

Nguyên Lý Thống Kê Kinh Tế

Chương 7 Phân Phối Mẫu

Chương 7 Phân phối mẫu

- Lý do tìm hiểu phân phối mẫu
Ứng dụng lý thuyết xác suất cho quá trình suy luận thống kê.
- Tham số của tổng thể và giá trị thống kê mẫu.

Tham số của tổng thể

■ Trung bình tổng thể
(Population Mean)

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

■ Tỷ lệ tổng thể
(Population Proportion)

$$P = \frac{X}{N}$$

■ Phương sai tổng thể
(Population Variance)

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

Giá trị thống kê mẫu

Trung bình mẫu (Sample Mean):

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Tỉ lệ mẫu (Sample Proportion):

$$\hat{p} = \frac{X}{n}$$

Phương sai mẫu (Sample Variance):

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Phân phối của trung bình mẫu

Thuộc tính của trung bình mẫu

- Không chệch (Unbiasedness)
- Hiệu quả (Efficiency)
- Chắc chắn (Consistency)
- Đầy đủ (Sufficiency)

Không chệch:

$$E(\bar{x}) = \mu_{\bar{x}} = \mu$$

Mức lương ngày của các nhân viên

Tên nhân viên Mức lương ngày (1000 đ)

A 70

B 70

C 80

D 80

E 70

G 80

H 90

Trung bình tổng thể: $\mu = \frac{\sum_{i=1}^7 x_i}{7} = 77,1429$

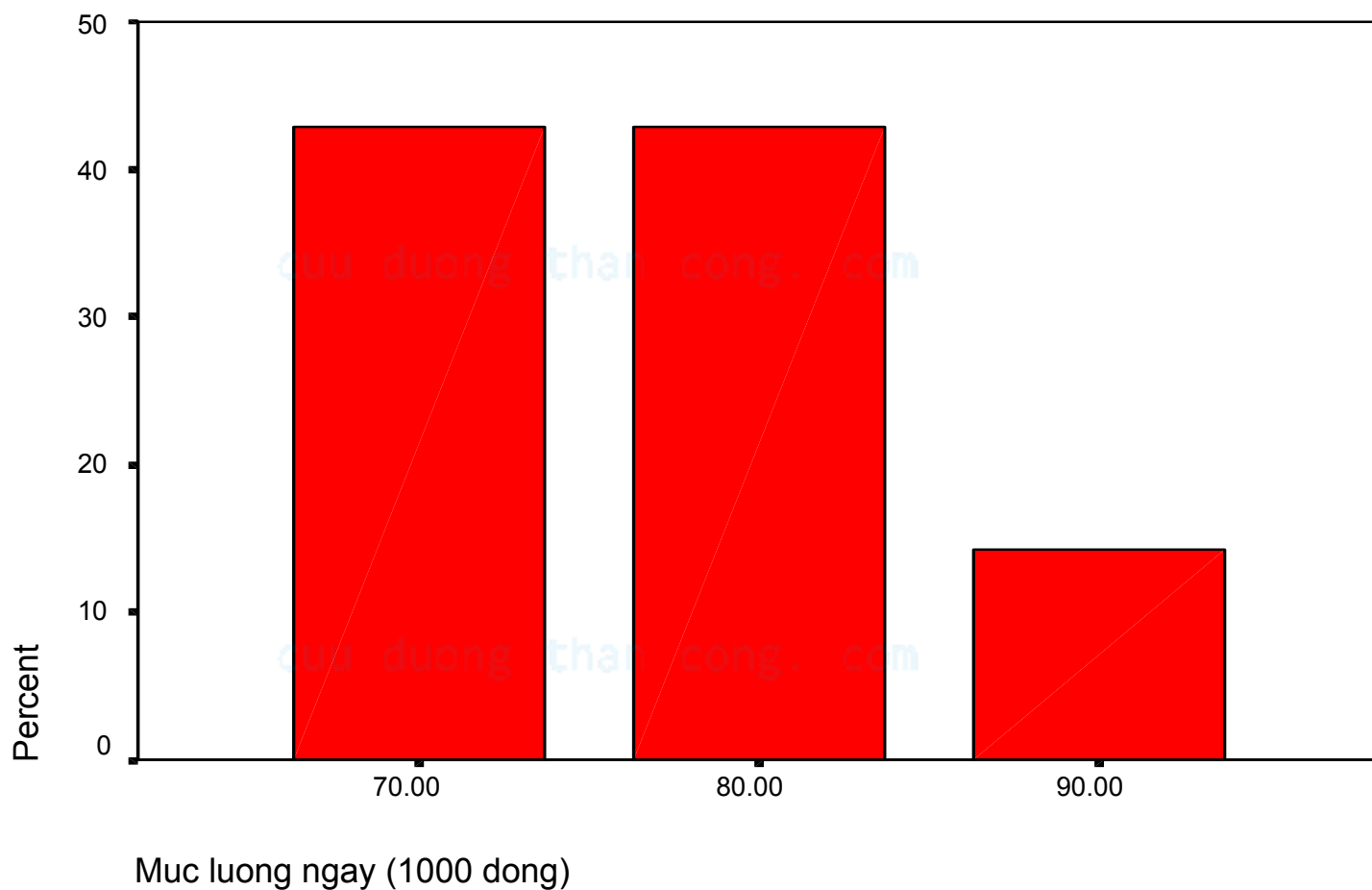
Các mẫu có thể thành lập với kích thước $n=2$

