CHỦ ĐỀ 3.2 KIỂM ĐỊNH GIẢ THUYẾT HAI MẪU

Trong chủ đề này, các em phải nắm và thực hiện được các nội dung như sau

1) Nắm được 5 bước cơ bản trong bài toán kiểm định gồm

- ☑ **B1.** Phát biểu giả thuyết kiểm định.
- ☑ **B2.** Xác định mức ý nghĩa α.
- ☑ **B3.** Tính giá trị thống kê kiểm định.
- ☑ **B4.** Xác định miền bác bỏ hoặc tính $p-$giá trị.
- ☑ **B5.** So sánh và kết luận.

2) Thực hiện được kiểm định giả thuyết so sánh hai trung bình khi phương sai đã biết.

3) Thực hiện được kiểm định giả thuyết so sánh hai trung bình khi phương sai bằng nhau chưa biết.

4) Thực hiện được kiểm định giả thuyết so sánh hai trung bình khi phương sai khác nhau chưa biết.

5) Thực hiện được kiểm định giả thuyết so sánh hai tỷ lệ.

1. KIỂM ĐỊNH SO SÁNH HAI TRUNG BÌNH Ở HAI MẪU ĐỘC LẬP

1.1. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TRUNG BÌNH KHI PHƯƠNG SAI ĐÃ BIẾT

Mẫu 1: $n_1 = ...; \bar{x}_1 = ...; \sigma_1^2 = ...$: đã biết.
Mẫu 2: $n_2 = ...; \bar{x}_2 = ...; \sigma_2^2 = ...$: đã biết.

- ☑ **B1.** Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = ...$
- ☑ **B2.** Xác định mức ý nghĩa α.
- ☑ **B3.** Tính giá trị thống kê kiểm định (TKKĐ)

$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- ☑ **B4.** Xác định miền bác bỏ hoặc tính $p-$ giá trị

Đối thuyết	**Miền bác bỏ H_0**	**$p-$giá trị**
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$\lvert z_0 \rvert > z_{1-\alpha/2}$	$p-$giá trị $= 2\left[1 - \Phi(\lvert z_0 \rvert)\right]$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$z_0 < -z_{1-\alpha}$	$p-$giá trị $= \Phi(z_0)$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$z_0 > z_{1-\alpha}$	$p-$giá trị $= 1 - \Phi(z_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0.

- ☑ **B5.** So sánh và kết luận.

1.2. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TRUNG BÌNH KHI PHƯƠNG SAI BẰNG NHAU CHƯA BIẾT

Mẫu 1: $n_1 = ...$; $\bar{x}_1 = ...$; $s_1^2 = ...$
Mẫu 2: $n_2 = ...$; $\bar{x}_2 = ...$; $s_2^2 = ...$
Phương sai không biết giá trị nhưng biết rằng $\sigma_1^2 = \sigma_2^2$.

- ☑ **B1.** Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = ...$
- ☑ **B2.** Xác định mức ý nghĩa α.
- ☑ **B3.** Tính giá trị thống kê kiểm định (TKKĐ)

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\dfrac{S_p^2}{n_1} + \dfrac{S_p^2}{n_2}}}$$

- ☑ **B4.** Xác định miền bác bỏ hoặc tính $p-$giá trị

Đối thuyết	**Miền bác bỏ**	$p-$**giá trị**
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$\lvert t_0 \rvert > t_{\alpha/2,n_1+n_2-2}$	$p-$giá trị $= 2\mathbb{P}(T \geq \lvert t_0 \rvert)$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha,n_1+n_2-2}$	$p-$giá trị $= \mathbb{P}(T \leq t_0)$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha,n_1+n_2-2}$	$p-$giá trị $= \mathbb{P}(T \geq t_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0.

- ☑ **B5.** So sánh và kết luận.

1.3. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TRUNG BÌNH KHI PHƯƠNG SAI KHÁC NHAU CHƯA BIẾT

Mẫu 1: $n_1 = ...$; $\bar{x}_1 = ...$; $s_1^2 = ...$
Mẫu 2: $n_2 = ...$; $\bar{x}_2 = ...$; $s_2^2 = ...$
Phương sai không biết giá trị nhưng biết rằng $\sigma_1^2 \neq \sigma_2^2$.

