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) → 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) → $p(A_1A_2...A_n) = p(A_1).p(A_2 / A_1)...p(A_n / A_1A_2..A_{n-1})$

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

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:

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

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

2. Biến ngẫu nhiên:

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

2.2. Hàm mật độ xác suất ($f(x)$) (biễn 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 <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. $E(x) = x_1p_1 + x_2p_2 + ... + x_np_n$ (từ bảng phân phối xác suất)

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

2.5. Phương sai:

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

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

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. $ModX=MedX=\mu; E(x)=\mu, V(x)=\sigma^2$

3.1.4. $p(a\le x\le b)=\varphi(\frac{b-\mu}{\sigma})-\varphi(\frac{a-\varphi}{\sigma})$

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. Đổi biến $T=\frac{X-\mu}{\sigma}$

3.1.5.4. $p(a\le x\le b)=\varphi(b)-\varphi(a)$

3.2.Phân phối Poisson: $X\sim P(\lambda)$, λ ***>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. $p(X=k)=p_n(k)=C_n^kp^kq^{n-k}, p+q=1$

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

3.3.3. $E(x)=np, ModX=x_0, np-q\le x_0\le np+q$

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

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

3.3.5.2. Bằng phân phối chuẩn:

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

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

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.

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

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

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

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

3.5.*Biến ngẫu nhiên 2 chiều: X* và Y độc lập $\Leftrightarrow P_{ij} = p(x_i).q(y_j)$ 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 ($\overline{X_n}$): $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} x_i$

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: $S^2 = \frac{1}{n-1}[\sum_{1}^{k} n_i x_i^2 - n(\overline{X})^2]$

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-α cho trước, 1 mẫu kích thước n.

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

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

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 thoả:

$p(p_1 \le p \le p_2) = 1-\alpha$, $p_{1,2} = f_n \mp \varepsilon$ 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-α 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 $\sigma^2 \in [\frac{\sum n_i(x_i-\mu)}{\chi_1^2}, \frac{\sum n_i(x_i-\mu)}{\chi_2^2}]$, 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)
$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$W_\alpha = \{u = \frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}, \lvert u \rvert > u_{\frac{\alpha}{2}}\}$
$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$W_\alpha = \{u = \frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}$,u<-u_α }
$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$	$W_\alpha = \{u = \frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}$,u>u_α }

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

Giả thuyết thống kê	W_α (miền bác bỏ H_0)
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$W_\alpha = \{u = \frac{\overline{X} - \mu_0}{s}\sqrt{n}, \lvert u\rvert > u_{\frac{\alpha}{2}}\}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$W_\alpha = \{u = \frac{\overline{X} - \mu_0}{s}\sqrt{n}, u < -u_\alpha\}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$W_\alpha = \{u = \frac{\overline{X} - \mu_0}{s}\sqrt{n}, u > u_\alpha\}$

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

Giả thuyết thống kê	W_α (miền bác bỏ H_0)
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$W_\alpha = \{t = \frac{\overline{X} - \mu_0}{s}\sqrt{n}, \lvert t\rvert > t_{(n-1,\frac{\alpha}{2})}\}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$W_\alpha = \{t = \frac{\overline{X} - \mu_0}{s}\sqrt{n}, t < -t_{(n-1,\frac{\alpha}{2})}\}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$W_\alpha = \{t = \frac{\overline{X} - \mu_0}{s}\sqrt{n}, t > t_{(n-1,\frac{\alpha}{2})}\}$

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

Giả thuyết thống kê	W_α (miền bác bỏ H_0)
$H_{0:} p = p_0$ $H_{1:} p \neq p_0$	$W_\alpha = \{u = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, \lvert u\rvert > u_{\frac{\alpha}{2}}\}$
$H_{0:} p = p_0$ $H_{1:} p < p_0$	$W_\alpha = \{u = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, u < -u_\alpha\}$
$H_{0:} p = p_0$ $H_{1:} p > p_0$	$W_\alpha = \{u = \frac{f - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, u > u_\alpha\}$

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)
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$W_\alpha = \{\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}, \chi^2 < \chi_1^2$ hoặc $\chi^2 > \chi_2^2$ $\chi_1^2 = \chi^2_{(n-1,1-\frac{\alpha}{2})}, \chi_2^2 = \chi^2_{(n-1,\frac{\alpha}{2})}$

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

1.2.3.3.2.TH2: μ biết.

