim\limits_{x \to +\infty} F(x) = 1$
3. $F(x)$ là hàm tăng,
4. $F(x)$ liên tục trái, nghĩa là $\lim\limits_{x \to a^-} F(x) = F(a)$;
5. $P(X < b) = F(b)$,
6. $P(a \leq X < b) = F(b) - F(a)$
7. $F'(x) = f(x)$

Ví dụ 74

Cho biến ngẫu nhiên X có bảng phân phối sau

X	-6	-3	0	3
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Tìm hàm phân phối của X.

Ví dụ 75

Cho biến ngẫu nhiên X có hàm mật độ

$$f(x) = \begin{cases} 0, & x < 0 \\ \frac{6}{5}x, & 0 \le x \le 1 \\ \frac{6}{5x^4}, & x > 1 \end{cases}$$

Tìm hàm phân phối của X.

Ví dụ 76

Giả sử rằng trong một ngày công ty sản xuất được 850 sản phẩm, trong đó có 50 sản phẩm không đạt yêu cầu khách hàng. Lấy ngẫu nhiên lần lượt 2 sản phẩm (không hoàn lại) để kiểm tra. Gọi X là số sản phẩm không đạt yêu cầu, lập hàm phân phối của X.

Ví dụ 77

Một giáo sư đại học không bao giờ kết thúc bài giảng của mình trước khi hết giờ và luôn hoàn thành bài giảng của mình trong vòng 2 phút sau giờ học.
Gọi X là thời gian trôi qua giữa thời điểm hết tiết học và kết thúc bài giảng của giáo sư. Giả sử hàm mật độ

$$f(x) = \begin{cases} kx^2, & 0 \le x \le 2 \\ 0, & \text{chỗ khác} \end{cases}$$

1. Tìm k và vẽ hàm mật độ tương ứng.
2. Hãy tính xác suất bài giảng kết thúc trong vòng 1 phút sau khi giờ học kết thúc.
3. Hãy tính xác suất bài giảng tiếp tục diễn ra sau khi giờ học kết thúc từ 60s tới 90s.

4. Xác suất mà bài giảng tiếp tục trong ít nhất 90s ngoài giờ kết thúc là bao nhiêu?
5. Viết hàm phân phối xác suất.

IV. Các đặc trưng của BNN

Trong cả phần này, ta chỉ xét
- BNN rời rạc X có bảng phân phối

X	x_1	x_2	...	...	x_n
P	p_1	p_2	...	...	p_n

-BNN liên tục X có hàm mật độ $f(x)$.

IV.1. Mode

Định nghĩa 0.1

Mod(X) là giá trị x_i của X mà $P(x_i)$ lớn nhất.

- X là BNN rời rạc:

$$Mod(X) = x_i \Leftrightarrow P(X = x_i) \text{ lớn nhất.}$$

- X là BNN liên tục:

$$Mod(X) = x_i \Leftrightarrow f(x_i) \text{ lớn nhất.}$$

Chú ý:

- Mod(X) có thể nhận nhiều giá trị.
- Tìm $Mod(X)$ BNN rời rạc $\Rightarrow$ tra bảng.
- Tìm $Mod(X)$ BNN liên tục $\Rightarrow$ đạo hàm.

IV.2. Median

Định nghĩa 0.2

Med(X) hay trung vị của X là điểm chia đôi phân phối xác suất của BNN X.

- X là BNN rời rạc :

$$med(X) = x_i \Leftrightarrow F(x_i) \le \frac{1}{2} \le F(x_{i+1})$$

- X là BNN liên tục :

$$med(X) = x_i \Leftrightarrow F(x_i) = \int_{-\infty}^{x_i} f(t)\,dt = 0,5$$

Ví dụ 78

Cho BNN X có bảng phân phối

X	-1	0	1	2
P	0,25	0,15	0,3	0,3

Tìm Mod(X), Med(X).