- ☑ **B1.** Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = ...$
- ☑ **B2.** Xác định mức ý nghĩa α.
- ☑ **B3.** Tính giá trị thống kê kiểm định (TKKĐ)

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$$

.

☑ **B4.** Xác định miền bác bỏ hoặc tính $p-$giá trị Với

$$v = \left\lfloor \frac{\left[(S_1^2/n_1) + (S_2^2/n_2)\right]^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}} \right\rfloor$$

trong đó $\lfloor a \rfloor$ là phép lấy phần nguyên và nhỏ hơn a, ví dụ $\lfloor 38.89 \rfloor = 38$.

Đối thuyết	**Miền bác bỏ**	$p-$**giá trị**
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	$\lvert t_0 \rvert > t_{\alpha/2,v}$	$p-$giá trị $= 2\mathbb{P}(T \geq \lvert t_0 \rvert)$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha,v}$	$p-$giá trị $= \mathbb{P}(T \leq t_0)$
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha,v}$	$p-$giá trị $= \mathbb{P}(T \geq t_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0.

☑ **B5.** So sánh và kết luận.

BÀI 3.1 (Câu 4 - Đề 1 CKI 20-21). Hai máy được sử dụng để rót đầy các chai nhựa. Khối lượng được rót từ hai máy được giả định có phân phối chuẩn với độ lệch chuẩn lần lượt là $\sigma_1 = 0.020$ và $\sigma_2 = 0.025$ ounce. Một nhân viên kỹ thuật cho rằng khối lượng trung bình được rót hai máy là bằng nhau. Hai mẫu ngẫu nhiên từ mỗi máy như sau

Máy 1	16.03	16.01	16.04	15.96	16.05	15.98	16.05	16.02	16.02	15.99
Máy 2	16.02	16.03	15.97	16.04	15.96	16.02	16.01	16.01	15.99	16.00

Kiểm định ý kiến của nhân viên kỹ thuật trên. Sử dụng $\alpha = 0.05$. Tìm $p-$giá trị.

✍ **LỜI GIẢI.**

Gọi X_1 (ounce) là khối lượng được rót vào chai từ máy 1;

X_2 (ounce) là khối lượng được rót vào chai từ máy 2.

Theo đề bài $X_1 \sim N(\mu_1; \sigma_1^2)$; $X_2 \sim N(\mu_2; \sigma_2^2)$

$\sigma_1 = 0.020; \sigma_2 = 0.025$: đã biết.

Mẫu 1: $n_1 = 10$; $\bar{x}_1 = 16.015$.

Mẫu 2: $n_2 = 10$; $\bar{x}_2 = 16.005$.

• **Giả thuyết kiểm định** $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$; $\Rightarrow \Delta_0 = \mu_1 - \mu_2 = 0$; $\alpha = 0.05$.

• **Giá trị thống kê kiểm định**

$$z_0 = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{16.015 - 16.005 - 0}{\sqrt{\dfrac{0.020^2}{10} + \dfrac{0.025^2}{10}}} \approx 0.9877.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $|z_0| > z_{1-\alpha/2}$.

Ta có $\alpha = 0.05$

$\Rightarrow z_{1-\alpha/2} = z_{0.975} = 1.96$.

• **So sánh và kết luận:**

Ta có: $|z_0| > z_{1-\alpha/2}$

$\Leftrightarrow |0.9877| > 1.96$

$\Leftrightarrow 0.9877 > 1.96$ (sai)

$\Rightarrow$ chưa đủ cơ sở để bác bỏ $H_0 : \mu_1 = \mu_2$.

Kết luận: Với mức ý nghĩa 5% thì khối lượng trung bình được rót hai máy là bằng nhau hay ý kiến của nhân viên kỹ thuật trên là đúng.
⋇ Tính p-giá trị:

$$z_0 = 0.9877 \Rightarrow p\text{-giá trị} = 2\left(1 - \Phi(|0.9877|)\right) = 2\left(1 - 0.8384\right) = 03232.$$

□

BÀI 3.2 (Câu 5 - Đề 2 CKI 19-20). Hai công thức khác nhau của nhiên liệu động cơ ôxy hóa đang được thử nghiệm để nghiên cứu số octane của chúng. Phương sai chỉ số octane của công thức thứ nhất $\sigma_1^2 = 1.5$ và công thức thứ hai $\sigma_2^2 = 1.2$. Hai mẫu ngẫu nhiên có cỡ mẫu $n_1 = 15$ và $n_2 = 20$ được nghiên cứu có chỉ số octane trung bình lần lượt là $\bar{x}_1 = 89.6$ và $\bar{x}_2 = 92.5$. Với giả sử có phân phối chuẩn.
Nếu công thức 2 tạo ra một số octane cao hơn so với công thức 1, nhà sản xuất muốn phát hiện nó. Xây dựng và kiểm định giả thuyết thích hợp sử dụng $\alpha = 0.05$ và tính $p-$value.

