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).
   - $A_1, A_2, \ldots, A_n$ xung khắc từng đôi $\Leftrightarrow$ $P(A_1+A_2+\ldots+A_n)=P(A_1)+P(A_2)+\ldots+P(A_n)$.
   - 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).
   - $A_1, A_2, \ldots, A_n$ độc lập với nhau $\Leftrightarrow$ $P(A_1.A_2.\ldots.A_n)=P(A_1).P(A_2).\ldots.P(A_n)$.
   - 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 $A_1, A_2, \ldots, A_n$ được gọi là một phép phân hoạch của Ω $\Leftrightarrow \begin{cases} A_i.A_j = \Phi, \forall i \neq j; i, j \in \overline{1,n} \\ A_1 + A_2 + \ldots + A_n = \Omega \end{cases}$
     - Công thức xác suất đầy đủ:

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

     - Công thức Bayes:

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

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

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
  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 $\Rightarrow \int_{-\infty}^{+\infty} f(x)dx = 1$,
  $$P\{a \le X \le b\} = \int_a^b f(x).dx$$
- 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 $f(x,\mu,\sigma) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\Rightarrow$ Với $\mu = 0, \sigma = 1$:

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

- $P(a \le X \le b) = \varphi(\dfrac{b-\mu}{\sigma}) - \varphi(\dfrac{a-\mu}{\sigma})$ với $\varphi(x) = \displaystyle\int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ (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 $\Leftrightarrow \lambda$-1 $\le$ k $\le \lambda$
- $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k \in \mathbb{N}$

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

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

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

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

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

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

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

Siêu bội: X~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~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~ P(λ)

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

n≥30, np ≥ 5, nq ≥ 5

0,1<p<0,9

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

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

với $\mu=np,\ \sigma=\sqrt{npq}$

Chuẩn: X~ $N(\mu;\sigma^2)$

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

$$Y=\frac{X-\mu}{\sigma}$$

Chuẩn chuẩn tắc: Y~ 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	$\overline{X} = \dfrac{X_1 + \ldots + X_n}{n}$	$\overline{x} = \dfrac{x_1 + \ldots + x_n}{n}$
Phương sai không hiệu chỉnh	$\hat{S}_X^2 = \dfrac{(X_1 - \overline{X})^2 + \ldots + (X_n - \overline{X})^2}{n}$	$\hat{s}_x^2 = \dfrac{(x_1 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n}$
Phương sai hiệu chỉnh	$S_X^2 = \dfrac{(X_1 - \overline{X})^2 + \ldots + (X_n - \overline{X})^2}{n-1}$	$s_x^2 = \dfrac{(x_1 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n-1}$

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	$\overline{x} = \dfrac{x_1 n_1 + \ldots + x_k n_k}{n}$
Phương sai không hiệu chỉnh	$\hat{s}_x^2 = \dfrac{(x_1 - \overline{x})^2 n_1 + \ldots + (x_k - \overline{x})^2 n_k}{n}$ $= \sigma_n^2$
Phương sai hiệu chỉnh	$s_x^2 = \dfrac{(x_1 - \overline{x})^2 n_1 + \ldots + (x_k - \overline{x})^2 n_k}{n-1}$ $= \sigma_{n-1}^2$

$\boxed{\sigma_{n-1}^2 = \frac{n}{n-1}\sigma_n^2}$

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+	X | FREQ x_1 = | n_1 = $\vdots$ | $\vdots$ x_k = | n_k =

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.

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

- Ước lượng chệch trái.

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

- Ước lượng chệch phải.

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

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

- Ước lượng đối xứng.

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

- Ước lượng chệch trái.

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

- Ước lượng chệch phải.

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

Trường hợp 3. (σ chưa biết, n<30) (Bảng pp student)

- Ước lượng đối xứng.

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

- Ước lượng chệch trái.

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

- 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. (Bảng Laplace)

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

- 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.

Trường hợp 1. (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.

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

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

- 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.

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

Trường hợp 2. (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.

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

- Ước lượng chệch phải.

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

3. Kiểm định tham số.

   a) Kiểm định giá trị trung bình.

<u>Trường hợp 1</u>. (σ đã biết)

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

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

  - 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.

