Chương 2. BIẾN NGẪU NHIÊN, LUẬT PHÂN PHỐI XÁC SUẤT

§1. Luật phân phối xác suất của biến ngẫu nhiên

1. Khái niệm biến ngẫu nhiên (BNN)

a) Bài toán: Làm thế nào để dự báo được: số khách đến thuê khách sạn vào các ngày trong tuần; nhiệt độ trong các ngày của tuần … ?

b) *Biến ngẫu nhiên* (BNN)

Khái niệm: Cho F là không gian mẫu của một phép thử, *Biến ngẫu nhiên* (BNN) là hàm số xác định trên F:

i) cho tương ứng mỗi biến cố của F với một số (hay hàm số xác định trên không gian mẫu F)

ii) nhận các giá trị số với một xác suất

c) Có hai loại biến ngẫu nhiên (BNN)

Kí hiệu X(Ω) là tập giá trị của BNN X.

+ BNN rời rạc nếu X(Ω) là hữu hạn hoặc đếm được

+ BNN liên tục nếu X(Ω) là một khoảng hay một số khoảng

Ví dụ.	Tập các giá trị của X	X nhận các giá trị với xác suất	Biến ngẫu nhiên
1) Một xí nghiệp có 2 máy hoạt động. Gọi X là số máy hỏng trong một ngày làm việc của xí nghiệp	$X(\Omega)=\{ 0, 1, 2\}$	$P(X=0)= 0.9$ $P(X=1)= 0.09$ $P(X=2)=0.01$	X là rời rạc
2) X là nhiệt độ trung bình ngày mai tại TP. Hồ Chí Minh được dự báo	$X(\Omega)= (5, 50)$ (độ C)	$P(X=30)=0.7$	X liên tục
3) X là chiều cao của một người Việt Nam	$X(\Omega)= (20, 250)$ (cm)	$P(X=165)=0.45$	X liên tục

2. BNN nhiều chiều (vec tơ ngẫu nhiên)

Các BNN $(X1, X2, \ldots, Xn)$ được xét đồng thời gọi là BNN n chiều.

Biến ngẫu nhiên nhiều chiều được gọi là liên tục(rời rạc) nếu tất cả các BNN thành phần đều liên tục (rời rạc)

2. Luật phân phối xác suất của biến ngẫu nhiên (BNN)

2.1 Phân phối xác suất

1) Phân phối xác suất (probability mass function -pmf) của BNN rời rạc

Cho BNN rời rạc X có tập giá trị $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$

- **Phân phối xác suất của X là hàm P_X sao cho**

  i) Xác suất $P_X(xi) = P(X=xi) = \mathbf{p_i}$

  ii) $\sum_{xi \in X(\Omega)} \mathbf{p_i} = 1$

Hay còn gọi là bảng phân phối xác suất

X	x_1	x_2	…	x_n
$\mathbf{p_i}$	$\mathbf{p_1}$	$\mathbf{p_2}$	…	$\mathbf{p}_n$

Với $\sum_{xi \in X(\Omega)} \mathbf{p_i} = 1$

- **Tính xác suất: $P(a \leq X \leq b) = \sum_{x_i \in [a,b]} \mathbf{p_i}$ (2.1)**

_Ví dụ 1._ Mộp hộp có 10 sản phẩm trong đó có 3 phế phẩm. Lấy ngẫu nhiên 2 sản phẩm từ hộp để kiểm tra. Gọi X là số phế phẩm có trong 2 sản phẩm lấy ra.

a) Lập bảng phân phối XS của X

b) Tính XS của các biến cố:

+ P(-2 ≤ X ≤ 1,3)

+ P(-2 ≤ X < 0)

+ P(0,5 ≤ X)

Giải. a) + Tập giá trị của X ∈ (0, 1, 2)

+ P(X=0)= $\frac{C_7^2}{C_{10}^2}$ **= 21/45= 7/15**

+ P(X=1)= $\frac{3C_7^1}{C_{10}^2}$ **=21/45=7/15**

+P(X=2) = $\frac{C_3^2}{C_{10}^2}$ **=1/15**

Ví dụ 2. **Có 2 hộp sản phẩm. Hộp 1 có 10 sản phẩm trong đó dó 2 phế phẩm. Hộp 2 có 12 sản phẩm trong đó có 4 phế phẩm.**

1) **Lấy ngẫu nhiên mỗi hộp 1 sản phẩm. Lập bảng phân phối xác suất của X - số phế phẩm lấy được**