✍ **LỜI GIẢI.**

Gọi X_1 : số octane được tạo ra từ công thức 1.

X_2 : số octane được tạo ra từ công thức 2.

Mẫu 1: $n_1 = 15$; $\bar{x}_1 = 89.6$; $\sigma_1^2 = 1.5$
Mẫu 2: $n_2 = 20$; $\bar{x}_2 = 92.5$; $\sigma_2^2 = 1.2$
Phương sai $\sigma_1^2; \sigma_2^2$: đã biết.

• **Giả thuyết KĐ:** $\begin{cases} H_0 : \mu_1 \geq \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$: KĐ 1 phía, $\mu_0 = \mu_1 - \mu_2 = 0$.

• **Mức ý nghĩa:** $\alpha = 0.05$.

• **Giá trị Thống kê kiểm định:**

$$z_0 = \frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{89.6 - 92.5 - 0}{\sqrt{\dfrac{1.5^2}{15} + \dfrac{1.2^2}{20}}} \approx -6.15491.$$

• **Miền bác bỏ:** Bác bỏ H_0 nếu $z_0 < -z_{1-\alpha}$.
Ta có $\alpha = 0.05$
$\Rightarrow -z_{1-\alpha} = -z_{0.95} = -1.64485$.

• **So sánh và kết luận:**

Ta có: $z_0 < -z_{1-\alpha}$

$\Leftrightarrow -6.15491 < -1.64485$ (đúng)

$\Rightarrow$ Bác bỏ $H_0 : \mu_1 \geq \mu_2$.
Kết luận: Với mức ý nghĩa 5%, số octane trung bình được tạo ra từ công thức 2 cao hơn công thức 1.

• **$p-$giá trị**

$p-$giá trị $= \Phi(z_0) = \Phi(-6.15491) \approx 3.771 \cdot 10^{-10} \approx 0$. □

BÀI 3.3 (Câu 4 - Đề 1 - CKI 19-20). Hai loại nhựa phù hợp để sử dụng cho một nhà sản xuất linh kiện điện tử. Sức chịu phá hủy của loại nhựa này là quan trọng và được giả sử có phân phối chuẩn. Được biết, $\sigma_1 = \sigma_2 = 1$ psi. Từ một mẫu ngẫu nhiên có kích thước $n_1 = 10$ và $n_2 = 12$, ta có được $\bar{x}_1 = 162.5$ và $\bar{x}_2 = 155.0$. Công ty sẽ không áp dụng nhựa loại 1 trừ khi sức chịu phá hủy trung bình của nó vượt quá nhựa loại 2 ít nhất 10 psi.
Trên cơ sở thông tin đó, ta có nên sử dụng nhựa loại 1? Sử dụng $\alpha = 0.05$ đưa ra câu trả lời.

✍ **LỜI GIẢI.**

Gọi X_1 : sức chịu phá hủy của nhựa loại 1.

X_2 : sức chịu phá hủy của nhựa loại 2.

Mẫu 1: $n_1 = 10$; $\bar{x}_1 = 162.5$.
Mẫu 2: $n_2 = 12$; $\bar{x}_2 = 155.0$.
$\sigma_1 = \sigma_2 = 1$ đã biết.

• **Giả Thuyết KĐ:** $\begin{cases} H_0: \mu_1 - \mu_2 \geq 10 \\ H_1: \mu_1 - \mu_2 < 10 \end{cases}$: KĐ 1 phía, $\mu_0 = \mu_1 - \mu_2 = 10$.

• **Mức ý nghĩa:** $\alpha = 0.05$.

• **Giá trị Thống kê kiểm định:**

$$z_0 = \frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{162.5 - 155.0 - 10}{\sqrt{\frac{1^2}{10} + \frac{1^2}{12}}} \approx -5.83874.$$

• **Miền bác bỏ:** Bác bỏ H_0 nếu $z_0 < -z_{1-\alpha}$.

