Chương 2
Biến ngẫu nhiên và phân phối xác suất

Mục lục

1 Các quy luật phân phối xác suất

1.1 Bảng phân phối xác suất của BNN rời rạc

+)

X	x_1	...	x_n
$p_X(x) = P(X = x)$	$P(X = x_1)$	...	$P(X = x_n)$

$$P(X = x_i) \geq 0, i = 1,2,...,n$$

$$\sum_1^n P(X = x_i) = 1$$

+) Dạng bài tìm bảng PPXS của $Z = g(X)$: Tìm tất cả các giá trị của Z bằng cách thay từng giá trị x_i vào $g(x)$

$\Rightarrow P(Z = z) = \sum_{x:g(x)=z} p_X(X)$

$Z = g(X)$	$g(x_1)$	...	$g(x_n)$
p	$P(X = X_1)$	...	$P(X = X_n)$

Ví dụ:

Bài 1: Một hộp có 10 sản phẩm, trong đó có 7 sản phẩm loại A. Lấy ngẫu nhiên 6 sản phẩm để kiểm tra. Lập bảng phân phối xác suất của X là biến ngẫu nhiên chỉ số sản phẩm loại A được lấy ra.

Bài 2: Cho biến ngẫu nhiên X có bảng phân phối xác suất:

X	-1	0	1	2
$P(X=x)$	a	0,18	0,34	$2a$

Lập bảng phân phối xác suất của: X^2

Bài 3: Một lô hàng có 18 sản phẩm, trong đó có 3 phế phẩm và 15 sản phẩm tốt. Chọn lần lượt ra 3 sản phẩm (không hoàn lại).

a) Gọi X là số phế phẩm trong 3 sản phẩm được chọn. Lập bảng PPXS của X.

b) Gọi Y là số phế phẩm trong 3 sản phẩm được chọn và đặt Z = 1 + 2Y. Lập bảng PPXS của Z.

1.2 Hàm phân phối xác suất

+) Hàm phân phối xác suất của biến ngẫu nhiên X (chung cho cả rời rạc và liên tục:)

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

Ta có: $P(\alpha \le X \le \beta) = F(\beta) - F(\alpha)$

+) Đối với BNN rời rạc:

$$F(x) = \begin{cases} 0, & x \le x_1 \\ P(X = x_1), & x_1 < x \le x_2 \\ P(X = x_1) + P(X = x_2), & x_2 < x \le x_3 \\ \dots \\ 1, & x > x_n \end{cases}$$

Ở đây mỗi điểm X_i là một điểm gián đoạn loại 1.

Chú ý dấu "=" phải chính xác

+) Đối với BNN liên tục:

Chú ý $F(x)$ liên tục trên $\mathbb{R}$: $\lim\limits_{x \to a^+} F(x) = \lim\limits_{x \to a^-} F(x)$

$P(\alpha \le X \le \beta) = P(\alpha < X \le \beta) = P(\alpha \le X < \beta) = P(\alpha < X < \beta)$

Ví dụ:

Bài 1: Một kiện hàng có 7 chiếc tivi, trong đó có 2 chiếc bị lỗi. Một khách sạn mua ngẫu nhiên 3 chiếc tivi trong số đó. Nếu x là số tivi bị lỗi mà khách sạn mua. Tìm hàm phân phối xác suất $F(X)$ của X. Dùng $F(X)$ để tìm: $P(0 < X \le 2)$

Bài 2: Cho biến ngẫu nhiên liên tục X có hàm phân phối:

$$F(x) = \begin{cases} 0, & x < \dfrac{-\pi}{2} \\ a + b\sin x, & \dfrac{-\pi}{2} \le x \le \dfrac{\pi}{2} \\ 1, & x > \dfrac{\pi}{2} \end{cases}$$

Với a, b là hằng số. Tìm a, b

Bài 3: Cho hàm phân phối xác suất của một biến ngẫu nhiên liên tục X có dạng:

$$F(x)=\begin{cases}0, & x\le 2\\ a(x-2)^2, & 2<x\le 4\\ 1, & x>4\end{cases}$$

Xác định hằng số a và tính $P(2\le x<3)$

1.3 Hàm mật độ xác suất của BNN liên tục

+) $f(x)=\dfrac{dF(x)}{dx}$.

