XÁC SUẤT CỦA BIẾN CỐ

I. CÁC TÍNH CHẤT

Quy ước: với A, B, C là 3 biến cố bất kỳ

$\bar{\bar{A}} = A$

$A + \Omega = \Omega$ $\quad$ $A.\Omega = \mathrm{A}$

$A + \varnothing = \mathrm{A}$ $\quad$ $A.\varnothing = \varnothing$

$A + \mathrm{A} = \mathrm{A}$ $\quad$ $A.\mathrm{A} = \mathrm{A}$

$A + \mathrm{B} = \mathrm{B} + \mathrm{A}$ $\quad$ $A.\mathrm{B} = \mathrm{B}.\mathrm{A}$

$A(\mathrm{B} + \mathrm{C}) = \mathrm{AB} + \mathrm{AC}$ $\quad$ $A + \mathrm{BC} = (\mathrm{A} + \mathrm{B})(\mathrm{A} + \mathrm{C})$

$\overline{A + \mathrm{B}} = \overline{A}.\overline{\mathrm{B}}$ $\quad$ $\overline{A.\mathrm{B}} = \overline{A} + \overline{\mathrm{B}}$

$\overline{A + \mathrm{B} + \mathrm{C}} = \overline{A}.\overline{\mathrm{B}}.\overline{C}$ $\quad$ $\overline{A.\mathrm{B}.\mathrm{C}} = \overline{A} + \overline{\mathrm{B}} + \overline{C}$

II. CÁC CÔNG THỨC TÍNH XÁC SUẤT

<u>Định lý 1:</u>

$0 \leq \mathrm{P}(A) \leq 1$

$\mathrm{P}(\Omega) = 1 \quad , \quad \mathrm{P}(\varnothing) = 0$

$\mathrm{P}(A + \mathrm{B}) = \mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B})$ nếu $\mathrm{A}.\mathrm{B} = \varnothing$

$P(\overline{A}) = 1 - P(A)$

Định lý 2:

Cho $A_1, A_2, \ldots, A_n$ là một họ xung khắc

Ta có: $P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$

Định lý 3: (Công thức cộng xác suất)

$P(A + B) = P(A) + P(B) – P(AB)$

Mở rộng:

1. $P(A + B + C) = P(A) + P(B) + P(C) – P(AB) – P(BC) – P(CA) + P(ABC)$
2. $P(A + B + C + D) = P(A) + P(B) + P(C) + P(D) – P(AB) – P(BC) – P(BD) – P(CA) – P(CD) – P(AD) + P(ABC) + P(BCD) + P(CDA) + P(DAB) – P(ABCD)$

III. XÁC SUẤT CÓ ĐIỀU KIỆN

$$P(A/B) = \frac{P(AB)}{P(B)}$$

Định lý 4: (Công thức nhân xác suất)

$P(AB) = P(A) . P(A/B) = P(B) . P(B/A)$

Tổng quát:

$P(ABC) = P(A/BC). P(B/C). P(C)$

$P(ABCD) = P(A/BCD). P(B/CD). P(C/D). P(D)$

$P(A_1A_2\ldots A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ \ldots P(A_n/A_1 \ldots A_{n-1})$

IV. SỰ ĐỘC LẬP

A, B độc lập nhau nếu: $P(AB) = P(A).P(B)$

V. CÔNG THỨC XÁC SUẤT ĐẦY ĐỦ

<u>Định lý 6:</u> (Công thức đầy đủ)

$$P(F) = P(A_1)P(F/A_1) + \cdots + P(A_n)P(F/A_n)$$

<u>Định lý 7:</u> (Công thức Bayès)

$$P(A_i/F) = \frac{P(A_i)P(F/A_i)}{P(F)}$$

ĐẠI LƯỢNG NGẪU NHIÊN VÀ HÀM CỦA ĐẠI LƯỢNG NGẪU NHIÊN

I. BẢNG PHÂN PHỐI XÁC SUẤT

Cho X là ĐLNN rời rạc, ta có:

$X(\Omega) = \{ x_1, x_2, \dots, x_n \}$ và $P(X = x_i) = p_i$

Bảng sau đây:

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

được gọi là <u>bảng phân phối xác suất</u> của ĐLNN rời rạc X.