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

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

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

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

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

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

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

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

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

1.2.4.2. So sánh 2 tỷ lệ:

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

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

1.2.4.3. So sánh 2 phương sai:

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

Tóm tắt công thức Xác Suất - Thống Kê

I. Phần Xác Suất

1. Xác suất cổ điển
   - Công thức cộng xác suất: P(A+B)=P(A)+P(B)-P(AB).
   - $A_1, A_2, \dots, A_n$ xung khắc từng đôi $\Leftrightarrow P(A_1+A_2+\dots+A_n)=P(A_1)+P(A_2)+\dots+P(A_n)$.
   - Ta có
     - A, B xung khắc $\Leftrightarrow$ P(A+B)=P(A)+P(B).
     - A, B, C xung khắc từng đôi $\Leftrightarrow$ P(A+B+C)=P(A)+P(B)+P(C).
     - $P(\overline{A}) = 1 - P(A)$.
   - Công thức xác suất có điều kiện: $P(A/B) = \dfrac{P(AB)}{P(B)}$, $P(B/A) = \dfrac{P(AB)}{P(A)}$.
   - Công thức nhân xác suất: P(AB)=P(A).P(B/A)=P(B).P(A/B).
   - $A_1, A_2, \dots, A_n$ độc lập với nhau $\Leftrightarrow P(A_1.A_2.\dots.A_n)=P(A_1).P(A_2).\dots.P(A_n)$.
   - Ta có
     - A, B độc lập $\Leftrightarrow$ P(AB)=P(A).P(B).
     - A, B, C độc lập với nhau $\Leftrightarrow$ P(A.B.C)=P(A).P(B).P(C).
   - Công thức Bernoulli: $B(k;n;p) = C_n^k p^k q^{n-k}$, với p=P(A): xác suất để biến cố A xảy ra ở mỗi phép thử và q=1-p.
   - Công thức xác suất đầy đủ - Công thức Bayes
     - Hệ biến cố gồm n phần tử $A_1, A_2, \dots, A_n$ được gọi là một phép phân hoạch của $\Omega \Leftrightarrow \begin{cases} A_i.A_j = \Phi, \forall i \neq j; i, j \in \overline{1,n} \\ A_1 + A_2 + \dots + A_n = \Omega \end{cases}$
     - Công thức xác suất đầy đủ:

$$P(B) = \sum_{i=1}^{n} P(A_i).P(B/A_i) = P(A_1).P(B/A_1) + P(A_2).P(B/A_2) + \dots + P(A_n).P(B/A_n)$$

     - Công thức Bayes:

$$P(A_i/B) = \frac{P(A_i).P(B/A_i)}{P(B)}$$

với $P(B) = P(A_1).P(B/A_1) + P(A_2).P(B/A_2) + \dots + P(A_n).P(B/A_n)$

2. Biến ngẫu nhiên

   a. Biến ngẫu nhiên rời rạc
   - Luật phân phối xác suất

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

với $p_i = P(X = x_i), i = \overline{1,n}$.

Ta có:

$$\sum_{i=1}^{n} p_i = 1 \text{ và } P\{a \le f(X) \le b\} = \sum_{a \le f(x_i) \le b} p_i$$

- Hàm phân phối xác suất
  $F_X(x) = P(X < x) = \sum_{x_i < x} p_i$
- Mode
  $\text{ModX} = x_0 \Leftrightarrow p_0 = \max\{p_i : i = \overline{1,n}\}$
- Median
  $$\text{MedX} = x_e \Leftrightarrow \begin{cases} P(X < x_e) \le 0,5 \\ P(X > x_e) \le 0,5 \end{cases} \Leftrightarrow \begin{cases} \sum_{x_i < x_e} p_i \le 0,5 \\ \sum_{x_i > x_e} p_i \le 0,5 \end{cases}$$
- Kỳ vọng
  $$EX = \sum_{i=1}^{n} (x_i.p_i) = x_1.p_1 + x_2.p_2 + ... + x_n.p_n$$
  $$E(\varphi(X)) = \sum_{i=1}^{n} (\varphi(x_i).p_i) = \varphi(x_1).p_1 + \varphi(x_2).p_2 + ... + \varphi(x_n).p_n$$
- Phương sai
  $VarX = E(X^2) - (EX)^2$
  với $E(X^2) = \sum_{i=1}^{n} (x_i^2.p_i) = x_1^2.p_1 + x_2^2.p_2 + ... + x_n^2.p_n$

b. Biến ngẫu nhiên liên tục.

- f(x) là hàm mật độ xác suất của X $\Rightarrow \int_{-\infty}^{+\infty} f(x)dx = 1$,
  $$P\{\text{a} \le \text{X} \le \text{b}\} = \int_a^b f(x).dx$$
- Hàm phân phối xác suất
  $$F_X(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt$$
- Mode
  $ModX = x_0 \Leftrightarrow$ Hàm mật độ xác suất f(x) của X đạt cực đại tại x_0.
- Median
  $$MedX = x_e \Leftrightarrow F_X(x_e) = \frac{1}{2} \Leftrightarrow \int_{-\infty}^{x_e} f(x)dx = \frac{1}{2}.$$
- Kỳ vọng
  $$\text{EX} = \int_{-\infty}^{+\infty} x.f(x)dx.$$
  $$E(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(x).f(x)dx$$

- Phương sai

$$VarX = E(X^2) - (EX)^2 \text{ với } EX^2 = \int_{-\infty}^{+\infty} x^2.f(x)dx.$$

c. Tính chất

- $E(C) = C, Var(C) = 0$, C là một hằng số.
- $E(kX) = kEX, Var(kX) = k^2VarX$
- $E(aX + bY) = aEX + bEY$
- Nếu X, Y độc lập thì $E(XY) = EX.EY, Var(aX + bY) = a^2VarX + b^2VarY$
- $\sigma(X) = \sqrt{VarX}$: Độ lệch chuẩn của X, có cùng thứ nguyên với X và EX.

