

PHẦN I: XÁC SUẤT

1. Biến cố ngẫu nhiên & xác suất của biến cố:

1.1. Công thức cộng xác suất:

1.1.1. $p(A+B)=p(A)+p(B)$ (2 biến cố xung khắc)

1.1.2. $p(A+B)=p(A)+p(B)-p(A.B) \rightarrow p(A+B+C)=p(A)+p(B)+p(C)-[p(AB)+p(AC)+p(BC)]+p(ABC)$

1.2. Công thức nhân xác suất:

1.2.1. $p(A.B)=p(A).p(B)$ (2 biến cố độc lập)

1.2.2. $p(A.B)=p(A).p(B/A) \rightarrow p(A_1A_2...A_n) = p(A_1).p(A_2 / A_1)...p(A_n / A_1A_2..A_{n-1})$

1.3. Công thức Bernoulli: cho 2 biến cố A và $\bar{A}$

1.3.1. $p_n(x) = C_n^x p^x q^{n-x}$, $p=p(A)$, $q=1-p$

1.4. Công thức xác suất đầy đủ:

$$p(F) = p(A_1).p(F / A_1) + p(A_2).p(F / A_2) + \dots + p(A_n).p(F / A_n)$$

1.5. Công thức Bayes: $p(A_i / F) = \frac{p(A_i.F)}{p(F)} = \frac{p(A_i).p(F / A_i)}{p(F)}$

2. Biến ngẫu nhiên:

2.1. Bảng phân phối xác suất (biến ngẫu nhiên rời rạc)

2.2. Hàm mật độ xác suất ($f(x)$) (biến ngẫu nhiên liên tục)

2.2.1. $f(x) \geq 0$

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

2.2.3. $p(a \leq x \leq b) = \int_a^b f(x) dx$

2.3. Hàm phân phối xác suất ($F(x)$) (dùng cho cả 2 loại biến-thường là biến ngẫu nhiên liên tục)

2.3.1. $F(x) = p(F \leq x)$

2.3.2. $F'(x) = f(x)$

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

2.4. Kỳ vọng

2.4.1. $E(x) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n$ (từ bảng phân phối xác suất)

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

2.5. Phương sai:

2.5.1. $V(x) = E(x^2) - [E(x)]^2$

$$2.5.2. V(x) = \int_{-\infty}^{+\infty} x^2 f(x) dx - \left[\int_{-\infty}^{+\infty} x f(x) dx \right]^2$$

3. Một số phân phối xác suất thông dụng:

3.1. Phân phối chuẩn tổng quát: $X \sim N(\mu; \sigma^2)$

$$3.1.1. f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$3.1.2. \int_{-\infty}^{+\infty} f(x) dx = 1$$

$$3.1.3. \text{Mod} X = \text{Med} X = \mu; E(x) = \mu, V(x) = \sigma^2$$

$$3.1.4. p(a \leq x \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

3.1.5. Phân phối chuẩn tắc $\mu = 0, \sigma^2 = 1$

$$3.1.5.1. T \sim N(0,1)$$

$$3.1.5.2. f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

$$3.1.5.3. \text{Đổi biến } T = \frac{X - \mu}{\sigma}$$

$$3.1.5.4. p(a \leq x \leq b) = \Phi(b) - \Phi(a)$$

3.2. Phân phối Poisson: $X \sim P(\lambda), \lambda > 0$

$$3.2.1. p(\lambda = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

$$3.2.2. E(x) = V(x) = \lambda$$

3.3. Phân phối nhị thức: $X \sim B(n, p)$

$$3.3.1. p(X = k) = p_n(k) = C_n^k p^k q^{n-k}, p + q = 1$$

$$3.3.2. \sum_{k=0}^n p(X = k) = 1$$

$$3.3.3. E(x) = np, \text{Mod} X = x_0, np - q \leq x_0 \leq np + q$$

3.3.4. Khi $n=1$: $X \sim B(1, p)$: phân phối không-một

$$3.3.4.1. E(x) = p, E(x^2) = p, V(x) = pq$$

3.3.5. Xấp xỉ phân phối nhị thức:

3.3.5.1. Bằng phân phối Poisson: $n > 50, p < 0.1; X \sim B(n, p) \approx X \sim P(\lambda), \lambda = np$.

$$p(x = k) = C_n^k p^k q^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

3.3.5.2. Bảng phân phối chuẩn:

$$np \geq 0.5, nq \geq 0.5, \mu = np, \sigma = \sqrt{npq} \cdot X \sim B(n, p) \approx X \sim N(np, npq)$$

$$p(X = k) = \frac{1}{\sigma} f\left(\frac{k - \mu}{\sigma}\right); p(k_1 < X < k_2) = \varphi\left(\frac{k_2 - \mu}{\sigma}\right) - \varphi\left(\frac{k_1 - \mu}{\sigma}\right)$$