Ta có $\alpha = 0.05$

$\Rightarrow -z_{1-\alpha} = -z_{0.95} = -1.64485$.

• **So sánh và kết luận:**

Ta có: $z_0 < -z_{1-\alpha}$

$\Leftrightarrow -5.83874 < -1.64485$ (đúng)

$\Rightarrow$ Bác bỏ $H_0: \mu_1 - \mu_2 \geq 10$.

Kết luận: Với mức ý nghĩa 5%, sức chịu phá hủy trung bình của nhựa loại 1 không vượt quá nhựa loại 2 ít nhất 10 psi, hay công ty sẽ không áp dụng nhựa loại 1. □

BÀI 3.4 (Câu 4 - Đề 2 CKI 20-21). Đường kính của các thanh thép được sản xuất trên hai máy đúc khác nhau đang được nghiên cứu. Hai mẫu ngẫu nhiên có cỡ mẫu $n_1 = 15$, $n_2 = 17$ được chọn có trung bình và phương sai mẫu $\bar{x}_1 = 8{,}73$; $s_1^2 = 0{,}35$ và $\bar{x}_2 = 8{,}68$; $s_2^2 = 0{,}40$. Giả sử rằng $\sigma_1^2 = \sigma_2^2$ và tổng thể có phân phối chuẩn. Có bằng chứng để khẳng định rằng hai máy sản xuất thanh thép có đường kính trung bình khác nhau không? Sử dụng $\alpha = 0{,}05$ khi đưa ra kết luận này. Tìm $p-$giá trị.

✍ **LỜI GIẢI.**

Gọi X_1: Đường kính của các thanh thép được sản xuất trên máy đúc thứ nhất.

X_2: Đường kính của các thanh thép được sản xuất trên máy đúc thứ hai.

Mẫu 1: $n_1 = 15$; $\bar{x}_1 = 8.73$; $s_1^2 = 0.35$.

Mẫu 2: $n_2 = 17$; $\bar{x}_2 = 8.68$; $s_2^2 = 0.40$.

Phương sai $\sigma_1^2 = \sigma_2^2$ chưa biết.

• **GTKĐ:** $\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$: KĐ 2 phía, $\mu_0 = \mu_1 - \mu_2 = 0$.

• **Mức ý nghĩa:** $\alpha = 0.05$.

• **Giá trị Thống kê kiểm định:**

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(15 - 1)0.35 + (17 - 1)0.40}{15 + 17 - 2} = \frac{113}{300}.$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{8.73 - 8.68 - 0}{\sqrt{\frac{113/300}{15} + \frac{113/300}{17}}} \approx 0.22998.$$

• **Miền bác bỏ:** Bác bỏ H_0 nếu $|t_0| > t_{\frac{\alpha}{2};n_1+n_2-2}$.

Ta có $\alpha = 0.05$

$\Rightarrow t_{\frac{\alpha}{2};n_1+n_2-2} = t_{0.025;30} = 2.042$.

• **So sánh và kết luận:**

Ta có: $|t_0| > t_{\frac{\alpha}{2};n_1+n_2-2}$

$\Leftrightarrow |0.22998| > 2.042$

$\Leftrightarrow 0.22998 > 2.042$ (sai)

$\Rightarrow$ Chưa đủ cơ sở để bác bỏ $H_0: \mu_1 = \mu_2$.

Kết luận: Với mức ý nghĩa 5%, đường kính trung bình của các thanh thép được sản xuất trên máy đúc

thứ nhất không khác máy thứ hai.
• $p-$**giá trị**
Vì bậc tự do $n_1 + n_2 - 2 = 30 \geq 30$ nên ta có thể xấp xỉ $p-$giá trị bằng phân phối chuẩn.
$p-$giá trị $\approx 2\left[1 - \Phi(|z_0|)\right] = 2\left[1 - \Phi(|0.22998|)\right] = 2\left[1 - \Phi(0.22998)\right] = 2\,(1 - 0.59095) = 0.8181.$ □