Mẫu	x_i	\bar{x}	Mẫu	x_i	\bar{x}
1 A,B	70,70	70	12 C,D	80,80	80
2 A,C	70,80	75	13 C,E	80,70	75
3 A,D	70,80	75	14 C,G	80,80	80
4 A,E	70,70	70	15 C,H	80,90	85
5 A,G	70,80	75	16 D,E	80,70	75
6 A,H	70,90	80	17 D,G	80,80	80
7 B,C	70,80	75	18 D,H	80,90	85
8 B,D	70,80	75	19 E,G	70,80	75
9 B,E	70,70	70	20 E,H	70,90	80
10 B,G	70,80	75	21 G,H	80,90	85
11 B,H	70,90	80			

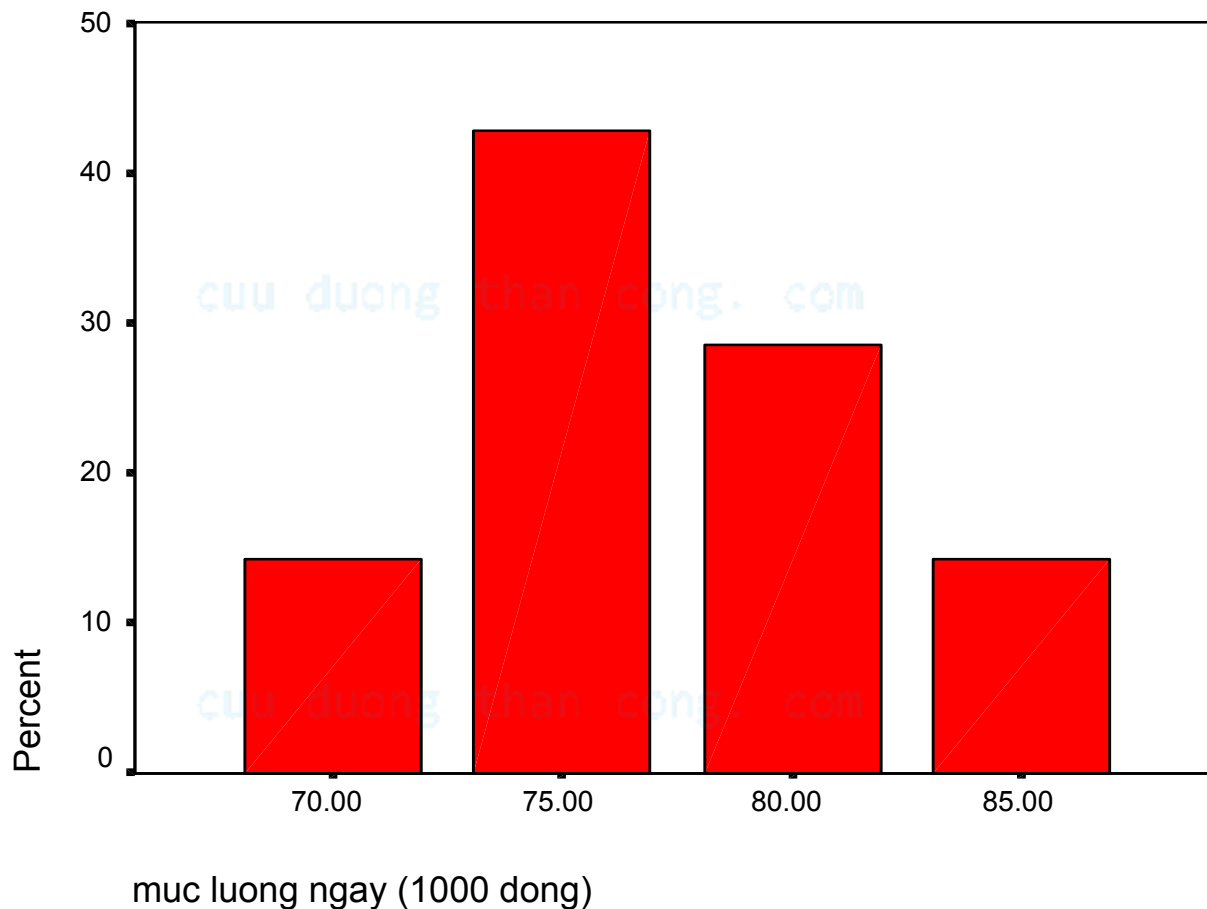
So sánh phân phối của tổng thể với phân phối của trung bình mẫu