3.4. Phân phối siêu bội: $X \sim H(N, N_A, n)$ [N: tổng số phần tử, N_A : Số phần tử có tính chất A trong N, n: số phần tử lấy ngẫu nhiên]. Gọi X là số phần tử có tính chất A trong n.

$$p(X = k) = \frac{C_{N_A}^k \cdot C_{N - N_A}^{n - k}}{C_N^n}$$

$$3.4.1. E(X) = np, p = \frac{N_A}{N}; V(X) = npq \cdot \frac{N - n}{N - 1}, q = 1 - p$$

3.4.2. Xấp xỉ phân phối siêu bội bằng phân phối nhị thức:

$$n \leq 0.05N \Rightarrow X \sim B(n, p); p(X = k) = C_n^k p^k q^{n - k}, p = \frac{N_A}{N}$$

3.5. Biến ngẫu nhiên 2 chiều: X và Y độc lập $\Leftrightarrow P_{ij} = p(x_i) \cdot q(y_j)$ với mọi i, j

3.6. Hiệp phương sai và hệ số tương quan:

3.6.1. Hiệp phương sai(cov): $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$

3.6.2. Hệ số tương quan $\rho_{X,Y}$: $\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)}$

PHẦN 2: THỐNG KÊ

1. Tổng thể và mẫu

1.1. Thực hành tính toán trên mẫu:

1.1.1. Tính trung bình ($\bar{X}_n$): $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n x_i$

1.1.2. Tính tỷ lệ mẫu: (f_n); $f_n = \frac{m_A}{n}$ (m_A : số phần tử mang tính chất A; n: kích thước mẫu)

1.1.3. Tính phương sai mẫu: $s^2 = \frac{1}{n - 1} \left[\sum_{i=1}^n n_i x_i^2 - n(\bar{X})^2 \right]$

1.2. Ước lượng tham số của tổng thể:

1.2.1. Ước lượng điểm: $E(X_n) = \mu, E(f_n) = p, E(s^2) = \sigma^2$

1.2.2. Ước lượng khoảng:

1.2.2.1. Ước lượng khoảng cho trung bình: Với độ tin cậy $1 - \alpha$ cho trước, 1 mẫu kích thước n.

$n \geq 30, \sigma^2$ biết	$n \geq 30, \sigma^2$ chưa biết
$\bar{X}, \sigma$ $\mu_1 = \bar{X} - \varepsilon, \mu_2 = \bar{X} + \varepsilon$	$\bar{X}, s$ $\mu_1 = \bar{X} - \varepsilon, \mu_2 = \bar{X} + \varepsilon$

$\varepsilon = u_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ $(1 - \alpha \rightarrow 0.5 - \frac{\alpha}{2} \rightarrow u_{\frac{\alpha}{2}})$	$\varepsilon = u_{\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$ $(1 - \alpha \rightarrow 0.5 - \frac{\alpha}{2} \rightarrow u_{\frac{\alpha}{2}})$
$n < 30, \sigma^2$ biết	$n < 30, \sigma^2$ chưa biết
Như TH1	$\bar{X}, s$ $\mu_1 = \bar{X} - \varepsilon, \mu_2 = \bar{X} + \varepsilon$ $\varepsilon = t_{(n-1, \frac{\alpha}{2})} \cdot \frac{s}{\sqrt{n}}$

1.2.2.2. Ước lượng khoảng cho tỷ lệ: tổng thể có tỷ lệ p chưa biết, với độ tin cậy $1 - \alpha$ cho trước, với 1 mẫu kích thước n , tỷ lệ mẫu f_n . Tìm 2 số p_1, p_2 thỏa:

$$p(p_1 \leq p \leq p_2) = 1 - \alpha, p_{1,2} = f_n \mp \varepsilon \quad \text{Công thức: } \varepsilon = u_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}}$$

1.2.2.3. Ước lượng khoảng cho phương sai: Giả sử tổng thể có σ^2 chưa biết. Dựa vào 1 mẫu kích thước n , với độ tin cậy $1 - \alpha$ cho trước.