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

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

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

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

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

  - 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 chưa biết, $n \ge 30$)

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

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

  - 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.

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

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

- Nếu $z < -z_\alpha$: Bác bỏ H$_o$, chấp nhận H$_1$.
- Nếu $z \ge -z_\alpha$: Chấp nhận H$_o$.

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

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

  - Nếu $z > z_\alpha$: Bác bỏ H$_o$, chấp nhận H$_1$.
  - Nếu $z \le z_\alpha$: Chấp nhận H$_o$.

Trường hợp 3. (σ chưa biết, n<30) (PP Student)

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

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

  - Nếu $|t| > t_{(n-1;\frac{\alpha}{2})}$: Bác bỏ H$_o$, chấp nhận H$_1$.
  - Nếu $|t| \le t_{(n-1;\frac{\alpha}{2})}$: 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 $t \ge -t_{(n-1;\alpha)}$: 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 $t \le t_{(n-1;\alpha)}$: Chấp nhận H$_o$.

b) Kiểm định tỉ lệ.

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

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

  - Nếu $|z| > z_{\frac{\alpha}{2}}$: Bác bỏ H$_o$, chấp nhận H$_1$.
  - Nếu $|z| \le z_{\frac{\alpha}{2}}$: 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.

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

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

- 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.

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

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

- 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.

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

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

- 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.

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

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

- 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)

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

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

- 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.

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

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

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

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

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

- 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, $n_1, n_2 \ge 30$)

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

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

- 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.

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

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

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

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

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

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

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

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

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

  - 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.

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

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

  - 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.

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

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

  - 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.

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

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

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

  - 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:

Tác vụ	Dòng CASIO MS	Dòng CASIO ES
Bật chế độ nhập tần số	Không cần	Shift Mode ↓ 4 1
Khởi động gói Hồi quy tuyến tính	Mode…(tìm)…REG Lin	Mode…(tìm)…STAT A+BX
Nhập số liệu	x_1 , y_1 Shift , n_1 M+ ⋮ x_k , y_k Shift , n_k M+ $n_i = 1$ thì chỉ cần nhấn x_i , y_i M+	X / Y / FREQ x_1 = / y_1 = / n_1 = ⋮ / ⋮ / ⋮ x_k = / y_k = / n_k =
Xóa màn hình hiển thị	AC	AC
Xác định: • Hệ số tương quan mẫu (r) • Hệ số hằng: A • Hệ số ẩn (x): B	Shift 2 ⟶⟶ 3 = Shift 2 ⟶⟶ 1 = Shift 2 ⟶⟶ 2 =	Shift 1 7 3 = Shift 1 7 1 = Shift 1 7 2 =
Thoát khỏi gói Hồi quy	Mode 1	Mode 1

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.

....

Phân phối của vectơ (X,Y)

X \ Y	y_1 .. y_j ...	P_X
x_1 ⋮ x_i	p_{ij}	
P_Y		

$p_{ij} = P(X=x_i, Y=y_j)$, $\sum p_{ij} = 1$

X,Y độc lập $\Leftrightarrow P(X=x_i, Y=y_j) = P(X=x_i).P(Y=y_j), \forall i,j$

Hàm mật độ XS đồng thời của (X,Y): $\begin{cases} f(x,y) \geq 0, \forall (x,y) \in \mathbb{R}^2 \\ \iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1 \end{cases}$

. Xác suất $(X,Y) \in D \subset \mathbb{R}^2$: $P[(X,Y) \in D] = \iint_D f(x,y)\,dx\,dy$

. Mật độ XS của X: $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$

. Mật độ XS của Y: $f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx$

. Mật độ XS của X với đk $Y=y_i$: $\underbrace{f_X(x/y_i)}_{= f_{X/y_i}(x)} = \frac{f(x,y_i)}{f_Y(y_i)}$

. Mật độ XS của Y với đk $X=x_i$: $\underbrace{f_Y(y/x_i)}_{= f_{(Y/x_i)}(y)} = \frac{f(x_i,y)}{f_X(x_i)}$

• X,Y độc lập khi $f(x,y) = f_X(x).f_Y(y)$