Ví dụ 79

Cho BNN X có hàm mật độ

$$f(x) = \begin{cases} \frac{2}{9}x(3-x), & x \in [0,3] \\ 0, & x \notin [0,3] \end{cases}$$

Tìm Mod(X), Med(X).

IV.3. Kỳ vọng

Kỳ vọng của BNN X, ký hiệu $E(X) \equiv \mu_X$.

- BNN rời rạc X:

$$E(X) = \sum_{i=1}^{n} x_i.p_i = x_1.p_1 + x_2.p_2 + \cdots + x_n.p_n$$

- BNN liên tục X:

$$E(X) = \int_{-\infty}^{+\infty} x.f(x)\,dx$$

Ý nghĩa

- Kỳ vọng là giá trị trung bình của BNN. Nó phản ánh giá trị trung tâm của phân phối xác suất của BNN.

Tính chất 3

.

1. $E(c) = c$.
2. $E(cX) = cE(X)$.
3. $E(X \pm Y) = E(X) \pm E(Y)$.
4. Nếu X, Y độc lập: $E(XY) = E(X).E(Y)$
5. Nếu $Y = \varphi(X)$ thì

$$E(Y) = \begin{cases} \sum_i \varphi(x_i).p_i, & X \text{rời rạc} \\ \int_{-\infty}^{+\infty} \varphi(x).f(x)\,dx, & X \text{liên tục} \end{cases}$$

Ví dụ 80

Tung đồng thời 3 súc sắc. Nếu suất hiện 3 mặt nhất thì được 1.000 USD, xuất hiện 2 mặt nhất thì được 500 USD, 1 mặt nhất thì được 100 USD, không xuất hiện mặt nhất thì không được gì. Mỗi lần chơi đóng a đồng. Hỏi a bằng bao nhiêu để trò chơi công bằng.

Giải

Gọi X là số tiền được(thua) trong 1 ván. Trò chơi công bằng nếu $E(X) = 0$. Ta có

X	-a	100-a	500-a	1000-a
P	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$

Suy ra

$E(X) = \dfrac{-216a + 16000}{216}(?)$

Vì $E(X) = 0 \Rightarrow a \approx 74$ USD.

Ví dụ 81

Cho BNN X với hàm mật độ

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

1. Tìm $E(X)$.
2. Tìm kỳ vọng của $Y = X^5 - \dfrac{2}{X}$.

IV.4. Phương sai, độ lệch chuẩn

Phương sai của BNN X, ký hiệu $Var(X) \equiv \sigma^2$

$$Var(X) = E(X^2) - [E(X)]^2$$

trong đó

-$E(X^2) = \sum_{i=1}^{n} {x_i}^2.p_i$ nếu X là BNN rời rạc.

-$E(X^2) = \int_{-\infty}^{+\infty} x^2.f(x)\,dx$ nếu X là BNN liên tục.

Độ lệch chuẩn của BNN X : $\sigma = \sqrt{Var(X)} = \sqrt{\sigma_X^2}$

Ý nghĩa

- Do $X - E(X)$ là độ lệch giữa giá trị của X so với trung bình của nó, do đó phương sai chính là trung bình của bình phương độ lệch đó.
- Phương sai đặc trưng cho độ phân tán của BNN quanh giá trị trung bình của nó: Phương sai nhỏ, độ phân tán nhỏ, độ tập trung lớn. Phương sai lớn, độ phân tán lớn, độ tập trung nhỏ.
- Trong kỹ thuật phương sai đặc trưng cho sai số thiết bị. Trong kinh doanh phương sai đặc trưng cho độ rủi ro của quyết định..

Tính chất 4

.

1. $Var(c) = 0$, c hằng số.
2. $Var(cX) = c^2.Var(X)$, c hằng số.
3. Nếu X, Y độc lập:
$var(X \pm Y) = Var(X) \pm Var(Y)$
4. $Var(X + c) = Var(X)$, c = hằng số.