   ĐS. $X \in \{0, 1, 2\}$, $P(X=0)= 0.5333$, $P(X=1)=0.4$, $P(X=2)=0.0667$

2) **Lấy ngẫu nhiên 1 hộp từ đó lấy ngẫu nhiên ra 2 sản phẩm. Lập bảng phân phối xác suất của Y - số phế phẩm lấy được**

   ĐS. $Y \in \{0, 1, 2\}$, $P(Y=0)= 0.5232$, $P(Y=1)=0.4202$, $P(Y=2)=0.0566$

3) **Lấy ngẫu nhiên 2 sản phẩm từ hộp 1 bỏ sang hộp 2, sau đó lấy ngẫu nhiên ra một sản phẩm của hộp 2. Lập bảng phân phối xác suất của Z- số phế phẩm lấy được**

   ĐS. $Z \in \{0, 1, 2\}$, $P(Z=0)=1864/4095$, $P(Z=1)=1888/4095$, $P(Z=2)=49/585$

b) Hàm mật độ XS của biến ngẫu nhiên (BNN) liên tục

- **Định nghĩa. Hàm f(x) được gọi là hàm mật độ của BNN liên tục X nếu thỏa mãn 3 điều kiện sau:**

  1) $f(x) \geq 0$ với mọi $x \in X(\Omega)$

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

  3) $P(a \leq X \leq b) = \int_a^b f(x)dx$

- ***Chú ý:* Nếu X liên tục thì P(X=a)=0 với mọi a**

***Ví dụ 1.* Cho BNN liên tục X có hàm mật độ XS**

$$f(x) = \begin{cases} \dfrac{c}{x^2}, & x \geq 1 \\ 0, & x < 1 \end{cases}$$

1) **Tìm C**
2) **Tính XS: a) $P(-2 < X < 3)$; b) $P(X \geq 4)$**

2) Hàm phân phối của biến ngẫu nhiên (BNN)

a) Định nghĩa. Hàm phân phối của BNN X là hàm số

F(x)=P(X ≤ x)

Ta có:

$$\mathrm{F}(x) = \begin{cases} \sum_{x_i \le x} \boldsymbol{p_i}, & \textbf{với X rời rạc} \\ \int_{-\infty}^{x} \mathbf{f(y)dy}, & \textbf{với X liên tục} \end{cases}$$

b) Tính chất của hàm phân phối XS

i) 0 ≤ F(x) ≤ 1

ii) F(-∞)=0 ; F(+∞)=1

iii) F(x) không giảm, liên tục phải (Khi X BNN liên tục thì F(x) liên tục)

iv) P(a < X ≤ b) = F(b) – F(a)

c) Quan hệ giữa phân phối XS và hàm phân phối XS

i) Có phân phối XS suy ra hàm phân phối XS

$$\mathrm{F}(x) = \begin{cases} \sum_{x_i \le x} p_i \,, & \textbf{với X rời rạc} \\ \int_{-\infty}^{x} \mathbf{f(y)dy} \,, & \textbf{với X liên tục} \end{cases}$$

ii) Có hàm phân phối XS suy ra phân phối XS

+ Rời rạc: pi = F(xi)-F(x(i-1))

+Liên tục: f(x) = F(x)’ tại những điểm x mà f(x) liên tục,bằng 0 tại những điểm còn lại

d) Sự độc lập của các BNN

- **Hai biến ngẫu nhiên X và Y gọi là độc lập nếu**

P(X≤ x,Y ≤ y) = P(X ≤ x) P(Y ≤ y) với mọi x, y

Ví dụ 2. Cho BNN rời rạc X có bảng phân phối XS

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

Tìm hàm phân phối XS của X

Ví dụ 3. **Cho biến ngẫu nhiên liên tục Y có phân phối đều trên khoảng (1, 4) như sau**

$$f(x) = \begin{cases} \frac{1}{3} & khi\ x \in (1,4) \\ 0 & khi\ x\ \text{còn lại} \end{cases}$$

a) Tìm hàm phân phối của Y

b) Tính xác suất P(0.5 < Y < 3)

Chú ý.