3. Luật phân phối xác suất

a. Phân phối Chuẩn $(X \sim N(\mu;\sigma^2))$

- $X(\Omega) = \mathbb{R}$, EX=ModX=MedX=μ, $VarX = \sigma^2$
- Hàm mđxs $f(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \Rightarrow$ Với $\mu = 0, \sigma = 1$:

$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (Hàm Gauss)

- $P(a \le X \le b) = \varphi(\frac{b-\mu}{\sigma}) - \varphi(\frac{a-\mu}{\sigma})$ với $\varphi(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ (Hàm Laplace)
- Cách sử dụng máy tính bỏ túi để tính giá trị hàm Laplace, hàm phân phối xác suất của phân phối chuẩn chuẩn tắc

Tác vụ	**Máy CASIO 570MS**	**Máy CASIO 570ES**
Khởi động gói Thống kê	Mode…(tìm)…SD	Mode…(tìm)…STAT 1-Var
Tính $\varphi(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$	Shift 3 2 x) =	Shift 1 7 2 x) =
$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$	Shift 3 1 x) =	Shift 1 7 1 x) =
Thoát khỏi gói Thống kê	Mode 1	Mode 1

Lưu ý: $F(x) = 0,5 + \varphi(x)$

b. Phân phối Poisson $(X \sim P(\lambda))$

- $X(\Omega) = \mathbb{N}$, $EX = VarX = \lambda$. ModX=k $\Leftrightarrow \lambda\text{-}1 \le k \le \lambda$
- $P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$, $k \in \mathbb{N}$

c. Phân phối Nhị thức $(X \sim B(n;p))$

- $X(\Omega)=\{0..n\}$, EX=np, VarX=npq, ModX=k $\Leftrightarrow (n+1)p-1 \le k \le (n+1)p$
- $P(\text{X=k})=\text{C}_n^k.p^k.q^{n-k}$, $\text{q}=1-\text{p}, 0 \le k \le n,\ k \in \mathbb{N}$
- Nếu $(n \ge 30; 0{,}1 < p < 0{,}9; np \ge 5, nq \ge 5)$ thì $X \sim B(n;p) \approx N(\mu;\sigma^2)$ với $\mu = n.p, \sigma = \sqrt{npq}$
  - $P(\text{X=k}) \approx \dfrac{1}{\sigma} f(\dfrac{k-\mu}{\sigma}),\ 0 \le k \le n,\ k \in \mathbb{N}$
  - $P(\text{a} \le \text{X<b}) \approx \varphi(\dfrac{b-\mu}{\sigma}) - \varphi(\dfrac{a-\mu}{\sigma})$
- Nếu $(n \ge 30,\ p \le 0{,}1, np < 5)$ thì $X \sim B(n;p) \approx P(\lambda)$ với $\lambda = np$
  - $P(\text{X=k}) \approx \text{e}^{-\lambda} \dfrac{\lambda^k}{k!},\ k \in \mathbb{N}$
- Nếu $(n \ge 30,\ p \ge 0{,}9, nq < 5)$

  $P(\text{X=k}) \approx \text{e}^{-\lambda} \dfrac{\lambda^{n-k}}{(n-k)!}, k \in \mathbb{R}$ với $\lambda = nq$

d. Phân phối Siêu bội $(X \sim H(N;N_A;n))$

- $X(\Omega)=\{\max\{0;n-(N-N_A)\}..\min\{\text{n};\text{N}_A\}\}$
- EX=np, VarX=npq$\dfrac{N-n}{N-1}$ với $p=\dfrac{N_A}{N}$, q=1-p.
- $ModX = k \Leftrightarrow \dfrac{(N_A+1)(n+1)+2}{N+2} - 1 \le k \le \dfrac{(N_A+1)(n+1)+2}{N+2}$.
- $P(\text{X=k})=\dfrac{C_{N_A}^k C_{N-N_A}^{n-k}}{C_N^n},\ k \in X(\Omega)$
- Nếu $\dfrac{N}{n} > 20$ thì $X \sim H(N;N_A;n) \approx B(n;p)$ với $p=\dfrac{N_A}{N}$.

  $P(\text{X=k}) \approx \text{C}_n^k.p^k.q^{n-k},\ k \in X(\Omega), q=1-p$.

Sơ đồ tóm tắt các dạng phân phối xác suất thông dụng:

Siêu bội: $X \sim H(N;N_A;n)$

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

↓ *N>20n*, $p=\frac{N_A}{N}$, *q=1-p*

Nhị thức: $X \sim B(n;p)$

$$P(X=k)=C_n^k.p^k.q^{n-k}$$

→ *n≥30, np<5*, *p≤0,1*, $\lambda=np$

Poisson: $X \sim P(\lambda)$

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

↓ *n≥30*, $np \geq 5$, $nq \geq 5$

0,1<p<0,9

$$P(X=k)\approx\frac{1}{\sigma}f\left(\frac{k-\mu}{\sigma}\right)$$

$$P(a\leq X<b)\approx\varphi\left(\frac{b-\mu}{\sigma}\right)-\varphi\left(\frac{a-\mu}{\sigma}\right)$$

với $\mu=np,\ \sigma=\sqrt{npq}$

Chuẩn: $X \sim N(\mu;\sigma^2)$

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

→ $Y=\frac{X-\mu}{\sigma}$

Chuẩn chuẩn tắc: $Y \sim N(0;1)$

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

II. Phần Thống Kê.

1. Lý thuyết mẫu.

a. Các công thức cơ bản.