II. HÀM PHÂN PHỐI XÁC SUẤT

<u>Định nghĩa:</u> Cho X là một ĐLNN.

Ánh xạ F: $\mathbb{R} \longrightarrow [0,1]$ định bởi

$$x \longrightarrow F(x) = P(X < x)$$

được gọi là hàm phân phối xác suất của ĐLNN X.

<u>Mệnh đề 1:</u> Cho X là ĐLNN rời rạc, có:

$X(\Omega) = \{ x_1, x_2, \dots, x_n \}$ và $p_i = P(X = x_i)$, và F(x) là hàm phân phối xác suất của X.

Ta có:

$$F(x) = \begin{cases} 0 & nếu\ x < x_1 \\ p_1 + p_2 + \cdots + p_i & nếu\ x_i < x < x_{i+1} \\ p_1 + p_2 + \cdots + p_n = 1 & nếu\ x_n < x \end{cases}$$

Mệnh đề 2: Cho X là ĐLNN liên tục có F(x) là hàm phân phối xác suất của nó. Ta có:

1) F(x) là hàm tăng và liên tục.

2) $\lim_{x \to -\infty} F(x) = 0$

3) $\lim_{x \to +\infty} F(x) = 1$

Mệnh đề 3: Cho X là ĐLNN rời rạc, có:

$X(\Omega) = \{ x_1, x_2, \ldots, x_n \}$, $p_i = P(X = x_i)$, và F(x) là hàm phân phối xác suất của X.

Ta có:

1) $P(X = x_i) = F(x_{i+1}) - F(x_i)$
2) $P(a \le X < b) = F(b) - F(a)$

Mệnh đề 4: Cho X là ĐLNN liên tục có F(x) là hàm phân phối xác suất của nó. Ta có:

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

III. HÀM MẬT ĐỘ

Định nghĩa: Cho X là ĐLNN liên tục có F(x) là hàm phân phối xác suất của nó. Hàm sau đây:

$$f(x) = F'(x), \qquad \forall\, x \in \mathbb{R}$$

được gọi là hàm mật độ của ĐLNN liên tục X.

Định lý: (Tính chất của hàm mật độ)

Cho f(x) là hàm mật độ và F(x) là hàm phân phối xác suất của ĐLNN liên tục X. Ta có:

$$\mathbf{1)}\qquad \mathbf{f}(\boldsymbol{x}) \geq \mathbf{0}, \forall \boldsymbol{x} \in \mathbb{R}$$

$$2)\qquad F(x) = \int_{-\infty}^{x} f(t)dt$$

$$3)\ P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

$$4)\qquad \int_{-\infty}^{+\infty} f(x)dx = 1$$

IV. KỲ VỌNG, PHƯƠNG SAI, ĐỘ LỆCH CHUẨN

❖ X là ĐLNN rời rạc có bảng phân phối xác suất:

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Kỳ vọng của X :

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

Kỳ vọng của X^2 :

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

Phương sai của X :

$$D(X)=var(X)=\sum_{i=1}^{n}\left(x_i-E(X)\right)^2 p_i$$

❖ X là ĐLNN liên tục có hàm mật độ xác suất f(x):

Kỳ vọng của X :

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

Kỳ vọng của X^2 :

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

Phương sai của X :

$$D(X)=var(X)=\int_{-\infty}^{+\infty}\left(x_i-E(X)\right)^2 f(x)dx$$

❖ Độ lệch chuẩn của X: $\sigma(X) = \sqrt{D(X)}$

$\sigma(X)$ cùng đơn vị đo với X

<u>Định lý 1:</u>

1) $E(C) = C$ với C : ĐLNN hằng số
2) $E(X \pm Y) = E(X) \pm E(Y)$
3) $E(\lambda X) = \lambda E(X)$ $\lambda \in \mathbb{R}$
4) $E(X.Y) = E(X).E(Y)$ nếu X, Y độc lập nhau

<u>Định lý 2:</u>

1) $D(C) = 0$ với C : ĐLNN hằng số
2) $D(X) = E(X^2) - [E(X)]^2$
3) $D(\lambda X) = \lambda^2.D(X)$ $\lambda \in \mathbb{R}$
4) $D(X + \lambda) = D(X)$ $\lambda \in \mathbb{R}$
5) $D(X) \geq 0$, $\forall X$