Ví dụ 82

Năng suất 2 máy tương ứng là hai biến ngẫu nhiên X, Y (đơn vị : sản phẩm/phút) có phân phối xác suất

X	1	2	3	4
P	0,3	0,1	0,5	0,1

Y	2	3	4	5
P	0,1	0,4	0,4	0,1

1. Tìm kỳ vọng, phương sai của X, Y.
2. Nếu mua, ta nên chọn máy nào trong 2 máy.

Giải

1.

$EX = 1.0,3 + 2.0,1 + 3.0,5 + 4.0,1 = 2,4$

$E(X^2) = 1^2.0,3 + 2^2.0,1 + 3^2.0,5 + 4^2.0,1 = 6,8$

$Var(X) = E(X^2) - (EX)^2 = 6,8 - 2,4^2 = 1,04$

Tương tự

$EY = 2.0,1 + 3.0,4 + 4.0,4 + 5.0,1 = 3,5$

$E(Y^2) = 2^2.0,1 + 3^2.0,4 + 4^2.0,4 + 5^2.0,1 = 12,9$

$Var(Y) = E(Y^2) - (EY)^2 12,9 - 3,5^2 = 0,65$

2.

Ta thấy: $EX < EY$: năng suất trung bình của Y cao hơn X.

$Var(Y) < Var(X)$: năng suất Y ổng định hơn của X.

Vậy ta chọn máy Y.

Ví dụ 83

Trọng lượng của 1 loại sản phẩm là X(kg) có hàm mật độ.

$$f(x) = \begin{cases} \frac{3}{16}(x^2 - 1), & x \in [2,3] \\ 0, & x \notin [2,3] \end{cases}$$

Tính trọng lượng trung bình, độ lệch tiêu chuẩn của X.

Giải

$$EX = \int_{-\infty}^{+\infty} xf(x)\,dx = \frac{3}{16}\int_2^3 x(x^2-1)\,dx = 2,5781$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx = \frac{3}{16}\int_2^3 x^2(x^2-1)\,dx = 6,725$$

$\Rightarrow Var(X) = E(X^2) - (EX)^2 = 0,0784$

$\Rightarrow \sigma = \sqrt{Var(X)} = 0,28$

V. BNN hai chiều

Trong thực tế, chúng ta thường xuyên xét đồng thời nhiều BNN có quan hệ với nhau gọi là BNN nhiều chiều hay vecto ngẫu nhiên.

Ví dụ 84

1. Một máy sản suất một loại sản phẩm. Nếu ta quan tâm tới chiều dài X, chiều rộng Y của sản phẩm thì ta có một BNN hai chiều (X,Y).

2. Chọn ngẫu nhiên 1 SV. Gọi X là điểm Toán, Y là điểm Lý, Z là điểm Hóa thì ta có BNN 3 chiều (X, Y, Z).

Định nghĩa 0.3

BNN n chiều là bộ có thứ tự $V = (X_1, X_2, .., X_n)$ trong đó $X_1, ..., X_n$ là các BNN.

Nếu $X_1, ..., X_n$ là các BNN rời rạc, ta nói $V = (X_1, X_2, .., X_n)$ là BNN n chiều rời rạc.

Nếu $X_1, ..., X_n$ là các BNN liên tục, ta nói $V = (X_1, X_2, .., X_n)$ là BNN n chiều liên tục.

Trong phần này, ta chỉ xét trường hợp $V(X, Y)$ là BNN 2 chiều rời rạc.

V.1. Bảng PPXS đồng thời của BNN 2 chiều V(X,Y)

Cho BNN 2 chiều $V(X, Y)$, trong đó

$X = \{x_1, ..., x_m\}, \quad Y = \{y_1, ..., n_n\}$

Bảng PPSX đồng thời của X, Y có dạng

X \ Y	y_1	y_2	...	y_n	tổng hàng P_X
x_1	p_{11}	p_{12}	$\cdots$	p_{1n}	p_1
x_2	p_{21}	p_{22}	$\cdots$	p_{2n}	p_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	
x_m	p_{m1}	p_{m2}	$\cdots$	p_{mn}	p_m
tổng cột P_Y	p_1	p_2	$\cdots$	p_n	$\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} = 1$