Các giá trị đặc trưng	**Mẫu ngẫu nhiên**	**Mẫu cụ thể**
Giá trị trung bình	$\overline{X}=\dfrac{X_1+...+X_n}{n}$	$\overline{x}=\dfrac{x_1+...+x_n}{n}$
Phương sai không hiệu chỉnh	$\hat{S}_X^2=\dfrac{(X_1-\overline{X})^2+...+(X_n-\overline{X})^2}{n}$	$\hat{s}_x^2=\dfrac{(x_1-\overline{x})^2+...+(x_n-\overline{x})^2}{n}$
Phương sai hiệu chỉnh	$S_X^2=\dfrac{(X_1-\overline{X})^2+...+(X_n-\overline{X})^2}{n-1}$	$s_x^2=\dfrac{(x_1-\overline{x})^2+...+(x_n-\overline{x})^2}{n-1}$

b. Để dễ xử lý ta viết số liệu của mẫu cụ thể dưới dạng tần số như sau:

x_i	x_1	x_2	$\cdots$	x_k
n_i	n_1	n_2	$\cdots$	n_k

Khi đó

Các giá trị đặc trưng	**Mẫu cụ thể**
Giá trị trung bình	$\overline{x}=\dfrac{x_1n_1+...+x_kn_k}{n}$
Phương sai không hiệu chỉnh	$\hat{s}_x^2=\dfrac{(x_1-\overline{x})^2n_1+...+(x_k-\overline{x})^2n_k}{n}$
Phương sai hiệu chỉnh	$s_x^2=\dfrac{(x_1-\overline{x})^2n_1+...+(x_k-\overline{x})^2n_k}{n-1}$

c. Cách sử dụng máy tính bỏ túi để tính các giá trị đặc trưng mẫu

- Nếu số liệu thống kê thu thập theo miền $[a;b)$ hay $(a;b]$ thì ta sử dụng giá trị đại diện cho miền đó là $\dfrac{a+b}{2}$ để tính toán.

Tác vụ	**Dòng CASIO MS**	**Dòng CASIO ES**
Bật chế độ nhập tần số	Không cần	Shift Mode ↓ 4 1
Khởi động gói Thống kê	Mode…(tìm)…SD	Mode…(tìm)…STAT 1-Var
Nhập số liệu	x_1 Shift , n_1 M+ $\vdots$ x_k Shift , n_k M+ Nếu $n_i=1$ thì chỉ cần nhấn x_i M+	<table><tr><th>X</th><th>FREQ</th></tr><tr><td>x_1 = $\vdots$ x_k =</td><td>n_1 = $\vdots$ n_k =</td></tr></table>

Xóa màn hình hiển thị	AC	AC
Xác định: • Kích thước mẫu (n) • Giá trị trung bình ($\overline{x}$) • Độ lệch chuẩn không hiệu chỉnh ($\hat{s}_x$) • Độ lệch chuẩn hiệu chỉnh (s_x)	 Shift 1 3 = Shift 2 1 = Shift 2 2 = Shift 2 3 =	 Shift 1 5 1 = Shift 1 5 2 = Shift 1 5 3 = Shift 1 5 4 =
Thoát khỏi gói Thống kê	Mode 1	Mode 1

2. Ước lượng khoảng.

a) Khoảng tin cậy cho giá trị trung bình.

Trường hợp 1. (σ đã biết)

- Ước lượng đối xứng.

$$\varphi(z_{\frac{\alpha}{2}}) = \frac{1-\alpha}{2} \rightarrow z_{\frac{\alpha}{2}} \Rightarrow \varepsilon = z_{\frac{\alpha}{2}}.\frac{\sigma}{\sqrt{n}} \Rightarrow (\overline{x}-\varepsilon\,;\overline{x}+\varepsilon)$$

- Ước lượng chệch trái.

$$\varphi(z_\alpha) = 0,5-\alpha \rightarrow z_\alpha \Rightarrow \varepsilon = z_\alpha.\frac{\sigma}{\sqrt{n}} \Rightarrow (-\infty;\overline{x}+\varepsilon)$$

- Ước lượng chệch phải.

$$\varphi(z_\alpha) = 0,5-\alpha \rightarrow z_\alpha \Rightarrow \varepsilon = z_\alpha.\frac{\sigma}{\sqrt{n}} \Rightarrow (\overline{x}-\varepsilon;+\infty)$$

Trường hợp 2. (σ chưa biết, $n \geq 30$)

- Ước lượng đối xứng.

$$\varphi(z_{\frac{\alpha}{2}}) = \frac{1-\alpha}{2} \rightarrow z_{\frac{\alpha}{2}} \Rightarrow \varepsilon = z_{\frac{\alpha}{2}}.\frac{s}{\sqrt{n}} \Rightarrow (\overline{x}-\varepsilon\,;\overline{x}+\varepsilon)$$

- Ước lượng chệch trái.

$$\varphi(z_\alpha) = 0,5-\alpha \rightarrow z_\alpha \Rightarrow \varepsilon = z_\alpha.\frac{s}{\sqrt{n}} \Rightarrow (-\infty;\overline{x}+\varepsilon)$$

- Ước lượng chệch phải.

$$\varphi(z_\alpha) = 0,5-\alpha \rightarrow z_\alpha \Rightarrow \varepsilon = z_\alpha.\frac{s}{\sqrt{n}} \Rightarrow (\overline{x}-\varepsilon;+\infty)$$

Trường hợp 3. (σ chưa biết, n<30)

- Ước lượng đối xứng.

$$1-\alpha \rightarrow \frac{\alpha}{2} \rightarrow t_{(n-1;\frac{\alpha}{2})} \Rightarrow \varepsilon = t_{(n-1;\frac{\alpha}{2})}.\frac{s}{\sqrt{n}} \Rightarrow (\overline{x}-\varepsilon\,;\overline{x}+\varepsilon)$$

- Ước lượng chệch trái.

$$1-\alpha \rightarrow \alpha \rightarrow t_{(n-1;\alpha)} \Rightarrow \varepsilon = t_{(n-1;\alpha)}.\frac{s}{\sqrt{n}} \Rightarrow (-\infty\,;\overline{x}+\varepsilon)$$

- Ước lượng chệch phải.

$$1-\alpha \to \alpha \to t_{(n-1;\alpha)} \Rightarrow \varepsilon = t_{(n-1;\alpha)}.\frac{s}{\sqrt{n}} \Rightarrow (\overline{x}-\varepsilon\,;+\infty)$$

b) Khoảng tin cậy cho tỉ lệ.