BÀI 3.5 (Câu 2 - Đề 1 CKII 21-22). Hai chất xúc tác có thể được sử dụng trong một phản ứng hóa học. Mười hai phản ứng được cho sử dụng chất xúc tác 1, dẫn đến hiệu suất trung bình là 86 (đv: %) và độ lệch chuẩn mẫu là 3. Mười lăm phản ứng được cho sử dụng chất xúc tác 2 và kết quả là hiệu suất trung bình 89 với độ lệch chuẩn mẫu là 2. Giả sử hiệu suất các phản ứng xấp xỉ phân phối chuẩn với cùng độ lệch chuẩn. Có bằng chứng để khẳng định rằng chất xúc tác 2 tạo ra hiệu suất trung bình cao hơn chất xúc tác 1 hay không? Sử dụng $\alpha = 0.01$. (*Yêu cầu dùng cả 2 phương pháp: miền bác bỏ và p-giá trị*)

✍ **LỜI GIẢI.**
Gọi X_1 (%) là hiệu suất phản ứng khi sử dụng chất xúc tác 1

X_2 (%) là hiệu suất phản ứng khi sử dụng chất xúc tác 2.
Theo đề bài $X_1 \sim N(\mu_1; \sigma_1^2)$; $X_2 \sim N(\mu_2; \sigma_2^2)$
$\sigma_1 = \sigma_2$ chưa biết.
Mẫu 1: $n_1 = 12$; $\bar{x}_1 = 86$ và $s_1 = 3$.
Mẫu 2: $n_2 = 15$; $\bar{x}_2 = 89$ và $s_2 = 2$.
• **Giả thuyết kiểm định** $\begin{cases} H_0 : \mu_1 \geq \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$; $\quad \Rightarrow \mu_1 - \mu_2 = 0$; $\alpha = 0.01$.
• **Giá trị thống kê kiểm định**

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(12-1).3^2 + (15-1).2^2}{12+15-2} = 6.2$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}} = \frac{86 - 89 - 0}{\sqrt{\dfrac{6.2}{12} + \dfrac{6.2}{15}}} = -3.1109.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $t_0 < -t_{\alpha;n_1+n_2-2}$.
Ta có $\alpha = 0.01$
$\Rightarrow -t_{\alpha;n_1+n_2-2} = -t_{0.01;25} = -2.4851.$
• **So sánh và kết luận:**
Ta có: $\quad t_0 < -t_{\alpha;n_1+n_2-2}$

$-3.1109 < -2.4851 \quad$ (đúng)
$\Rightarrow$ Bác bỏ $H_0 : \mu_1 \geq \mu_2$.
Kết luận: Với mức ý nghĩa 1% thì có bằng chứng để khẳng định được rằng sử dụng chất xúc tác 2 tạo ra hiệu suất trung bình cao hơn khi dùng chất xúc tác 1.
⁎ Cách 2: Dùng p - giá trị:
$t_0 = -3.1109 \Rightarrow p$-giá trị $= \mathbb{P}\left(T_{12+15-2} \leq t_0\right) = \mathbb{P}\left(T_{12+15-2} \leq -3.1109\right) \approx 0.0023.$
Vì $p-$giá trị $\approx 0.0023 < 0.01 = \alpha$ nên bác bỏ $H_0 : \mu_1 \geq \mu_2$.
□

BÀI 3.6 (Câu 5 - Đề 1 CKI 21-22). Trong sản xuất chất bán dẫn, khắc hóa chất ướt thường được sử dụng để loại bỏ silic từ mặt sau của tấm wafer trước khi kim loại hóa. Tỷ lệ ăn mòn (etch) là một đặc tính quan trọng trong quá trình này và được biết là tuân theo phân phối chuẩn. Hai phương pháp khắc khác nhau đã được so sánh bằng cách sử dụng hai mẫu ngẫu nhiên gồm 8 tấm wafer cho mỗi dung dịch. Gọi X và Y lần lượt là tỷ lệ ăn mòn tương ứng với phương pháp khắc I và II. Kết quả quan sát được như sau

Phương pháp I (x_i)	11.1	11.2	11.4	11.1	11.7	11.4	11.2	11.1
Phương pháp II (y_i)	9.1	8.9	8.9	9.1	8.9	9.2	9	9.3

Giả sử rằng phương sai của hai tổng thể bằng nhau. Với mức ý nghĩa 8%, hỏi dữ liệu trên có hỗ trợ cho tuyên bố rằng "tỷ lệ ăn mòn trung bình là giống nhau cho cả hai phương pháp" không?