TH1: μ chưa biết, biết s^2 . Khi đó ta có $\sigma^2 \in [\frac{(n-1)s^2}{\chi_1^2}, \frac{(n-1)s^2}{\chi_2^2}]$ trong đó

$$\chi_1^2 = \chi^2(n-1, \frac{\alpha}{2}), \chi_2^2 = \chi^2(n-1, 1 - \frac{\alpha}{2})$$

TH2: μ biết. Khi đó $\sigma^2 \in [\frac{\sum n_i(x_i - \mu)^2}{\chi_1^2}, \frac{\sum n_i(x_i - \mu)^2}{\chi_2^2}]$, trong đó

$$\chi_1^2 = \chi^2(n, \frac{\alpha}{2}), \chi_2^2 = \chi^2(n, 1 - \frac{\alpha}{2})$$

1.2.3. Kiểm định giả thuyết thống kê:

1.2.3.1. Kiểm định giả thuyết thống kê cho μ

1.2.3.1.1. TH1: σ^2 biết

Giả thuyết thống kê	$W_\alpha : \sigma^2$ biết (miền bác bỏ H_0)
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$W_\alpha = \{u = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}, u > u_{\frac{\alpha}{2}}\}$
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$W_\alpha = \{u = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}, u < -u_\alpha\}$
$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$	$W_\alpha = \{u = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}, u > u_\alpha\}$

1.2.3.1.2. TH2: $n \geq 30, \sigma^2$ không biết

Giả thuyết thống kê	w_α (miền bác bỏ H_0)
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$w_\alpha = \left\{ u = \frac{\bar{X} - \mu_0}{s} \sqrt{n}, u > u_{\frac{\alpha}{2}} \right\}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$w_\alpha = \left\{ u = \frac{\bar{X} - \mu_0}{s} \sqrt{n}, u < -u_\alpha \right\}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$w_\alpha = \left\{ u = \frac{\bar{X} - \mu_0}{s} \sqrt{n}, u > u_\alpha \right\}$

1.2.3.1.3. TH3: $n < 30, \sigma^2$ không biết

Giả thuyết thống kê	w_α (miền bác bỏ H_0)
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$w_\alpha = \left\{ t = \frac{\bar{X} - \mu_0}{s} \sqrt{n}, t > t_{(n-1, \frac{\alpha}{2})} \right\}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$w_\alpha = \left\{ t = \frac{\bar{X} - \mu_0}{s} \sqrt{n}, t < -t_{(n-1, \frac{\alpha}{2})} \right\}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$w_\alpha = \left\{ t = \frac{\bar{X} - \mu_0}{s} \sqrt{n}, t > t_{(n-1, \frac{\alpha}{2})} \right\}$

cuuduongthancong.com

1.2.3.2. Kiểm định giả thuyết thống kê cho tỷ lệ:

Giả thuyết thống kê	w_α (miền bác bỏ H_0)
$H_0: p = p_0$ $H_1: p \neq p_0$	$w_\alpha = \left\{ u = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, u > u_{\frac{\alpha}{2}} \right\}$
$H_0: p = p_0$ $H_1: p < p_0$	$w_\alpha = \left\{ u = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, u < -u_\alpha \right\}$
$H_0: p = p_0$ $H_1: p > p_0$	$w_\alpha = \left\{ u = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, u > u_\alpha \right\}$

cuuduongthancong.com

1.2.3.3. Kiểm định giả thuyết thống kê cho phương sai:

1.2.3.3.1. TH1: μ chưa biết

Giả thuyết thống kê	w_α (miền bác bỏ H_0)
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$w_\alpha = \left\{ \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \chi^2 < \chi_{(n-1, \frac{\alpha}{2})}^2 \text{ hoặc } \chi^2 > \chi_{(n-1, \frac{\alpha}{2})}^2 \right\}$

$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$	$W_\alpha = \{ \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \chi^2 < \chi_{(n-1, 1-\alpha)}^2 \}$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$	$W_\alpha = \{ \chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \chi^2 > \chi_{(n-1, \alpha)}^2 \}$

1.2.3.3.2.TH2: μ biết.