- Ước lượng đối xứng.

$$\varphi(z_{\frac{\alpha}{2}}) = \frac{1-\alpha}{2} \to z_{\frac{\alpha}{2}} \Rightarrow \varepsilon = z_{\frac{\alpha}{2}}.\frac{\sqrt{f(1-f)}}{\sqrt{n}} \Rightarrow (f-\varepsilon\,;\,f+\varepsilon)$$

- Ước lượng chệch trái.

$$\varphi(z_{\alpha}) = 0,5-\alpha \to z_{\alpha} \Rightarrow \varepsilon = z_{\alpha}.\frac{\sqrt{f(1-f)}}{\sqrt{n}} \Rightarrow (0;\,f+\varepsilon)$$

- Ước lượng chệch phải.

$$\varphi(z_{\alpha}) = 0,5-\alpha \to z_{\alpha} \Rightarrow \varepsilon = z_{\alpha}.\frac{\sqrt{f(1-f)}}{\sqrt{n}} \Rightarrow (\,f-\varepsilon;1)$$

c) Khoảng tin cậy cho phương sai.

<u>Trường hợp 1</u>. (μ chưa biết)

- Nếu đề bài chưa cho s mà cho mẫu cụ thể thì phải xác định s (bằng máy tính).

- Ước lượng không chệch.

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

$$\Rightarrow \left(\frac{(n-1)s^2}{\chi_2};\frac{(n-1)s^2}{\chi_1}\right)$$

- Ước lượng chệch trái.

$$1-\alpha \to \chi_1 = \chi^2_{(n-1;1-\alpha)} \Rightarrow \left(0;\frac{(n-1)s^2}{\chi_1}\right)$$

- Ước lượng chệch phải.

$$1-\alpha \to \alpha \to \chi_2 = \chi^2_{(n-1;\alpha)} \Rightarrow \left(\frac{(n-1)s^2}{\chi_2};+\infty\right)$$

<u>Trường hợp 2</u>. (μ đã biết)

- Tính $(n-1)s^2 = \sum_{i=1}^{k} n_i.(x_i-\mu)^2$

- Ước lượng không chệch.

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

$$\Rightarrow \left(\frac{(n-1)s^2}{\chi_2};\frac{(n-1)s^2}{\chi_1}\right)$$

- Ước lượng chệch trái.

  $1-\alpha \to \chi_1 = \chi^2_{(n;1-\alpha)} \Rightarrow (0;\dfrac{(n-1)s^2}{\chi_1})$

- Ước lượng chệch phải.

  $1-\alpha \to \alpha \to \chi_2 = \chi^2_{(n;\alpha)} \Rightarrow (\dfrac{(n-1)s^2}{\chi_2};+\infty)$

3. Kiểm định tham số.
   a) Kiểm định giá trị trung bình.

Trường hợp 1. (σ đã biết)

- $H_o : \mu = \mu_o, H_1 : \mu \neq \mu_o$

  $\varphi(z_{\frac{\alpha}{2}}) = \dfrac{1-\alpha}{2} \to z_{\frac{\alpha}{2}}, z = \dfrac{\overline{x}-\mu_o}{\sigma}.\sqrt{n}$

  - Nếu $|z| > z_{\frac{\alpha}{2}}$: Bác bỏ H_o, chấp nhận H_1.

  - Nếu $|z| \leq z_{\frac{\alpha}{2}}$: Chấp nhận H_o.

- $H_o : \mu = \mu_o, H_1 : \mu < \mu_o$

  $\varphi(z_{\alpha}) = 0,5-\alpha \to z_{\alpha}, z = \dfrac{\overline{x}-\mu_o}{\sigma}.\sqrt{n}$

  - Nếu $z < -z_{\alpha}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \geq -z_{\alpha}$: Chấp nhận H_o.

- $H_o : \mu = \mu_o, H_1 : \mu > \mu_o$

  $\varphi(z_{\alpha}) = 0,5-\alpha \to z_{\alpha}, z = \dfrac{\overline{x}-\mu_o}{\sigma}.\sqrt{n}$

  - Nếu $z > z_{\alpha}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \leq z_{\alpha}$: Chấp nhận H_o.

Trường hợp 2. (σ chưa biết, $n \geq 30$)

- $H_o : \mu = \mu_o, H_1 : \mu \neq \mu_o$

  $\varphi(z_{\frac{\alpha}{2}}) = \dfrac{1-\alpha}{2} \to z_{\frac{\alpha}{2}}, z = \dfrac{\overline{x}-\mu_o}{s}.\sqrt{n}$

  - Nếu $|z| > z_{\frac{\alpha}{2}}$: Bác bỏ H_o, chấp nhận H_1.

  - Nếu $|z| \leq z_{\frac{\alpha}{2}}$: Chấp nhận H_o.

- $H_o : \mu = \mu_o, H_1 : \mu < \mu_o$

  $\varphi(z_{\alpha}) = 0,5-\alpha \to z_{\alpha}, z = \dfrac{\overline{x}-\mu_o}{s}.\sqrt{n}$

- Nếu $z < -z_{\alpha}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $z \geq -z_{\alpha}$: Chấp nhận H_o.