Cho biến ngẫu nhiên liên tục Y được gọi là có phân phối đều trên khoảng (a, b) (a<b) nếu hàm mật độ của nó có dạng sau

$$f(x) = \begin{cases} \frac{1}{b-a} & khi\ x \in (a,b) \\ 0 & khi\ x\ \text{còn lại} \end{cases}$$

§2. Các số đặc trưng của BNN

1. Mốt (Mode) của BNN X. Kí hiệu ModX

a) Trường hợp rời rạc: Mod là giá trị của BNN tại đó có XS lớn nhất
   Mod X=xi $\leftrightarrow$ pi=max(pj)

Ví dụ 1. Cho BNN X có bảng phân phối xác suất:

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

Tìm ModX

+ *Ý nghĩa:* Mốt là giá trị mà BNN nhiều khả năng nhận nhất; đáng tin nhất

b) Trường hợp liên tục: ModX =a $\leftrightarrow$ f(a) = max f(x)

Ý nghĩa: Mốt là giá trị của BNN thường xuất hiện nhất trong khoảng chứa nó.

2.Kì vọng (trung bình của BNN) và mômen

2.1 Kì vọng

a) Định nghĩa. Kì vọng của BNN X, kí hiệu EX,là một số bằng

$$EX=\begin{cases} \sum_{x_i} x_i p_i & \text{Khi X là rời rạc} \\ \int_{-\infty}^{\infty} xf(x)dx & \text{khi X là liên tục} \end{cases}$$

b) *Ý nghĩa:* Kì vọng của biến ngẫu nhiên là giá trị trung bình của biến ngẫu nhiên

c) Tính chất của kì vọng

1) E(C)=C

2) E(X±Y)=EX ± EY

3) E(cX)=cEX

4) $$E(h(X))=\begin{cases} \sum_{x_i} h(x_i) p_i & \text{khi X là rời rạc} \\ \int_{-\infty}^{\infty} h(x)f(x)dx & \text{khi X liên tục} \end{cases}$$

5) Nếu X, Y độc lập: E(XY)=EXEY

2.2 Mômen của BNN

- **Định nghĩa. Mômen bậc n của X là kì vọng của X^n nếu nó tồn tại**
- **Ta có**

$$E(X^n) = \begin{cases} \sum_{x_i} (x_i)^n p_i & \text{khi X là rời rạc} \\ \int_{-\infty}^{\infty} (x)^n f(x)\,dx & \text{khi X liên tục} \end{cases}$$

Ví dụ 2. **Cho biến ngẫu nhiên X có phân phối đều trên (1, 4), nghĩa là X có hàm mật độ**

$$f(x) = \begin{cases} \frac{1}{3} & \text{khi } x \in (1, 4) \\ 0 & \text{khi x còn lại} \end{cases}$$

a) **Tìm EX**
b) **Tìm EX^2**
c) **Tìm EY với Y = 3X - 1**

***Ví dụ 3*. Cho biến ngẫu nhiên X có bảng phân phối xác suất:**

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

a) Tìm EX
b) Tìm EX^2
c) Tính EY với Y=3X - 1

3. Phương sai và độ lệch tiêu chuẩn

a) Định nghĩa. Phương sai của biến ngẫu nhiên X là một số, kí hiệu là VX hay VarX, được xác định như sau:

$$\mathbf{VX = E(X - EX)^2}$$

Ta có: VX = $\mathbf{EX^2}$ - $\mathbf{(EX)^2}$

ở đây

$$\mathbf{EX^2} = \begin{cases} \sum_{x_i} x_i^2 \, p_i & \text{khi X rời rạc} \\ \int_{-\infty}^{\infty} x^2 \, f(x)dx & \text{khi X liên tục} \end{cases}$$

b) Ý nghĩa của phương sai:Phương sai đặc trưng cho độ phân tán của giá trị BNN quanh giá trị trung bình

c) Tính chất

1) V(C) =0

2) V(CX)=C^2VX

3) Nếu X, Y độc lập

V(X+Y)= V(X)+V(Y); V(X-Y)= V(X)+V(Y)

d) Độ lệch tiêu chuẩn (Sai số chuẩn)của BNN X

$$\sigma(X) = \sqrt{VX}$$

Ví dụ 4. Cho BNN X có bảng Phân Phối XS:

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

a) Tìm VX

b) Tìm $\sigma(X)$

c) Tính VY với $Y=3X - 1$

Ví du 5. **Cho BNN liên tục có hàm mật độ XS**

$$f(x)=\begin{cases} \frac{1}{3} \text{ khi } x \in (1,4) \\ 0 \text{ khi x còn lại} \end{cases}$$

a) **Tìm VX**
b) **Tính V(-2X-3)**

§ 3. Hàm đặc trưng

1. Đinh nghĩa, tính chất

1) Định nghĩa.