Như vây $P(\alpha\le x\le\beta)=\displaystyle\int_{\alpha}^{\beta} f(x)dx$ (diện tích phần chắn bởi đường $y=f(x)$, trục Ox và các đường thẳng tương ứng)

+) $\displaystyle\int_{-\infty}^{+\infty} f(x)dx=1$

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

+) Nếu $f(x)=0$ khi $x\notin(\alpha;\beta)$ thì $\displaystyle\int_{\beta}^{+\infty} f(x)dx=\int_{-\infty}^{\alpha} f(x)dx=0$

Như vậy $\displaystyle\int_{-\infty}^{+\infty} f(x)dx=\int_{\alpha}^{\beta} f(x)dx=1$

+) Cách tìm hàm phấn phối từ hàm mật độ:

(1) Xác định α,β

(2) Dùng công thức dấu +) số 3 và 4, ta có:

Với $x\le\alpha$: $\quad F(x)=\displaystyle\int_{-\infty}^{x} f(x)dx=0$

Với $\alpha<x\le\beta$: $\quad F(x)=\displaystyle\int_{\alpha}^{x} f(x)dx$

Với $x>\beta$: $\quad F(x)=\displaystyle\int_{\alpha}^{\beta} f(x)dx=1$

(3) Viết hàm $F(x)$ trên các khoảng

Ví dụ:

Bài 1: Cho X là biến ngẫu nhiên có hàm mật độ xác suất $f(x)=cx+d$ với $0\le x\le 1$ và $P(X>\frac{1}{2})=\frac{1}{3}$

a) Xác định hằng số c và d

b) Tìm hàm phân phối xác suất của X

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

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

b) Tìm hàm phân phối xác suất của X

c) Tính xác suất để X nhận giá trị trong khoảng $(0;\frac{1}{2})$

Bài 3: Cho X là biến ngẫu nhiên cóc hàm mật độ $f(x)$:

$$f(x) = \begin{cases} 1+x, & -1 \leq x \leq 0 \\ 1-x, & 0 < x \leq 1 \\ 0, |x| > 1 \end{cases}$$

Tính $P\left(\frac{-1}{2} < X < 1\right)$

2 Các tham số đặc trưng

- Kì vọng (Giá trị trung bình của biến ngẫu nhiên): $E(X)$

+ Rời rạc:

$$E(X) = \sum_i x_i P(X = x_i)$$

$$E(g(X)) = \sum_i g(x_i) P(X = x_i)$$

+ Ngẫu nhiên:

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

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

+ Tính chất

$E(aX+b) = aE(X)+b$

$E(g(X) \pm h(X)) = E(g(X)) \pm E(h(X))$

- Phương sai: $V(X)$

$$V(X) = \sum (x - E(X))^2 p_X(x) \quad \text{nếu } X \text{ rời rạc}$$

$$V(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx \quad \text{nếu } X \text{ liên tục}$$

+ Định lý: $V(X) = E(X^2) - [E(X)]^2$

+ $V(aX+b) = a^2 V(X)$

- Độ lệch chuẩn: $\sigma(X) = \sqrt{V(X)}$

- Trung vị: x_{med} hoặc $medX$

$$P(X < x_{med}) = P(X \geq x_{med}) = \frac{1}{2}$$

- Mốt: $x_{mod})$ hoặc $modX$: Đối với BNN rời rạc là giá trị của X có xác suất lớn nhất; Đối với BNN liên tục là giá trị để $f(x)$ max

$$P(X = x_{mod} \geq P(X = x) \forall x$$

Ví dụ:

Bài 1: Một cơ sở sản xuất các bao kẹo. Số kẹo trong mỗi bao là một biến ngẫu nhiên X có phân phối xác suất như sau:

X	18	19	20	21	22
P	0,14	0,24	0,32	0,21	0,09

a) Tìm trung bình và phương sai của số viên kẹo trong mỗi bao

b) Hai bao kẹo được chọn ngẫu nhiên. Tính xác suất để ít nhất một trong hai bao chứa ít nhất 20 viên kẹo.

c) Chi phí sản xuất của mỗi bao kẹo là $3X+16$. Tiền báo mỗi bao kẹo là $100.000VND$. Không phân biệt số kẹo trong bao. Tìm lợi nhuận trung bình và độ lệch chuẩn của lợi nhuận trong mỗi bao kẹo

Bài 2: X là biến ngẫu nhiên có hàm xác suất:

$$p_X(x) = \begin{cases} 0,01, & \text{x=1, 2, ..., 100} \\ 0, & \text{nếu trái lại} \end{cases}$$

a) Tìm $modX$. Nếu mốt không chỉ có một, hãy tìm tập hợp các mốt của X

b) Tìm $medX$. Nếu trung vị không chỉ có một, hãy tìm tập các trung vị của X.