Giả thuyết thống kê	W_α (miền bác bỏ H_0)
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$W_\alpha = \{ \chi^2 = \frac{\sum n_i (x_i - \mu)^2}{\sigma_0^2}, \chi^2 < \chi_1^2 \text{ hoặc } \chi^2 > \chi_2^2 \}$ $\chi_1^2 = \chi_{(n, 1-\frac{\alpha}{2})}^2, \chi_2^2 = \chi_{(n, \frac{\alpha}{2})}^2$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$	$W_\alpha = \{ \chi^2 = \frac{\sum n_i (x_i - \mu)^2}{\sigma_0^2}, \chi^2 < \chi_{(n, 1-\alpha)}^2 \}$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$	$W_\alpha = \{ \chi^2 = \frac{\sum n_i (x_i - \mu)^2}{\sigma_0^2}, \chi^2 > \chi_{(n, \alpha)}^2 \}$

1.2.4. So sánh 2 tham số của tổng thể:

1.2.4.1. So sánh 2 số trung bình:

1.2.4.1.1.TH1: $m \geq 30, n \geq 30, \sigma_1^2, \sigma_2^2$ biết

GTTK	W_α
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}; u > u_{\frac{\alpha}{2}} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}; u < -u_\alpha \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}; u > u_\alpha \right\}$

1.2.4.1.2.TH2: $m < 30, n < 30, \sigma_1^2, \sigma_2^2$ biết, X,Y có phân phối chuẩn

GTTK	W_α
------	------------

$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}; u > u_{\frac{\alpha}{2}} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}; u < -u_\alpha \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}; u > u_\alpha \right\}$

1.2.4.1.3. TH3: $m \geq 30, n \geq 30, \sigma_1^2, \sigma_2^2$ không biết

GTTK	W_α
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}; u > u_{\frac{\alpha}{2}} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}; u < -u_\alpha \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$W_\alpha = \left\{ u = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}; u > u_\alpha \right\}$

1.2.4.1.4. TH4: $m < 30, n < 30, X, Y$ có phân phối chuẩn, $\sigma_1^2 = \sigma_2^2$ không biết

GTTK	W_α
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$W_\alpha = \left\{ t = \frac{\bar{X} - \bar{Y}}{\sqrt{s^2 \left(\frac{1}{m} + \frac{1}{n} \right)}}; t > t_{\left(m+n-2, \frac{\alpha}{2} \right)} \right\} \quad s^2 = \frac{m-1 s_1^2 + n-1 s_2^2}{m+n-2}$

$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$W_\alpha = \left\{ t = \frac{\bar{X} - \bar{Y}}{\sqrt{s^2 \left(\frac{1}{m} + \frac{1}{n} \right)}}; t < -t_{m+n-2, \alpha} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$W_\alpha = \left\{ t = \frac{\bar{X} - \bar{Y}}{\sqrt{s^2 \left(\frac{1}{m} + \frac{1}{n} \right)}}; t > t_{m+n-2, \alpha} \right\}$

1.2.4.1.5. TH5: $m < 30, n < 30$, X, Y có phân phối chuẩn, $\sigma_1^2 \neq \sigma_2^2$ chưa biết

GTTK	W_α
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$W_\alpha = \left\{ g = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}; g > t; t_1 = t_{m-1, \frac{\alpha}{2}}, t_2 = t_{n-1, \frac{\alpha}{2}}; v_1 = \frac{s_1^2}{m}, v_2 = \frac{s_2^2}{n}; t = \frac{t_1 v_1 + t_2 v_2}{v_1 + v_2} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$W_\alpha = \left\{ g = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}; g < -t; t_1 = t_{m-1, \alpha}, t_2 = t_{(n-1, \alpha)} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$W_\alpha = \left\{ g = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}; g > t \right\}$

1.2.4.2. So sánh 2 tỷ lệ:

GTTK	W_α
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$W_\alpha = \left\{ u = \frac{f_1 - f_2}{\sqrt{f(1-f) \left(\frac{1}{m} + \frac{1}{n} \right)}}; u > u_{\frac{\alpha}{2}}; f_1 = \frac{k_1}{m}, f_2 = \frac{k_2}{n} \right\}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$W_\alpha = \left\{ u = \frac{f_1 - f_2}{\sqrt{f(1-f) \left(\frac{1}{m} + \frac{1}{n} \right)}}; u < -u_\alpha \right\}$

$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$W_\alpha = \left\{ u = \frac{f_1 - f_2}{\sqrt{f_1 - f_2 \left(\frac{1}{m} + \frac{1}{n} \right)}}; u > u_\alpha \right\}$
--	---

1.2.4.3. So sánh 2 phương sai:

GTTK	W_α
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	$W_\alpha = \left\{ g = \frac{s_1^2}{s_2^2}, g < \bar{f} \text{ hay } g > f; f = f_{\alpha/2}(m-1, n-1), \bar{f} = \frac{1}{f_{\alpha/2}(n-1, m-1)} \right\}$
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	$W_\alpha = \left\{ g = \frac{s_1^2}{s_2^2}, g > f_\alpha(m-1, n-1) \right\}$

cuu duong than cong . com

cuu duong than cong . com