Hàm đặc trưng của biến ngẫu nhiên X, kí hiệu ϕ_X(t) là kì vọng E$\{e^{itX}\}$.

- ϕ_X(t) = E$\{e^{itX}\}$= $\sum_k e^{itx_k} P_X(x_k)$, khi X là rời rạc (1)
- ϕ_X(t) = E$\{e^{itX}\}$= $\int_{-\infty}^{\infty} e^{itx} f_X$(x)dx, khi X là liên tục (2)

Ở đây t là số thực bất kì, i là đơn vị ảo: i=$\sqrt{-1}$

Chú ý. Do $|e^{itX}|$=| cos(tX) + iSin(tX)|=1 nên ϕ_X(t) tồn tại

2) Tính chất

i) $\phi_X(0) = 1$

ii) $\phi_X(-t) = \phi_X^*(t)$

iii) $|\phi_X(t)| \leq 1$

iv) Với mọi số thực a, b nếu Y=aX +b thì ϕ_Y(t) =e^{itb} ϕ_X(at) (3)

v) Cho biến ngẫu nghiên có môn men các cấp tồn tại thì

$$\phi_X(t) = 1 + \sum_{n=1}^{\infty} (it)^n \alpha_n \qquad (4)$$

Ở đây: α_n = EX^n

Từ (3) ta có:

$$\alpha_n = (i)^{-n} \phi^{(n)}(0), \quad n=1, 2, \ldots \qquad (5)$$

vi) Hàm đặc trưng ϕ_X(t) xác định duy nhất hàm phân phối

§4. Các dạng hội tụ của biến ngẫu nhiên

1. Các khái niệm

- ***Định nghĩa 1.*** **Cho dãy ($\mathbf{X_n}$) các biến ngẫu nhiên**

i) Nếu P($\omega : lim_n\, \mathrm{X_n}(\omega)$ tồn tại)=1 thì ta nói dãy ($\mathbf{X_n}$) hội tụ hầu chắc chắn (kí hiệu: h.c.c)

ii) Nếu X là biến ngẫu nhiên và P($\omega : lim_n\, \mathrm{X_n}(\omega)$ =X(ω))=1 thì ta nói dãy ($\mathbf{X_n}$) hội tụ hầu chắc chắn tới X , kí hiệu: $lim_n\, \mathbf{X_n}$ =X (h.c.c)

- ***Định nghĩa 2.*** **Cho dãy ($\mathbf{X_n}$) các biến ngẫu nhiên và biến ngẫu nhiên X. Nếu với mọi số ε >0 ta có lim_n P(|$\mathbf{X_n}$ - X |$\leq \varepsilon$) =1 thì ta nói dãy ($\mathbf{X_n}$) hội tụ tới X theo xác suất , kí hiệu: P-$lim_n\, \mathbf{X_n}$ = X hay $\mathbf{X_n} \xrightarrow{\mathbf{P}}$ X**

- *Định nghĩa 3.* Cho dãy ($\mathrm{X_n}$) các biến ngẫu nhiên và biến ngẫu nhiên X có mômen bặc p hữu hạn (P là một số dương). Dãy (X_n) hội tụ trung bình cấp p tới X nếu lim_n E($|\mathrm{X_n}$ – X$|^p$)=0
- *Định nghĩa 4.* Cho dãy ($\mathrm{X_n}$) các biến ngẫu nhiên và biến ngẫu nhiên X có các hàm phân phối F_n(x) và F(x). Ta nói dãy (X_n) hội tụ theo phân phối về X nếu tại mọi điểm x liên tục của F(x) thì $lim_n F_\mathrm{n}$(x) = F(x).

  Kí hiệu: $\mathrm{X_n} \xrightarrow{d}$ X

2. Mối quan hệ giữa các dạng hội tụ:

Định lí. Ta có mối quan hệ giữa các dạng hội tụ như sau:

i) Hội tụ (h.c.c) thì hội tụ theo xác suất (ngược lại *không* đúng)

ii) Hội tụ trung bình cấp p thì hội tụ theo xác suất

iii) Hội tụ theo xác suất thì hội tụ theo phân phối (ngược lại không đúng)