- $H_o : \mu = \mu_o, H_1 : \mu > \mu_o$

  $\varphi(z_{\alpha}) = 0,5 - \alpha \to z_{\alpha}, z = \dfrac{\overline{x} - \mu_o}{s}.\sqrt{n}$

  - Nếu $z > z_{\alpha}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \leq z_{\alpha}$: Chấp nhận H_o.

<u>Trường hợp 3</u>. (σ chưa biết, n<30)

- $H_o : \mu = \mu_o, H_1 : \mu \neq \mu_o$

  $\alpha \to \dfrac{\alpha}{2} \to t_{(n-1;\frac{\alpha}{2})}, t = \dfrac{\overline{x} - \mu_o}{s}.\sqrt{n}$

  - Nếu $|t| > t_{(n-1;\frac{\alpha}{2})}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $|t| \leq t_{(n-1;\frac{\alpha}{2})}$: Chấp nhận H_o.

- $H_o : \mu = \mu_o, H_1 : \mu < \mu_o$

  $\alpha \to t_{(n-1;\alpha)}, t = \dfrac{\overline{x} - \mu_o}{s}.\sqrt{n}$

  - Nếu $t < -t_{(n-1;\alpha)}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $t \geq -t_{(n-1;\alpha)}$: Chấp nhận H_o.

- $H_o : \mu = \mu_o, H_1 : \mu > \mu_o$

  $\alpha \to t_{(n-1;\alpha)}, t = \dfrac{\overline{x} - \mu_o}{s}.\sqrt{n}$

  - Nếu $t > t_{(n-1;\alpha)}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $t \leq t_{(n-1;\alpha)}$: Chấp nhận H_o.

b) Kiểm định tỉ lệ.

- $H_o : p = p_o, H_1 : p \neq p_o$

  $\varphi(z_{\frac{\alpha}{2}}) = \dfrac{1-\alpha}{2} \to z_{\frac{\alpha}{2}}, f = \dfrac{k}{n}, z = \dfrac{f - p_o}{\sqrt{p_o(1-p_o)}}.\sqrt{n}$

  - Nếu $|z| > z_{\frac{\alpha}{2}}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $|z| \leq z_{\frac{\alpha}{2}}$: Chấp nhận H_o.

- $H_o : p = p_o, H_1 : p < p_o$

  $\varphi(z_{\alpha}) = 0,5 - \alpha \to z_{\alpha}, f = \dfrac{k}{n}, z = \dfrac{f - p_o}{\sqrt{p_o(1-p_o)}}.\sqrt{n}$

- Nếu $z < -z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $z \ge -z_\alpha$: Chấp nhận H_o.

- $H_o : p = p_o, H_1 : p > p_o$

  $\varphi(z_\alpha) = 0,5 - \alpha \rightarrow z_\alpha,\ f = \dfrac{k}{n}, z = \dfrac{f - p_o}{\sqrt{p_o(1-p_o)}}.\sqrt{n}$

  - Nếu $z > z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \le z_\alpha$: Chấp nhận H_o.

c) Kiểm định phương sai.

Trường hợp 1. (μ chưa biết)

- Nếu đề chưa cho s mà cho mẫu cụ thể thì phải sử dụng máy tính để xác định s.

- $H_o : \sigma^2 = \sigma_o^2, H_1 : \sigma^2 \ne \sigma_o^2$

  $\alpha \rightarrow 1 - \dfrac{\alpha}{2} \rightarrow \chi_1^2 = \chi^2_{(n-1;1-\frac{\alpha}{2})},\ \alpha \rightarrow \dfrac{\alpha}{2} \rightarrow \chi_2^2 = \chi^2_{(n-1;\frac{\alpha}{2})},\ \chi^2 = \dfrac{(n-1)s^2}{\sigma_o^2}$

  - Nếu $\left[\begin{array}{l}\chi^2 > \chi_2^2 \\ \chi^2 < \chi_1^2\end{array}\right.$: Bác bỏ H_0, chấp nhận H_1.
  - Nếu $\chi_1^2 \le \chi^2 \le \chi_2^2$: Chấp nhận H_o.

- $H_o : \sigma^2 = \sigma_o^2, H_1 : \sigma^2 < \sigma_o^2$

  $\alpha \rightarrow 1 - \alpha \rightarrow \chi_1^2 = \chi^2_{(n-1;1-\alpha)},\ \chi^2 = \dfrac{(n-1)s^2}{\sigma_o^2}$

  - Nếu $\chi^2 < \chi_1^2$: Bác bỏ H_0, chấp nhận H_1.
  - Nếu $\chi^2 \ge \chi_1^2$: Chấp nhận H_o.

- $H_o : \sigma^2 = \sigma_o^2, H_1 : \sigma^2 > \sigma_o^2$

  $\alpha \rightarrow \chi_2^2 = \chi^2_{(n-1;\alpha)},\ \chi^2 = \dfrac{(n-1)s^2}{\sigma_o^2}$

  - Nếu $\chi^2 > \chi_2^2$: Bác bỏ H_0, chấp nhận H_1.
  - Nếu $\chi^2 \le \chi_2^2$: Chấp nhận H_o.

4. Kiểm định so sánh tham số.

a) Kiểm định so sánh giá trị trung bình.