   $D(X) = 0 \Leftrightarrow$ X : ĐLNN hằng số
6) $D(X + Y) = D(X) + D(Y)$ nếu X, Y độc lập nhau

CÁC QUY LUẬT PHÂN PHỐI XÁC SUẤT

I. PHÂN PHỐI NHỊ THỨC

Ký hiệu: X ~ B(n, p)

Công thức xác suất:

$$P(X = k) = C_n^k . p^k . q^{n-k} \quad ; \quad q = 1 - p \qquad k = 0, 1, 2, \dots . n$$

Tính chất:

$$E(X) = np \, ; \, D(X) = npq$$

$$np - q \leq \mathrm{mod}(X) \leq np - q + 1$$

II. PHÂN PHỐI SIÊU BỘI

Ký hiệu: X ~ H(N, M, n)

Công thức xác suất:

$$P(X = k) = \frac{C_M^k . C_{N-M}^{n-k}}{C_N^n}$$

Tính chất:

$$E(X) = np \left(\text{với } p = \frac{M}{N} \text{ và } q = 1 - p\right); \; D(X) = npq . \frac{N - n}{N - 1}$$

$$\frac{N-n}{N-1} \text{ gọi là hệ số hiệu chỉnh}$$

III. PHÂN PHỐI POISSON

Ký hiệu: X ~ P(λ)

Công thức xác suất:

$$P(X=k)=\frac{e^{-\lambda}.\lambda^{k}}{k!} \quad ; \quad (\lambda>0) \qquad k=0,1,2,\ldots.n$$

Tính chất:

$$E(X) = D(X) = \lambda$$

$$\lambda-1 \leq \mod(X) \leq \lambda$$

IV. PHÂN PHỐI CHUẨN

Ký hiệu: X ~ N(μ, σ^2)

$$\text{Hàm mật độ: } f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^{2}}$$

Tính chất 1:

$$E(X) = \mu \,;\; D(X) = \sigma^2 \,;\; \mod(X) = \operatorname{med}(X) = \mu$$

Tính chất 2:

$$P(\alpha < X < \beta) = \phi\left(\frac{\beta - \mu}{\sigma}\right) - \phi\left(\frac{\alpha - \mu}{\sigma}\right)$$

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

$$P(X > \alpha) = 1 - P(X \leq \alpha) = 0{,}5 - \phi\left(\frac{\alpha - \mu}{\sigma}\right)$$

$$P(|X - \mu| < k.\sigma) = 2\phi(k)$$

V. CÁC CÔNG THỨC XẤP XỈ

1. $X \sim H(N, M, n)$ (X có phân phối siêu bội)

   Khi n nhỏ hơn rất nhiều so với N ($n \ll N$) ta xấp xỉ: $X \sim B(n, p)$ với $p = M/N$

2. $X \sim B(n, p)$ (X có phân phối nhị thức)

a) Khi n lớn, p nhỏ gần 0 thì ta xấp xỉ: $X \sim P(np)$

Thông thường: $X \sim B(n, p)$ có $n \geq 30$, $p \leq 0{,}1$ và $np \leq 5$ thì ta xấp xỉ $X \sim P(np)$

b) Khi n lớn, p không quá gần 0 và 1 thì ta xấp xỉ: $X \sim N(np, npq)$ với $q = 1 - p$

$$P(X = k) = \frac{1}{\sqrt{npq}}\varphi\left(\frac{k - np}{\sqrt{npq}}\right)$$

$$P(k_1 < X < k_2) = \phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \phi\left(\frac{k_1 - np}{\sqrt{npq}}\right)$$

Thông thường: $X \sim B(n,p)$ có $n \geq 30$, p gần 0,5; $np \geq 5$ và $npq \geq 5$ thì ta xấp xỉ $X \sim N(np, npq)$

VI. CÁC ĐỊNH LÝ

X_1, X_2 là 2 đại lượng ngẫu nhiên độc lập

1) $X_1 \sim B(n_1, p)$ và $X_2 \sim B(n_2, p) \Rightarrow X_1 + X_2 \sim B(n_1 + n_2, p)$

2) $X_1 \sim P(\lambda_1)$ và $X_2 \sim P(\lambda_2) \Rightarrow X_1 + X_2 \sim P(\lambda_1 + \lambda_2)$