2. Luật số lớn và luật mạnh số lớn

Giả sử dãy các biến ngẫu nhiên ($\mathbf{X_n}$) độc lập, cùng phân phối. Đặt

$$\mathbf{S_n = X_1 + X_2 + \ldots + X_n}$$

***a)* Luật số lớn**

Định nghĩa 1. **Dãy ($\mathbf{X_n}$) các biến ngẫu nhiên được gọi là tuân theo *luật số lớn* đối với các dãy số thực ($a_\mathbf{n}$) và dãy số dương ($\mathbf{b_n}$) tăng ra vô hạn nếu**

$$\frac{\mathbf{S_n} - a_\mathbf{n}}{b_\mathbf{n}} \xrightarrow{\mathbf{P}} 0 \text{ , khi } n \to \infty$$

Ở đây $\mathbf{S_n} = \sum_{k=1}^{n} \mathbf{X_n}$

Bổ đề (Bất đẳng thức Trebưsep) **Nếu biến ngẫu nhiên X có kì vọng hữu hạn EX=μ và phương sai VX=σ^2 thì**

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Chú ý. **Nếu lấy ε =nσ từ bất đẳng thức *Trebưsep* suy ra**

$$P(|X - \mu| < n\sigma) \geq 1 - \frac{1}{n^2}$$

***Định lí 1.(Luật số lớn Trebưsep)* Nếu dãy biến ngẫu nhiên ($\mathbf{X_n}$) độc lập có kì vọng hữu hạn và phương sai bị chặn đều (V($\mathbf{X_n}$)$\leq$ c với mọi n) thì dãy ($\mathbf{X_n}$) tuân theo luật số lớn đối với các dãy số $\boldsymbol{a}_{\mathbf{n}}$ =E($\sum_{k=1}^{n} \mathbf{X_n}$), $\mathbf{b_n}$ =n, nghĩa là:**

$$\frac{\sum_{k=1}^{n} \mathrm{X_n} - \sum_{k=1}^{n} E\mathrm{X_n}}{n} \xrightarrow{\mathrm{P}} 0$$

***Hệ quả 1.* Nếu dãy các biến ngẫu nhiên ($\mathbf{X_n}$) độc lập, cùng phân phối và E$X_1^2 < \infty$, ta có:**

$$\overline{\mathrm{X}} = \frac{\mathrm{X_1} + \mathrm{X_2} + \ldots + \mathrm{X_n}}{n} \xrightarrow{\mathrm{P}} \mathrm{E}\mathrm{X_1}$$

***Hệ quả 2.(Luật số lớn Bernoulli)* Gọi $\mathbf{f_n}$(A) là tần suất xuất hiện biến cố A khi lặp lại n phép thử độc lập với xác suất xuất hiện biến cố A trong mỗi phép thử là p thì**

$$\mathrm{f_n}(A) \xrightarrow{\mathrm{P}} p$$

b) Luật mạnh số lớn

Định nghĩa 2. **Dãy ($\mathbf{X_n}$) các biến ngẫu nhiên được gọi là tuân theo *luật mạnh số lớn* đối với các dãy số thực ($a_{\mathbf{n}}$) và dãy số dương ($\mathbf{b_n}$) tăng ra vô hạn nếu**

$$\frac{S_n - a_n}{b_n} \rightarrow 0 \text{ (h.c.c) , khi } n \rightarrow \infty$$

Ở đây $S_n = \sum_{k=1}^{n} X_n$

Định lí 2. **Nếu dãy biến ngẫu nhiên (X_n) độc lập, E$(X_n)^2$ < ∞ với mọi n và chuỗi** $\sum_{k=1}^{\infty} \frac{V(X_n)}{n^2} < \infty$ **thì**

$$\frac{S_n - \sum_{k=1}^{n} EX_k}{n} \rightarrow 0 \text{ (h.c.c)}$$

Định lí 3. **Cho dãy biến ngẫu nhiên (X_n) độc lập,cùng phân phối và EX_1=a thì E| X_1 |< ∞ khi và chỉ khi** $\frac{S_n - na}{n} \rightarrow 0$ **(h.c.c)**

Hệ quả 1. **Nếu dãy các biến ngẫu nhiên ($\mathbf{X_n}$) độc lập, cùng phân phối và E$\mathbf{X_1}$< ∞, ta có:**

$$\overline{\mathbf{X}} = \frac{\mathbf{X_1 + X_2 + \ldots + X_n}}{n} \to \mathbf{EX_1} \quad \text{(h.c.c)}$$

Hệ quả 2. **Gọi $\mathbf{f_n}$(A) là tần suất xuất hiện biến cố A khi lặp lại n phép thử độc lập với xác suất xuất hiện biến cố A trong mỗi phép thừ là p thì**

$$\mathbf{f_n}(\mathrm{A}) \to \mathrm{p} \quad \text{(h.c.c)}$$