Trong đó
$p_{ij} = P(X = x_i, Y = y_j)$, $1 \le i \le m, 1 \le j \le n$.
= xác suất xảy ra đồng thời hai sự kiện $(X = x_i)$ và $(Y = y_j)$
Mặt khác, ta còn có thể biểu diễn bằng hàm mật độ đồng thời

$$f(x,y) = \begin{cases} p_{ij}, & \text{khi } (x,y) = (x_i, y_j) \\ 0, & \text{khi } (x,y) \neq (x_i, y_j) \end{cases}$$

Chú ý :
1. $f(x,y) \ge 0, \forall (x,y)$ và $\sum_{x,y} f(x,y) = 1$
2. X, Y độc lập $\Leftrightarrow p_{ij} = P(X = x_i).P(Y = y_j)$

Ví dụ 85

Tung 1 xúc xắc và 1 đồng xu. Lập bảng pp xác suất đồng thời của X và Y.

X \ Y	S	N	P_X
1			
2			
3			
4			
5			
6			
P_Y			

Bảng phân phối lề

Từ bảng phân phối đồng thời của $V(X, Y)$, ta có bảng phân phối lề sau

X	x_1	x_2	...	x_m
P_X	p_1	p_2	...	p_m

Y	y_1	y_2	...	y_n
P_Y	p_1	p_2	...	p_n

Chú ý: để lập bảng phân phối lề của biến X ta tính tổng hàng, của Y ta tính tổng cột.

Ví dụ 86

Cho BNN 2 chiều $V(X, Y)$ có quy luật phân phối đồng thời như sau

X \ Y	1	2	3
4	0,18	0,22	0,16
6	0,08	0,16	0,20

1. Tìm luật phân phối của từng biến.
2. Tính $P(X = 4, Y \le 2)$

Giải 1.

X \ Y	1	2	3	P_X
4	0,18	0,22	0,16	0,56
6	0,08	0,16	0,20	0,44
P_Y	0,26	0,38	0,36	1

Vậy

X	4	6
P_X	0,56	0,44

và

Y	1	2	3
P_Y	0,25	0,38	0,36

2. Ta có
$P(X = 4, Y \le 2) = P(X = 4, Y = 1) + P(X = 4, Y = 2) = \ldots$

Xác suất có điều kiện

Giả sử $V(X, Y)$ là BNN 2 chiều rời rạc.
P(X|Y) là xác suất để X sảy ra khi biết Y đã xảy ra.
Ta có

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i; Y = y_j)}{P(Y = y_j)}$$

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i; Y = y_j)}{P(X = x_i)}$$

và

Bảng PPXS của X khi biết $Y = y_j$

X	x_1	.	x_m
$P(X\|Y = y_j)$	$P(X = x_1\|Y = y_j)$	...	$P(X = x_m\|Y = y$

Bảng PPXS của Y khi biết $X = x_i$

Y	y_1	.	y_n
$P(Y\|X = x_i)$	$P(Y = y_1\|X = x_i)$	...	$P(Y = y_n\|X = x$

Ví dụ 87

Thống kê dân số 1 vùng theo 2 chỉ tiêu: giới tính X và học vấn Y như sau

X \ Y	Thất học (0)	Phổ thông (1)	Đại học (2)
Nam (0)	0,1	0,25	0,16
Nữ (1)	0,15	0,22	0,12

1. Lập bảng PPXS học vấn, giới tính.
2. học vấn có độc lập với giới tính không.
3. Tìm xác suất để lấy ngẫu nhiên 1 người thì người đó không thất học.
4. Lập bảng PPXS của nữ, tính trung bình học vấn

VI. Hàm của các BNN

1. Hàm của 1 BNN $Y = f(X)$

Cho BNN X có bảng PPXS

X	x_1	x_2	...	x_n
P_X	p_1	p_2	...	p_n

Lập bảng PPXS của BNN $Y = f(X)$

Các giải

Bước 1: Tìm các giá trị tương ứng của Y.