3) $X_1 \sim N(\mu_1\sigma_1^2,)$ và $X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

4) $X_1 \sim \chi^2(n_1)$ và $X_2 \sim \chi^2(n_2) \Rightarrow X_1 + X_2 \sim \chi^2 (n_1 + n_2)$

χ^2 : phân phối chi (khi) bình phương

5) $X_1 \sim N(0, 1)$ và $X_2 \sim N(0, 1) \Rightarrow X_1^2 + X_2^2 \sim \chi^2(2)$

TÓM TẮT CÔNG THỨC THỐNG KÊ

I. ƯỚC LƯỢNG THAM SỐ

A. Ước lượng trung bình μ với độ tin cậy γ

1) $n \geq 30$

   1.1 Biết σ :

$$\bar{x} - t_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

   1.2 Không biết σ :

$$\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}$$

2) $n < 30$

   2.1 Biết σ :

$$\bar{x} - t_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

   2.2 Không biết σ :

$$\bar{x} - \left[t_{\alpha/2}(n-1)\right]\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + \left[t_{\alpha/2}(n-1)\right]\frac{s}{\sqrt{n}}$$

B. Ước lượng tỷ lệ *p* với độ tin cậy γ

$$f - t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}} \leq p \leq f + t_{\alpha/2}\sqrt{\frac{f(1-f)}{n}}$$

C. Ước lượng phương sai σ^2 với độ tin cậy γ

1) Biết kỳ vọng μ:

$$\frac{\sum n_i(x_i-\mu)^2}{\chi^2_{\alpha/2}(n)} < \sigma^2 < \frac{\sum n_i(x_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}$$

2) Không biết kỳ vọng μ:

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}$$

II. KIỂM ĐỊNH THAM SỐ (KIỂM ĐỊNH 2 PHÍA)

A. Kiểm định trung bình, với mức ý nghĩa α

Đặt giả thiết:

$$\begin{cases} H: \mu = \mu_0 \\ \overline{H}: \mu \neq \mu_0 \end{cases}$$

1) $n \geq 30$, σ^2 đã biết (hoặc $n < 30$, σ đã biết, X có phân phối chuẩn)

$$|t| = \frac{|\overline{x}-\mu_0|\sqrt{n}}{\sigma}$$

Kết luận:

$|t| \leq t_{\alpha/2}$: chấp nhận H

$|t| > t_{\alpha/2}$: bác bỏ H

2) $n \geq 30$, σ^2 chưa biết

$$|t| = \frac{|\overline{x}-\mu_0|\sqrt{n}}{s}$$

Kết luận:

$|t| \leq t_{\alpha/2}$: chấp nhận H

$|t| > t_{\alpha/2}$: bác bỏ H

3) $n < 30$, σ chưa biết, X có phân phối chuẩn

$$|t| = \frac{|\overline{x} - \mu_0|\sqrt{n}}{s}$$

Kết luận:

$|t| \leq t_{\alpha/2}^{(n-1)}$: chấp nhận H

$|t| > t_{\alpha/2}^{(n-1)}$: bác bỏ H

B. Kiểm định tỷ lệ, với mức ý nghĩa α

Đặt giả thiết:

$$\begin{cases} H: p = p_0 \\ \overline{H}: p \neq p_0 \end{cases}$$

$$|t| = \frac{|f - p_0|\sqrt{n}}{\sqrt{p_0(1 - p_0)}}$$

Kết luận:

$|t| \leq t_{\alpha/2}$: chấp nhận H

$|t| > t_{\alpha/2}$: bác bỏ H

C. Kiểm định phương sai, với mức ý nghĩa α

Đặt giả thiết:

$$\begin{cases} H: \sigma^2 = \sigma_0^2 \\ \overline{H}: \sigma^2 \neq \sigma_0^2 \end{cases}$$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Kết luận:

$\chi^2_{1-\alpha/2}(n-1) \leq \chi^2 \leq \chi^2_{\alpha/2}(n-1)$: chấp nhận H

$\chi^2 < \chi^2_{1-\alpha/2}(n-1)$ hoặc $\chi^2 > \chi^2_{\alpha/2}(n-1)$: bác bỏ H