<i>Phân phối của tổng thể</i>				<i>Phân phối mẫu</i>		
X	Số n.viên	Xác suất	\bar{x}	Số mẫu	Xác suất	
70	3	0,4286	70	3	0,1429	
80	3	0.4286	75	9	0.4285	
90	1	0,1428	80	6	0,2857	
			85	3	0,1429	
Cộng	7	1,000	Cộng	21	1,0000	

Biểu đồ phân phối xác suất của tổng thể



Biểu đồ phân phối xác suất của trung bình mẫu



Sai số chuẩn của trung bình mẫu

- Khi mẫu chọn theo phương thức có trả lại, hoặc từ tổng thể vô hạn.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- Mẫu chọn từ tổng thể hữu hạn theo phương thức không trả lại.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}}$$

Sai số chuẩn của trung bình mẫu

- Khi mẫu chọn theo phương thức có trả lại, hoặc từ tổng thể vô hạn.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$


- Mẫu chọn từ tổng thể hữu hạn theo phương thức không trả lại.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}}$$

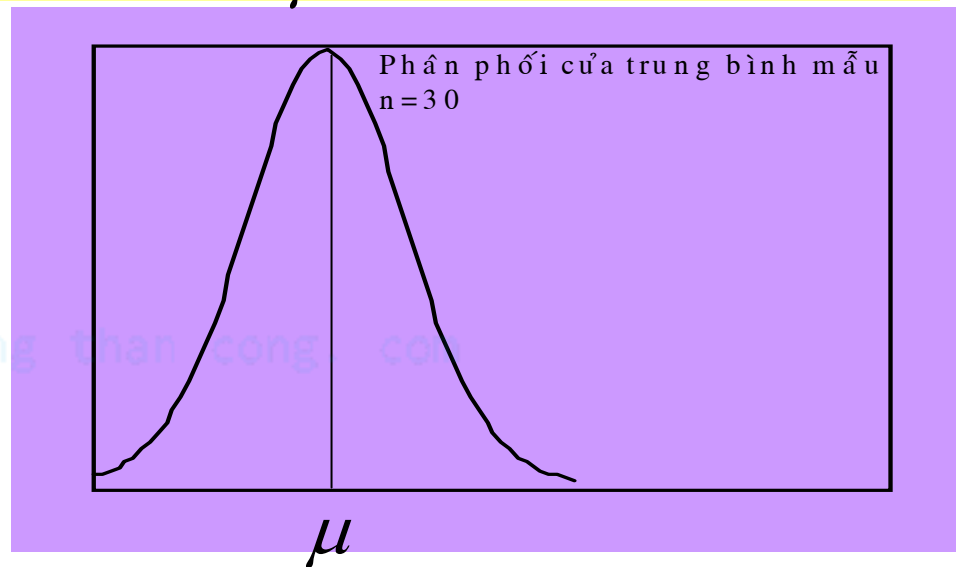
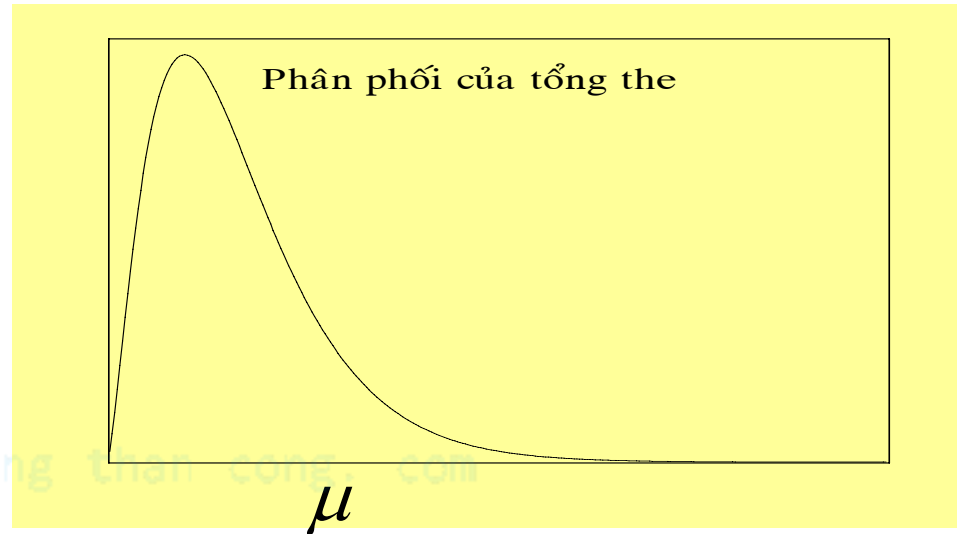
Mẫu chọn từ tổng thể có phân phối chuẩn