X	x_1	x_2	...	x_n
Y	$f(x_1)$	$f(x_2)$	...	$f(x_n)$

Bước 2: Gom giá trị của Y, tính xác suất tương ứng.

$$P(Y = y_j) = \sum_{f(x_i)=y_j} P(X = x_i)$$

Ví dụ 88

Cho BNN X có bảng phân phối

X	-1	0	1	2
P_X	0, 1	0,3	0,4	0,2

Lập bảng PPXS của BNN $Y = X^2 + 1$.

2. Hàm của 2 BNN $Z = f(X, Y)$

Cho BNN X, và Y có bảng PPXS

X	x_1	x_2	...	x_m
P_X	p_1	p_2	...	p_m

Y	y_1	y_2	...	y_n
P_Y	p_1	p_2	...	p_n

Lập bảng PPXS của BNN $Z = f(X, Y)$

Cách giải

B1: Tìm tất cả các giá trị tương ứng của Z.

Z; X \ Y	y_1	y_2	...	y_n
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\cdots$	$f(x_1, y_n)$
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\cdots$	$f(x_2, y_n)$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\cdots$	$f(x_m, y_n)$

B2: gom giá trị của Z, tính xác suất tương ứng.

$$P(Z = z_k) = \sum_{f(x_i,y_j)=z_k} P(X = x_i, Y = y_j)$$

Ví dụ 89

Cho bảng PPXS đồng thời của (X, Y) là

X \ Y	-1	0	1
0	0,1	0,2	0,3
1	0,2	0,1	0,1

Lập bảng PPXS của BNN Z = X -Y +1.

Giải

Các giá trị có thể nhận của Z.

X \ Y	-1	0	1
0			
1			

Bảng phân phối xác suất của Z.

Z	. . . .
P	

VII. Các tham số đặc trưng

1. Kỳ vọng BNN 2 chiều V(X,Y) là

$$E(V) = (E(X), E(Y)) \in \mathbb{R}^2$$

Chú ý: kỳ vọng BNN 2 chiều là 1 vecto.

2. Kỳ vọng hàm 1 BNN

$$E(Y) = E(f(X)) = \sum_i f(x_i).p_i$$

3. Kỳ vọng hàm 2 BNN rời rạc

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

4. Kỳ vọng có điều kiện

$$E(X|Y = y_j) = \sum_{i=1}^{m} x_i.\frac{P(X = x_i; Y = y_j)}{P(Y = y_j)}$$
$$E(Y|X = x_i) = \sum_{j=1}^{n} y_j.\frac{P(X = x_i; Y = y_j)}{P(X = x_i)}$$

Chú ý: **đối với phần 1,2,3,4** ta lập bảng PPXS tương ứng rồi làm.

5. Covarian

Cho BNN V(X,Y). *covarian* của V là

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

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

6. Hệ số tương quan.

$$\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X.\sigma_Y}$$

Tính chất 5

.

1. $|\rho(X,Y)| \le 1$
2. $Var(aX \pm bY) =$
$a^2.Var(X) + b^2.Var(Y) \pm 2ab\operatorname{cov}(X,Y)$
3. X , Y độc lập $\Rightarrow$ $Cov(X,Y) = 0$
4. $Cov(X,Y) \ne 0 \Rightarrow$ X, Y phụ thuộc.

Ví dụ 90

Tính hệ số tương quan giữa học vấn và giới tính trong ví dụ 87.

Ví dụ 91

Có ba đồng tiền gồm hai đồng công bằng và một đồng thiên vị (cả hai mặt đều ngửa). Tung cả ba đồng tiền đó, gọi X là số mặt sấp, Y là số mặt ngửa.
1. Lập bảng PPXS đồng thời của X và Y, X và Y có độc lập không.
2. Tính $cov(X, Y)$.
3. Tính r_{XY}.
4. Tính $D(X, Y)$.