✍ **LỜI GIẢI.**

Theo đề bài $X \sim N(\mu_X; \sigma_X^2)$; $Y \sim N(\mu_Y; \sigma_Y^2)$

$\sigma_X = \sigma_Y$ chưa biết.

Mẫu 1: $n_X = 8$; $\bar{x} = 11.275$ và $s_X^2 = 0.045$.

Mẫu 2: $n_Y = 8$; $\bar{y} = 9.05$ và $s_Y^2 \approx 0.0229$.

• **Giả thuyết kiểm định** $\begin{cases} H_0 : \mu_X = \mu_Y \\ H_1 : \mu_X \neq \mu_Y \end{cases}$; $\quad \Rightarrow \mu_X - \mu_Y = 0$; $\alpha = 8\% = 0.08$.

• **Giá trị thống kê kiểm định**

$$s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2} = \frac{(8-1)\cdot 0.045 + (8-1)\cdot 0.0229}{8+8-2} \approx 0.0340$$

$$t_0 = \frac{\bar{x} - \bar{y} - (\mu_X - \mu_Y)}{\sqrt{\dfrac{s_p^2}{n_X} + \dfrac{s_p^2}{n_Y}}} = \frac{11.275 - 9.05 - 0}{\sqrt{\dfrac{0.0340}{8} + \dfrac{0.0340}{8}}} = 24.1335.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $|t_0| > t_{\alpha/2; n_X + n_Y - 2}$.

Ta có $\alpha = 0.08$

$\Rightarrow t_{\alpha/2; n_X+n_Y-2} = t_{0.04;14} = 1.89$.

• **So sánh và kết luận:**

Ta có: $\quad |t_0| > t_{\alpha/2; n_X+n_Y-2}$

$|24.1335| > t_{\alpha/2; n_X+n_Y-2}$

$24.1335 > 1.89 \quad$ (đúng)

$\Rightarrow$ Bác bỏ $H_0 : \mu_X = \mu_Y$.

Kết luận: Với mức ý nghĩa 8% thì tỷ lệ ăn mòn trung bình của hai phương pháp là khác nhau. □

BÀI 3.7 (Câu 3 - Đề 2 CKII 20-21). Để tìm ra liệu một loại huyết thanh mới có kìm hãm được bệnh bạch cầu hay không, 9 con chuột, tất cả các con đều trong giai đoạn tiến triển của bệnh, được chọn. Năm con chuột nhận trị liệu và 4 con không. Thời gian sống, theo năm, từ thời điểm thí nghiệm bắt đầu là như sau

Trị liệu	2.1	5.3	1.4	4.6	0.9
Không trị liệu	1.9	0.5	2.8	3.1	

Tại mức ý nghĩa 5%, huyết thanh có thể được nói là có hiệu quả hay không? Giả sử hai tổng thể có phân phối chuẩn với các phương sai bằng nhau.

✍ **LỜI GIẢI.**

Gọi X_1 (năm) là thời gian sống của chuột được trị liệu bằng huyết thanh;

X_2 (năm) là thời gian sống của chuột không được trị liệu bằng huyết thanh.

Theo đề bài $X_1 \sim N(\mu_1; \sigma_1^2)$; $X_2 \sim N(\mu_2; \sigma_2^2)$

$\sigma_1^2 = \sigma_2^2$ chưa biết.

Mẫu 1: $n_1 = 5$; $\bar{x}_1 = 2.86$ và $s_1^2 = 3.883$.

Mẫu 2: $n_2 = 4$; $\bar{x}_2 = 2.075$ và $s_2^2 = 1.3625$.

• **Giả thuyết kiểm định** $\begin{cases} H_0 : \mu_1 \leq \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$; $\quad \Rightarrow \Delta_0 = \mu_1 - \mu_2 = 0$; $\alpha = 0.05$.