Bài 3: Cho X là biến ngẫu nhiên liên tục có hàm mật độ xác suất:

$$f(x) = \begin{cases} \frac{3}{4}x(2-x), & 0 \le x \le 2 \\ 0, & \text{nếu trái lại} \end{cases}$$

a) Tìm $medX$ và $modX$

b) Tìm $E(X)$

c) Tìm $V(X), E[X(X-1)]$

d) Tính $P(|X-1| > \frac{1}{2})$

Bài 4: Nếu $E(X) = 1,2$; $E(X^2) = 2,6$. Tính:

a) $E(2X-3)$

b) $E(3X^2-2X)$

Bài 5: Một người chơi trò tung con xúc xắc ba lần. Nếu cả ba lần dều xuất hiện mặt 6 thì thu về 36\$, nếu hai lần xuất hiện mặt 6 thì thu về $2,8$\$, nếu một lần xuất hiện mặt 6 thì thu về $0,4$\$. Biết rằng khi chơi người đó phải nộp x\$

a) Tìm x để ban tổ chức trò chơi không bị lỗ.

b) x bằng bao nhiêu thì trung bình mỗi lần chơi, người chơi mất 1\$?

Bài 6: Cho hàm mật độ xác suất:

$$f(x) = \begin{cases} 3e^{-3x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

của biến ngẫu nhiên liên tục X và định nghĩa $Y = [X]$ là số nguyên lớn nhất không vượt quá X (nghĩa là $[x] = 0$ nếu $0 \le x < 1$, $[x] = 1$ nếu $1 \le x < 2$...)

a) Tính $P(Y = 0)$

b) Tính $E(Y)$

3 Một số phân phối xác suất rời rạc thông dụng

3.1 Phân phối đều rời rạc: $X \sim \mathcal{U}(n)$

X	1	2	...	n
p	1/n	1/n	1/n	1/n

$$E(X) = \frac{n+1}{2} \qquad V(X) = \frac{n^2-1}{12}$$

Ví dụ: Tung một con xúc xắc. X là biến ngẫu nhiên chỉ số chấm xuất hiện. Lập bảng phân phối xác suất của X. Tìm $E(X)$ và $V(X)$

3.2 Phân phối nhị thức: $X \sim \mathcal{B}(n, p)$

trong đó n là số phép thử Bec-nu-li, p là xác suất thành công của một sự kiện A nào đó.

+) $P(X = k) = C_n^k p^k (1-p)^{n-k}$: Xác suất có k lần thành công trong n phép thử

+) $E(X) = np$ +) $V(X) = np(1-p)$

Ví dụ:

Bài 1: Bắn 5 viên đạn vào một mục tiêu. Xác suất trúng đích của mỗi lần bắn như nhau và bằng 0,2. Muốn phá hủy mục tiêu phải có ít nhất 3 viên trúng mục tiêu. Tìm xác suất mục tiêu bị phá hủy.

Bài 2: Tỷ lệ phế phẩm của lô hàng là 4%. Chọn ngẫu nhiên 20 sản phẩm để kiểm tra. Gọi X là số phế phẩm phát hiện được

a) X có phân phối gì?

b) Tính xác suất có đúng 5 phế phẩm phát hiện được

c) Lô hàng được xem là đạt tiêu chuẩn nếu số phế phẩm phát hiện được không nhiều hơn 2. Tính xác suất để lô hàng đạt tiêu chuẩn

Bài 3: Một bài thi trắc nghiệm gồm có 10 câu hỏi, mỗi câu có 4 phương án trả lời, trong đó chỉ có 1 phương án đúng. Giả sử mỗi câu trả lời đúng được 4 điểm và câu trả lời sai bị trừ 2 điểm. Một sinh viên kém làm bài bằng cách chọn ngẫu nhiên một phương án cho mỗi câu hỏi.
a) Tính xác suất để học sinh này được 4 điểm.

b) Tính xác suất để học sinh này bị điểm âm.

c) Gọi X là số câu trả lời đúng, tính $E(X)$ và $V(X)$.

d) Tính số câu sinh viên này có khả năng trả lời đúng lớn nhất.