Trường hợp 1. (σ_1, σ_2 đã biết)

- $H_o : \mu_1 = \mu_2, H_1 : \mu_1 \ne \mu_2$

  $\varphi(z_{\frac{\alpha}{2}}) = \dfrac{1-\alpha}{2} \rightarrow z_{\frac{\alpha}{2}}, z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

- Nếu $z > z_{\frac{\alpha}{2}}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $|z| \le z_{\frac{\alpha}{2}}$: Chấp nhận H_o.

* $H_o : \mu_1 = \mu_2, H_1 : \mu_1 < \mu_2$

  $\varphi(z_\alpha) = 0,5 - \alpha \to z_\alpha,\ z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

  - Nếu $z < -z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \ge -z_\alpha$: Chấp nhận H_o.

* $H_o : \mu_1 = \mu_2, H_1 : \mu_1 > \mu_2$

  $\varphi(z_\alpha) = 0,5 - \alpha \to z_\alpha,\ z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

  - Nếu $z > z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \le z_\alpha$: Chấp nhận H_o.

<u>Trường hợp 2</u>. (σ_1, σ_2 chưa biết, $n_1, n_2 \ge 30$)

* $H_o : \mu_1 = \mu_2, H_1 : \mu_1 \ne \mu_2$

  $\varphi(z_{\frac{\alpha}{2}}) = \dfrac{1-\alpha}{2} \to z_{\frac{\alpha}{2}},\ z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

  - Nếu $|z| > z_{\frac{\alpha}{2}}$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $|z| \le z_{\frac{\alpha}{2}}$: Chấp nhận H_o.

* $H_o : \mu_1 = \mu_2, H_1 : \mu_1 < \mu_2$

  $\varphi(z_\alpha) = 0,5 - \alpha \to z_\alpha,\ z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

  - Nếu $z < -z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \ge -z_\alpha$: Chấp nhận H_o.

* $H_o : \mu_1 = \mu_2, H_1 : \mu_1 > \mu_2$

  $\varphi(z_\alpha) = 0,5 - \alpha \to z_\alpha,\ z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

- Nếu $z > z_{\alpha}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $z \le z_{\alpha}$: Chấp nhận H_o.

<u>Trường hợp 3</u>. ($\sigma_1 = \sigma_2$ chưa biết, $n_1, n_2 < 30$)

- $H_o : \mu_1 = \mu_2, H_1 : \mu_1 \ne \mu_2$

$$\alpha \to \frac{\alpha}{2} \to t_{(n_1+n_2-2;\frac{\alpha}{2})}, t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}, \text{ với } s^2 = \frac{(n_1-1).s_1^2 + (n_2-1).s_2^2}{n_1+n_2-2}$$

- Nếu $|t| > t_{(n_1+n_2-2;\frac{\alpha}{2})}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $|t| \le t_{(n_1+n_2-2;\frac{\alpha}{2})}$: Chấp nhận H_o.

- $H_o : \mu_1 = \mu_2, H_1 : \mu_1 < \mu_2$

$$\alpha \to t_{(n_1+n_2-2;\alpha)}, t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}, \text{ với } s^2 = \frac{(n_1-1).s_1^2 + (n_2-1).s_2^2}{n_1+n_2-2}$$

- Nếu $t < -t_{(n_1+n_2-2;\frac{\alpha}{2})}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $t \ge -t_{(n_1+n_2-2;\frac{\alpha}{2})}$: Chấp nhận H_o.

- $H_o : \mu_1 = \mu_2, H_1 : \mu_1 > \mu_2$

$$\alpha \to t_{(n_1+n_2-2;\alpha)}, t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}, \text{ với } s^2 = \frac{(n_1-1).s_1^2 + (n_2-1).s_2^2}{n_1+n_2-2}$$

- Nếu $t > t_{(n_1+n_2-2;\frac{\alpha}{2})}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $t \le t_{(n_1+n_2-2;\frac{\alpha}{2})}$: Chấp nhận H_o.

b) Kiểm định so sánh tỉ lệ.

$$f_1 = \frac{k_1}{n_1}, \ f_2 = \frac{k_2}{n_2}, \ f = \frac{k_1+k_2}{n_1+n_2}$$

- $H_o : p_1 = p_2, H_1 : p_1 \ne p_2$

$$\varphi(z_{\frac{\alpha}{2}}) = \frac{1-\alpha}{2} \to z_{\frac{\alpha}{2}}, \ z = \frac{f_1 - f_2}{\sqrt{f(1-f).(\frac{1}{n_1} + \frac{1}{n_2})}}$$

- Nếu $|z| > z_{\frac{\alpha}{2}}$: Bác bỏ H_o, chấp nhận H_1.
- Nếu $|z| \le z_{\frac{\alpha}{2}}$: Chấp nhận H_o.

- $H_o: p_1 = p_2, H_1: p_1 < p_2$

  $\varphi(z_\alpha) = 0,5 - \alpha \to z_\alpha, z = \dfrac{f_1 - f_2}{\sqrt{f(1-f).(\frac{1}{n_1} + \frac{1}{n_2})}}$

  - Nếu $z < -z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \geq -z_\alpha$: Chấp nhận H_o.

- $H_o: p_1 = p_2, H_1: p_1 > p_2$

  $\varphi(z_\alpha) = 0,5 - \alpha \to z_\alpha, z = \dfrac{f_1 - f_2}{\sqrt{f(1-f).(\frac{1}{n_1} + \frac{1}{n_2})}}$

  - Nếu $z > z_\alpha$: Bác bỏ H_o, chấp nhận H_1.
  - Nếu $z \leq z_\alpha$: Chấp nhận H_o.

c. Kiểm định so sánh phương sai.