• **Giá trị thống kê kiểm định**

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(5-1)\cdot 3.883 + (4-1)\cdot 1.3625}{5+4-2} \approx 2.8028$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = \frac{2.86 - 2.075 - 0}{\sqrt{\frac{2.8028}{5} + \frac{2.8028}{4}}} \approx 0.6990.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $t_0 > t_{\alpha;n_1+n_2-2}$.
Ta có $\alpha = 0.05$
$\Rightarrow t_{\alpha;n_1+n_2-2} = t_{0.05;7} = 1.895$.
• **So sánh và kết luận:**
Ta có: $t_0 > t_{\alpha;n_1+n_2-2}$
$\Leftrightarrow 0.6990 > 1.895$ (sai)
$\Rightarrow$ chưa đủ cơ sở để bác bỏ $H_0 : \mu_1 \le \mu_2$.
Kết luận: Với mức ý nghĩa 5% thì huyết thanh không thể được nói là có hiệu quả. □

BÀI 3.8 (Câu 4 - Đề 1 HKI 22-23). Một nhà sản xuất công bố rằng độ bền kéo trung bình của sợi A cao hơn độ bền kéo trung bình của sợi B. Để kiểm định công bố này, 50 mẫu của mỗi loại sợi được kiểm tra dưới các điều kiện tương tự. Loại sợi A có độ bền kéo trung bình là 86,7 kg với độ lệch chuẩn mẫu 6,28 kg, trong khi loại sợi B có độ bền kéo trung bình là 85,4 kg với độ lệch chuẩn mẫu là 5,61 kg. Hãy kiểm định công bố của nhà sản xuất bằng cách sử dụng mức ý nghĩa 1%.

✍ **LỜI GIẢI.**
Gọi X_1 (kg) là độ bền kéo của loại sợi A,
X_2 (kg) là độ bền kéo của loại sợi B.
Giả sử: $X_1 \sim N(\mu_1; \sigma_1^2)$, $X_2 \sim N(\mu_2; \sigma_2^2)$
σ_1^2; σ_2^2: **chưa có thông tin**.
Mẫu 1: $n_1 = 50$; $\bar{x}_1 = 86{,}7$; $s_1 \approx 6{,}28$.
Mẫu 2: $n_2 = 50$; $\bar{x}_2 = 85{,}4$; $s_2 \approx 5{,}61$.
⁕ **Kiểm định phương sai**

- ☑ Giả thuyết kiểm định $\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : \sigma_1^2 \neq \sigma_2^2. \end{cases}$

- ☑ Mức ý nghĩa $\alpha = 0{,}01$.

- ☑ Giá trị thống kê kiểm định

$$f_0 = \frac{s_1^2}{s_2^2} = \frac{6{,}28^2}{5{,}61^2} \approx 1{,}2531.$$

- ☑ Miền bác bỏ: bác bỏ H_0 khi $f_0 > f_{\alpha/2,n_1-1,n_2-1}$ hoặc $f_0 < \dfrac{1}{f_{\alpha/2,n_2-1,n_1-1}}$

  Ta có $\alpha = 0{,}01 \Rightarrow f_{\alpha/2,n_1-1,n_2-1} = f_{\alpha/2,n_2-1,n_1-1} = f_{0{,}005,49,49} = 2{,}11$.

- ☑ So sánh:

  Ta có $f_0 > f_{\alpha/2,n_1-1,n_2-1}$ hoặc $f_0 < \dfrac{1}{f_{\alpha/2,n_2-1,n_1-1}}$

  $\Leftrightarrow 1{,}2531 > 2{,}11$ (sai)

  $\Leftrightarrow 1{,}2531 < \dfrac{1}{2{,}11}$

  $\Leftrightarrow 1{,}2531 < 0{,}4739$ (sai)

  $\Rightarrow$ chưa đủ cơ sở để bác bỏ $H_0 : \sigma_1^2 = \sigma_2^2$
  $\Rightarrow \sigma_1^2 = \sigma_2^2$.

⁕ **Kiểm định hai trung bình khi $\sigma_1^2 = \sigma_2^2$**

1) Giả thuyết kiểm định: $\begin{cases} H_0 : \mu_1 \le \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$ kiểm định 1 phía, $\Delta_0 = \mu_1 - \mu_2 = 0$.

2) Mức ý nghĩa: $\alpha = 0{,}01$.

3) Giá trị thống kê kiểm định

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(50-1)\cdot 6{,}28^2 + (50-1)\cdot 5{,}61^2}{50+50-2} = 35{,}45525$$

$$\Rightarrow t_0 = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1-\mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}} = \frac{86{,}7 - 85{,}4 - 0}{\sqrt{\dfrac{35{,}45525}{50} + \dfrac{35{,}45525}{50}}} \approx 1{,}0916.$$

4) Miền bác bỏ: bác bỏ H_0 khi $t_0 > t_{\alpha, n_1+n_2-2}$.
Ta có $\alpha = 0{,}01 \Rightarrow t_{\alpha,n_1+n_2-2} = t_{0{,}01{,}98} \approx z_{0{,}99} \approx 2{,}33$.