- $X \sim N(\mu, \sigma^2) \implies \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- Mẫu chọn từ tổng thể không có phân phối chuẩn.
- Định lý giới hạn trung tâm (CLT)

Central Limit Theorem

- Một tổng thể có phân phối bất kỳ, với điều kiện tượng tự nếu tất cả các mẫu được chọn có cùng kích thước n , phân phối của trung bình mẫu sẽ có phân phối xấp xỉ chuẩn khi n lớn.
- Nếu $n \geq 30$: CLT  phân phối của trung bình mẫu xấp xỉ phân phối chuẩn.
- Nếu $n < 30$ phân phối của trung bình mẫu có phân phối chuẩn nếu tổng thể có phân phối chuẩn

■ Định lý CLT



Phân phối của tỉ lệ mẫu

1. Thuộc tính không chệch.

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} nP$$

$$E(\hat{p}) = P$$

2. Sai số chuẩn của tỉ lệ mẫu

$$\sigma_{\hat{p}} = \sqrt{\frac{P(1-P)}{n}}$$

Trường hợp chọn mẫu không lặp:

$$\sigma_{\hat{p}} = \sqrt{\frac{P(1-P)}{n}} \sqrt{\frac{N-n}{N-1}}$$

Phân phối của phương sai mẫu.

- Phương sai mẫu: $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$
- Thuộc tính không chệch: $E(s^2) = \sigma^2$
- Biến ngẫu nhiên $\frac{(n-1)s^2}{\sigma^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma^2}$
có phân phối χ^2_{n-1}

Chương 8 Ước lượng (Estimation)

Nội dung của ước lượng:

Mẫu $\xrightarrow{\text{Suy diễn}}$ Tổng thể

8.1 Ước lượng điểm

- Trung bình

Mẫu $\xrightarrow{\quad}$ Tổng thể

$\bar{x} \xrightarrow{\quad} \mu$

- Tỷ lệ

$\hat{p} \xrightarrow{\quad} P$

$s^2 \xrightarrow{\quad} \sigma^2$

- Phương sai

8.2 Ước lượng khoảng

Gọi θ : Tham số của tổng thể.

$$P(A < \theta < B) = (1 - \alpha)$$

1. Tìm CI(1- α)100% của μ .

■ $X \sim N$

✿ Biết σ^2 cuu duong than cong. com

$$CI(1 - \alpha)100\% \mu = \bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

✿ Không biết σ^2

cuu duong than cong. com

$$CI(1 - \alpha)100\% \mu = \bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}$$

Ước lượng khoảng

- Không biết phân phối của tổng thể. Mẫu lớn ($n \geq 30$).