3.3 Phân phối Poisson: $X \sim \mathcal{P}(\lambda)$

$$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$$
$$E(X) = V(X) = \lambda$$

+) Xấp xỉ phân phối nhị thức bằng phân phối Poisson:

Dùng được khi số phép thử Bec-nu-li khá lớn và p khá bé: $n \geq 100$, $p \leq 0{,}01$ và $np < 7$. Khi đó $np \to \lambda$

+) Nếu các kết cục xảy ra trong khoảng thời gian T, số các kết cục xảy ra có phân phối Poisson $\mathcal{P}(\lambda T)$

Ví dụ:

Bài 1: Một ga ra cho thuê ôtô thấy rằng số người đến thuê ôtô vào thứ bảy cuối tuần là một biến ngẫu nhiên có phân bố Poisson với tham số $\lambda = 2$. Giả sử gara có 4 chiếc ôtô.

1. Tìm xác suất để tất cả 4 ôtô đều được thuê vào thứ 7.
2. Tìm xác suất gara không đáp ứng được yêu cầu (thiếu xe cho thuê) vào thứ 7.
3. Trung bình có bao nhiêu ôtô được thuê vào ngày thứ 7?

Bài 2: Xác suất để trong khi vận chuyển mỗi chai rượu bị vỡ là 0,001. Người ta tiến hành vận chuyển 2000 chai rượu đến cửa hàng. Tìm số chai vỡ trung bình khi vận chuyển.

Bài 3: Một tổng đài bưu điện có các cuộc điện thoại gọi đến xuất hiện ngẫu nhiên, độc lập với nhau và có tốc độ trung bình 2 cuộc gọi trong 1 phút. Tính xác suất để:

(a) Có đúng 5 cuộc gọi điện thoại trung 2 phút

(b) Không có cuộc điện thoại nào trong khoảng thời gian 30 giây

(c) Có ít nhất 1 cuộc điện thoại trong khoảng thời gian 10 giây.

4 Một số phân phối xác suất liên tục thông dụng

4.1 Phân phối đều liên tục: $X \sim \mathcal{U}([a;b])$

+) Hàm mật độ: $f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a;b] \\ 0, & x \notin [a;b] \end{cases}$

+) Hàm phân phối: $F(x) = \begin{cases} 0, & x \leq a \\ \dfrac{x-a}{b-a}, & a < x \leq b \\ 1, & x > b \end{cases}$

+) $E(X) = \dfrac{a+b}{2}$ $\qquad V(X) = \dfrac{(b-a)^2}{12}$

Ví dụ:

Bài 1: Lịch chạy của xe bus tại một trạm xe bus như sau: chiếc xe bus đầu tiên trong ngày sẽ khởi hành từ trạm này lúc 7 giờ, cứ sau 15 phút sẽ có một xe khác đến trạm. Giả sử một hành khách đến trạm ngẫu nhiên trong khoảng thời gian từ 7 giờ đến 7 giờ 30. Tìm xác suất để hành khách này chờ:

(a) Ít hơn 5 phút.

(b) Ít nhất 12 phút.

Bài 2: Lấy ngẫu nhiên một điểm M trên nửa đường tròn tâm O, đường kính $AB = 2a$. Biết rằng xác suất điểm M rơi vào cung CD bất kì của nửa đường tròn AMB chỉ phụ thuộc vào độ dài cung CD. Tìm hàm phân phối xác suất của biến ngẫu nhiên Y chỉ diện tích tam giác AMB.

4.2 Phân phối chuẩn

+) $X \sim \mathcal{N}(\mu, \sigma^2)$. $E(X) = \mu$, $V(X) = \sigma^2$

+) Đặt $Z = \dfrac{X-\mu}{\sigma}$ ta được phân phối chuẩn tắc $Z \sim \mathcal{N}(0;1)$

Hàm mật độ của $\mathcal{N}(0;1)$: $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (Hàm Gao-xơ - tra giá trị trong phụ lục)

Hàm phân phối của $\mathcal{N}(0;1)$: $\Phi(x)$

+) Hàm Laplace: $\phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^x e^{-\frac{t^2}{2}} dt$ (tra giá trị trong phụ lục). Ta có:

$$\Phi(x) = 0,5 + \phi(x)$$

+) $P(x_1 < X < x_2) = \phi\left(\dfrac{x_2-\mu}{\sigma}\right) - \phi\left(\dfrac{x_1-\mu}{\sigma}\right)$

+) Nếu $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ thì $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

+) Xấp xỉ phân phối nhị thức bằng phân phối Chuẩn:

Dùng khi n khá lớn và p không quá gần 0 hoặc 1. Khi đó $\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p))$

$$P(X = a) \cong \phi\left(\frac{a+0,5-\mu}{\sigma}\right) - \phi\left(\frac{a-0,5-\mu}{\sigma}\right)$$

$$P(x_1 \le X \le x_2) \cong \phi\left(\frac{x_1+0,5-\mu}{\sigma}\right) - \phi\left(\frac{x_2-0,5-\mu}{\sigma}\right)$$

Ví dụ:

Bài 1: Giả sử X là biến ngẫu nhiên tuân theo luật phân phối chuẩn với trung bình là 3 và phương sai là 0,16.

a) Hãy tính $P(X > 3)$, $P(X > 7,784)$

b) Tìm c sao cho $P(3 - c < X < 3 + c) = 0,9$

Bài 2: Các kết quả của bải kiểm tra chỉ số thông minh (IQ) cho các học sinh của một trường tiểu học cho thấy điểm IQ của các học sinh này tuân theo phân phối chuẩn với các tham số là $\mu = 100$ và $\sigma^2 = 225$. Tỉ lệ học sinh có điểm IQ nhỏ hơn 91 hoặc lớn hơn 130 là bao nhiêu?

Bài 3: Trọng lượng của một loại trái cây tuân theo luật phân phối chuẩn với trọng lượng trung bình là 250g, độ lệch chuẩn là 5g. Trái cây loại I là trái cây có trọng lượng không nhỏ hơn 260g. 1. Một người lấy 1 trái từ trong sọt trái cây ra. Tính xác suất người này lấy được trái cây loại I. 2. Nếu lấy được trái loại I thì người này sẽ mua sọt đó. Người ngày kiểm tra 100 sọt. Tính xác suất người này mua được 6 sọt.

Bài 4: Chiều cao của nam giới khi trưởng thành là biến ngẫu nhiên tuân theo luật phân phối chuẩn với chiều cao trung bình là 160cm và độ lệch chuẩn là 6cm. Tìm xác suất để đo ngẫu nhiên 4 người thì có ít nhất một người có chiều cao nằm trong khoảng (158–162) cm.

Bài 5: Kiểm tra chất lượng 1000 sản phẩm với tỷ lệ chính phẩm 0,95. Tìm xác suất để số chính phẩm trong lô kiểm tra từ 940 đến 960

Bài 6: Dùng hai phương pháp để tính sai số của một biến ngẫu nhiên. Phương pháp 1: Cho sai số đó bằng $2X$ với X là biến ngẫu nhiên có phân phối chuẩn $\mathcal{N}(0;25)$. Phương pháp 2: Cho sai số đó bằng tổng hai biến ngẫu nhiên độc lập $Y = Y1 + Y2$ trong đó $E(Y_1) = E(Y_2) = 0$ và $\sigma(Y_1) = \sigma(Y_2) = 5$. Hỏi phương pháp nào được ưa dùng hơn?

4.3 Phân phối mũ: $X \sim \mathcal{E}(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{trái lại} \end{cases} \qquad F(X) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{trái lại} \end{cases}$$

$$E(X) = \frac{1}{\lambda} \qquad V(X) = \frac{1}{\lambda^2}$$

Ví dụ:

Bài 1: Tuổi thọ X (đơn vị: năm) của sản phẩm do nhà máy M sản xuất là biến ngẫu nhiên có phân phối mũ với tham số $\lambda = 0.2$. Mua một sản phẩm của nhà máy M. Tính xác suất sản phẩm có tuổi thọ từ 2 đến 4 năm.

Bài 2: Một linh kiện điện tử có thời gian hoạt động$X$ là biến ngẫu nhiên có phân phối mũ với hàm mật độ X là: $f(x) = \lambda e^{-\lambda x}$, với $x > 0$

(a) Xác định phân phối xác suất cho thời gian hoạt động của một mạng gồm 2 linh kiện loại trên mắc song song.

(b) Tính kỳ vọng và phương sai của thời gian hoạt động của mạng đó

Bài 3: Tuổi thọ X (đơn vị: năm) của sản phẩm do nhà máy M sản xuất là biến ngẫu nhiên có phân phối mũ với tham số $\lambda = 0.1$. Mua một sản phẩm của nhà máy M đã sử dụng rồi. Tính xác suất sử dụng sản phẩm được thêm 10 năm nữa.

Bài 4: Thời gian một người được phục vụ ở một quán cà phê là một biến ngẫu nhiên có phân phối mũ với kì vọng là 4 phút. Tính xác suất một người được phục vụ ít hơn 3 phút vào ít nhất 4 ngày trong 6 ngày tiếp theo.