- μ_1, μ_2 chưa biết nên tính s_1 và s_2 từ mẫu (sử dụng máy tính) nếu đề bài chưa cho.

  - $H_o: \sigma_1{}^2 = \sigma_2^2, H_1: \sigma_1{}^2 \neq \sigma_2^2$
    - $f = \dfrac{s_1^2}{s_2^2}$, $f_1 = f(n_1 - 1; n_2 - 1; 1 - \dfrac{\alpha}{2})$, $f_2 = f(n_1 - 1; n_2 - 1; \dfrac{\alpha}{2})$
    - Nếu $\left[\begin{array}{l} f < f_1 \\ f > f_2 \end{array}\right.$: Bác bỏ H_o, chấp nhận H_1.
    - Nếu $f_1 \leq f \leq f_2$: Chấp nhận H_o.
  - $H_o: \sigma_1{}^2 = \sigma_2^2, H_1: \sigma_1{}^2 < \sigma_2^2$
    - $f = \dfrac{s_1^2}{s_2^2}$, $f_1 = f(n_1 - 1; n_2 - 1; 1 - \alpha)$
    - Nếu $f < f_1$: Bác bỏ H_o, chấp nhận H_1.
    - Nếu $f_1 \leq f$: Chấp nhận H_o.
  - $H_o: \sigma_1{}^2 = \sigma_2^2, H_1: \sigma_1{}^2 > \sigma_2^2$
    - $f = \dfrac{s_1^2}{s_2^2}$, $f_2 = f(n_1 - 1; n_2 - 1; \alpha)$
    - Nếu $f > f_2$: Bác bỏ H_o, chấp nhận H_1.
    - Nếu $f \leq f_2$: Chấp nhận H_o.

5. Hệ số tương quan mẫu và phương trình hồi quy tuyến tính mẫu.

a. Hệ số tương quan mẫu: $r = \dfrac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2}}$

Phương trình hồi quy tuyến tính mẫu: $\overline{y_x} = A + Bx$ với

$$B = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \text{ và } A = \frac{\sum_{i=1}^{n} y_i - B.\sum_{i=1}^{n} x_i}{n}.$$

b. Trong trường hợp sử dụng bảng tần số:

x_i	x_1	x_2	$\cdots$	x_k
y_i	y_1	y_2	$\cdots$	y_k
n_i	n_1	n_2	$\cdots$	n_k

Ta tính theo công thức thu gọn như sau:

Hệ số tương quan mẫu: $r = \dfrac{n\sum_{i=1}^{k} n_i x_i y_i - \sum_{i=1}^{k} n_i x_i \sum_{i=1}^{k} n_i y_i}{\sqrt{n\sum_{i=1}^{k} n_i x_i^2 - (\sum_{i=1}^{k} n_i x_i)^2}\sqrt{n\sum_{i=1}^{k} n_i y_i^2 - (\sum_{i=1}^{k} n_i y_i)^2}}$

Phương trình hồi quy tuyến tính mẫu: $\overline{y_x} = A + Bx$ với

$$B = \frac{n\sum_{i=1}^{k} n_i x_i y_i - \sum_{i=1}^{k} n_i x_i \sum_{i=1}^{k} n_i y_i}{n\sum_{i=1}^{k} n_i x_i^2 - (\sum_{i=1}^{k} n_i x_i)^2} \text{ và } A = \frac{\sum_{i=1}^{k} n_i y_i - B.\sum_{i=1}^{k} n_i x_i}{n}.$$

c. Sử dụng máy tính bỏ túi để tính hệ số tương quan mẫu và phương trình hồi quy tuyến tính mẫu:

<table>
<tr><th>Tác vụ</th><th>Dòng CASIO MS</th><th>Dòng CASIO ES</th></tr>
<tr><td>Bật chế độ nhập tần số</td><td>Không cần</td><td>Shift Mode ↓ 4 1</td></tr>
<tr><td>Khởi động gói Hồi quy tuyến tính</td><td>Mode…(tìm)…REG Lin</td><td>Mode…(tìm)…STAT A+BX</td></tr>
<tr><td>Nhập số liệu</td><td>x_1 , y_1 Shift , n_1 M+
$\vdots$
x_k , y_k Shift , n_k M+

$n_i = 1$ thì chỉ cần nhấn
x_i , y_i M+</td><td><table><tr><th>X</th><th>Y</th><th>FREQ</th></tr><tr><td>x_1 =
$\vdots$
x_k =</td><td>y_1 =
$\vdots$
y_k =</td><td>n_1 =
$\vdots$
n_k =</td></tr></table></td></tr>
<tr><td>Xóa màn hình hiển thị</td><td>AC</td><td>AC</td></tr>
<tr><td>Xác định:
• Hệ số tương quan mẫu (r)
• Hệ số hằng: A
• Hệ số ẩn (x): B</td><td>Shift 2 ⟶⟶ 3 =

Shift 2 ⟶⟶ 1 =
Shift 2 ⟶⟶ 2 =</td><td>Shift 1 7 3 =

Shift 1 7 1 =
Shift 1 7 2 =</td></tr>
<tr><td>Thoát khỏi gói Hồi quy</td><td>Mode 1</td><td>Mode 1</td></tr>
</table>

Lưu ý: Máy ES nếu đã kích hoạt chế độ nhập tần số ở phần Lý thuyết mẫu rồi thì không cần kích hoạt nữa.

……………………………………….