CLT \longrightarrow $CI(1-\alpha)100\% \text{ for } \mu = \bar{x} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$

2. Tìm $CI(1-\alpha)100\%$ của tỉ lệ tổng thể.

Với mẫu lớn ($n \geq 30$).

$$CI(1-\alpha)100\%P = \hat{p} \pm Z_{\alpha/2} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

- Trường hợp N hữu hạn , mẫu chọn không lặp

→ Hệ số $fpc = \sqrt{\frac{N-n}{N-1}}$

3. Tìm CI(1- α)100% của phương sai tổng thể.

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

8.3 Xác định kích thước mẫu

1. Xác định n khi tìm CI(1- α)100% của μ .

Công thức:

$$n = \frac{Z_{\alpha/2}^2 \sigma^2}{\varepsilon^2}$$

2. Xác định n khi tìm CI (1- α)100% của P .

$$n = \frac{Z_{\alpha/2}^2 0,25}{\varepsilon^2}$$

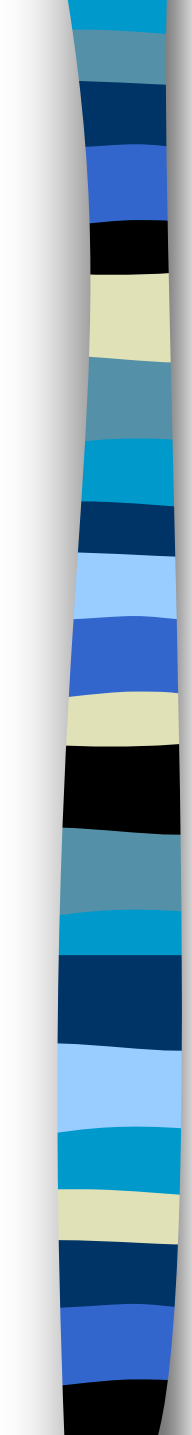
Chương 9 Kiểm định giả thuyết 1 (Hypothesis Testing)

Nội dung: Mẫu  Tổng thể

9.1 Bài toán mở đầu.

Theo thiết kế quá trình sản xuất là bình thường nếu sản phẩm có trọng lượng trung bình là 375 gam. Vào 9 giờ sáng của một ca sản xuất người ta chọn ngẫu nhiên 9 sản phẩm. Trọng lượng ghi nhận như sau:

370	374	372	378	376
371	370	377	375	



Giả sử trọng lượng sản phẩm có phân phối chuẩn với phương sai bằng 16. Với mức ý nghĩa 5%, có thể kết luận quá trình sản xuất là bình thường?

9.2 Một số khái niệm

1. Giả thuyết không và giả thuyết thay thế.

- Giả thuyết không(Null Hypothesis): H_0
- Giả thuyết thay thế (Alternative Hypothesis): H_1 .

2. Sai lầm trong kiểm định giả thuyết.

- Sai lầm loại 1: Bác bỏ H_0 đúng.
- Sai lầm loại 2: Chấp nhận H_0 sai.

Sai lầm trong kiểm định giả thuyết

Tình trạng H_0

H_0 đúng

H_0 sai

Không bác bỏ

Quyết định đúng,
xác suất: $(1 - \alpha)$

Sai lầm loại 2,
Xác-suất $= \beta$

Bác bỏ

Sai lầm loại 1,
Xác-suất $= \alpha$

Quyết định đúng,
Xác suất: $(1 - \beta)$

9.3 Kiểm định giả thuyết trung bình tổng thể

1. $X \sim N$.

■ Biết σ^2

Bước 1 Đặt giả thuyết.

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0 \text{ (Kiểm định 2 phía)}$$

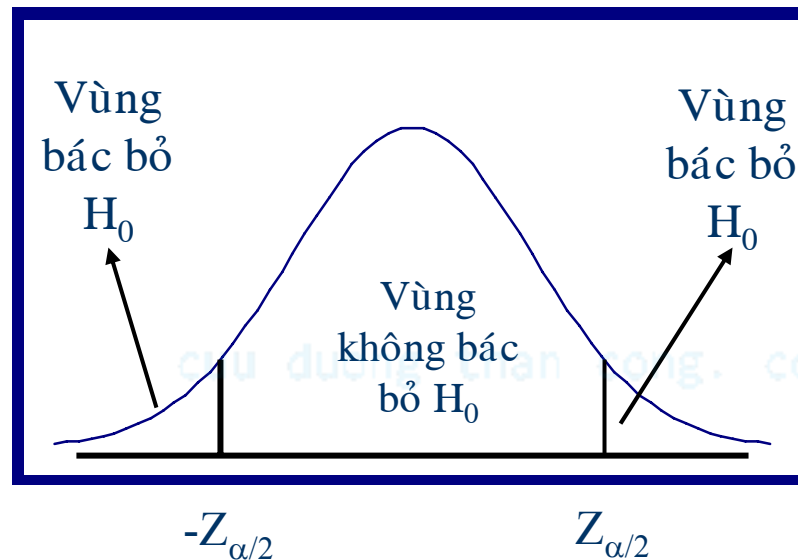
Bước 2 Chọn mức ý nghĩa α

Bước 3 Tính giá trị kiểm định

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Bước 4 Quy tắc quyết định

H_0 sẽ bị bác bỏ nếu giá trị kiểm định
 $Z > Z_{\alpha/2}$ hay $Z < -Z_{\alpha/2}$



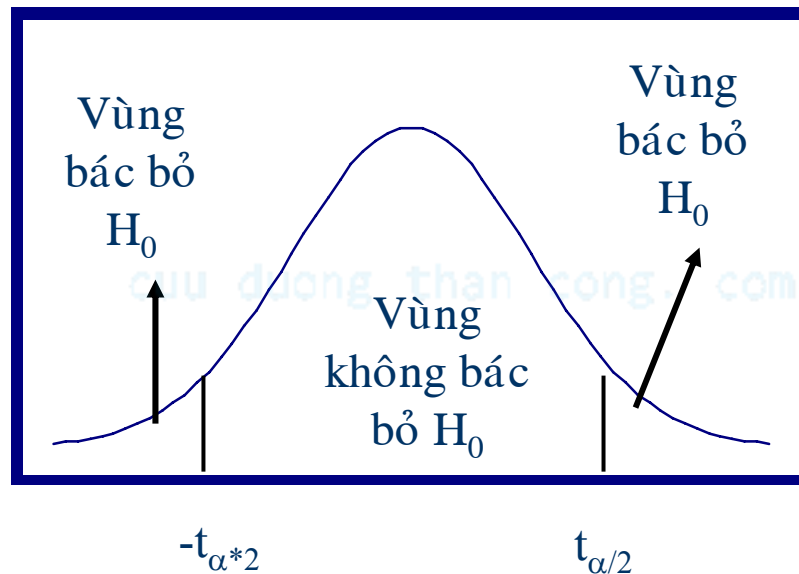
■ Không biết σ^2

Giá trị kiểm định $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

Quy tắc quyết định

H_0 sẽ bị bác bỏ nếu giá trị kiểm định $t > t_{\alpha/2}$
hay $t < -t_{\alpha/2}$

cuu duong than cong. com

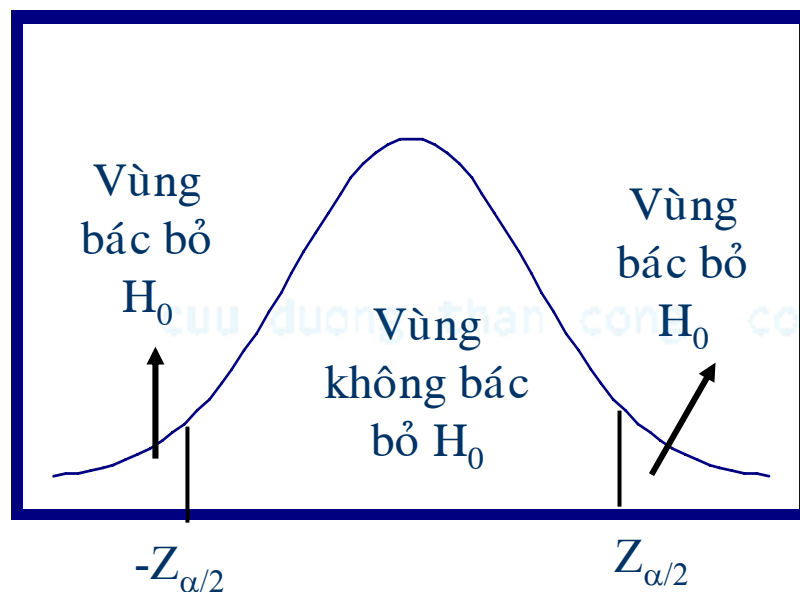


2. Tổng thể không có phân phối chuẩn, $n \geq 30$

Giá trị kiểm định:
$$Z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Quy tắc quyết định

H_0 sẽ bị bác bỏ nếu giá trị kiểm định $Z > Z_{\alpha/2}$ hay $Z < -Z_{\alpha/2}$



9.4 Phương pháp giá trị P (P-value)

- Khái niệm
- Xác định P- value:

Giá trị kiểm định



$$P - value = 2P(Z \geq |Z|)$$

- Quy tắc sử dụng giá trị P:

$P\text{-value} \leq \alpha \Rightarrow$ bác bỏ H_0

$P\text{-value} > \alpha \Rightarrow$ Không bác bỏ H_0

9.5 Kiểm định một phía (one-tail test)

- Người phân tích nhận định tham số của tổng thể theo một hướng nào đó (particular direction) Kiểm định một phía.

Kiểm định một phía, bên phải:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

Kiểm định một phía, bên trái:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0$$

9.6 Kiểm định giả thuyết tỉ lệ tổng thể (P)

Bước 1 Đặt giả thuyết

Kiểm định 2 phía:

$$H_0 : P = P_0$$

$$H_1 : P \neq P_0$$

Kiểm định 1 phía, bên phải:

$$H_0 : P = P_0$$

$$H_1 : P > P_0$$

Kiểm định 1 phía, bên trái:

$$H_0 : P = P_0$$

$$H_1 : P < P_0$$

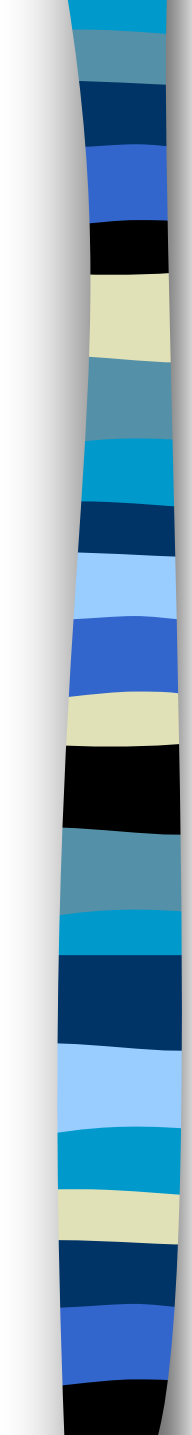
Bước 2 Chọn mức ý nghĩa α

Bước 3 Tính giá trị kiểm định:

$$Z = \frac{\hat{p} - P_0}{\sqrt{P_0(1 - P_0) / n}}$$

Bước 4 Quy tắc quyết định

- Kiểm định 2 phía: H_0 sẽ bị bác bỏ nếu giá trị kiểm định $Z > Z_{\alpha/2}$ hay $Z < -Z_{\alpha/2}$.

- 
- Kiểm định 1 phía, bên phải: H_0 sẽ bị bác bỏ nếu giá trị kiểm định $Z > Z_{\alpha}$
 - Kiểm định 1 phía, bên trái: H_0 sẽ bị bác bỏ nếu giá trị kiểm định $Z < -Z_{\alpha}$.

9.6 Kiểm định giả thuyết phương sai tổng thể.

■ Giả định: $X \sim N$

Bước 1 Đặt giả thuyết: $H_0 : \sigma^2 = \sigma_0^2$

$$H_1 : \sigma^2 \neq \sigma_0^2$$

✿ Kiểm định một phía: $H_0 : \sigma^2 = \sigma_0^2$

$$H_1 : \sigma^2 > \sigma_0^2$$

Bước 2 Chọn mức ý nghĩa α

Bước 3 Tính giá trị kiểm định:

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

Bước 4 Quy tắc quyết định

H_0 sẽ bị bác bỏ nếu:

$$\frac{(n-1)s^2}{\sigma_0^2} > \chi_{n-1, \alpha/2}^2 \quad \text{hay} \quad \frac{(n-1)s^2}{\sigma_0^2} < \chi_{n-1, 1-\alpha/2}^2$$

Chương 10 Kiểm định giả thuyết

2

10.1 Kiểm định giả thuyết trung bình của hai tổng thể.

10.1.1 Mẫu cặp (Pair samples)

Giả định: $d_i \sim N$

Bước 1 Đặt giả thuyết: $H_0 : \mu_x = \mu_y$

$H_1 : \mu_x \neq \mu_y$



cuu duong than cong. com

cuu duong than cong. com