5) So sánh
Ta có $\quad t_0 > t_{\alpha,n_1+n_2-2}$
$\Leftrightarrow 1{,}0916 > 2{,}33$ (sai)
$\Rightarrow$ chưa đủ cở sở để bác bỏ $H_0 : \mu_1 \le \mu_2$.
Kết luận: Với mức ý nghĩa 1%, độ bền kéo trung bình của sợi A không cao hơn so với độ bền kéo trung bình của sợi B hay công bố của nhà sản xuất không đúng.

□

BÀI 3.9 (Câu 3 - Đề 1 CKII 20-21). Một nghiên cứu được thực hiện để xem việc tăng nồng độ cơ chất có tác dụng đáng kể đến tốc độ của một phản ứng hóa học hay không. Với một nồng độ cơ chất 1.5 mol/l, phản ứng được chạy 15 lần, với tốc độ trung bình 7.5 micromoles mỗi 30 phút và độ lệch chuẩn 1.5. Với một nồng độ cơ chất 2.0 mol/l, 12 lần chạy được thực hiện, thu được tốc độ trung bình 8.8 micromoles mỗi 30 phút và độ lệch chuẩn 1.2. Có lý do để tin rằng việc tăng nồng độ cơ chất làm tăng tốc độ trung bình của phản ứng 0.5 micromoles mỗi 30 phút hay không? Sử dụng mức ý nghĩa 0.01 và giả sử rằng các tổng thể là xấp xỉ chuẩn với các phương sai bằng nhau.

✍ **LỜI GIẢI.**

Gọi X_1 (micromoles) là tốc độ phản ứng ứng với nồng độ cơ chất 1.5 mol/l;

X_2(micromoles) là tốc độ phản ứng ứng với nồng độ cơ chất 2 mol/l;.

Theo đề bài $X_1 \sim N(\mu_1; \sigma_1^2)$; $X_2 \sim N(\mu_2; \sigma_2^2)$
$\sigma_1 = \sigma_2$ chưa biết.
Mẫu 1: $n_1 = 15$; $\bar{x}_1 = 7.5$ và $s_1 = 1.5$.
Mẫu 2: $n_2 = 12$; $\bar{x}_2 = 8.8$ và $s_2 = 1.2$.

• **Giả thuyết kiểm định** $\begin{cases} H_0 : \mu_1 - \mu_2 = -0.5 \\ H_1 : \mu_1 - \mu_2 \neq -0.5 \end{cases}$; $\quad \Rightarrow \Delta_0 = \mu_1 - \mu_2 = -0.5$; $\alpha = 0.01$.

• **Giá trị thống kê kiểm định**

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(15-1)\cdot 1.5^2 + (12-1)\cdot 1.2^2}{15+12-2} = 1.8936$$

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}} = \frac{7.5 - 8.8 - (-0.5)}{\sqrt{\dfrac{1.8936}{12} + \dfrac{1.8936}{15}}} = -1.5011.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $|t_0| > t_{\alpha/2; n_1+n_2-2}$.
Ta có $\alpha = 0.01$

$\Rightarrow t_{\alpha/2;n_1+n_2-2} = t_{0.005;25} = 2.787.$
• So sánh và kết luận:
Ta có: $\quad |t_0| > t_{\alpha/2;n_1+n_2-2}$
$\quad 1.5011 > 2.787 \quad$ (sai)
$\Rightarrow$ chưa đủ cơ sở để bác bỏ $H_0 : \mu_1 - \mu_2 = -0.5$.
Kết luận: Với mức ý nghĩa 1% thì không có lý do để tin rằng việc tăng nồng độ cơ chất làm tăng tốc độ trung bình của phản ứng 0.5 micromoles mỗi 30 phút. □

BÀI 3.10 (Câu 4 - Đề 2 CKII 21-22 (CNTT hệ CLC)). Một công ty sản xuất xe ô tô chế tạo ra một dòng xe ô tô **Z**, công ty kiểm tra xe bằng cách cho 30 xe chạy trên cùng 1 quãng đường 100 km và đo lượng xăng sử dụng X (Đv: lít), kết quả cho ở bảng sau

X (lít)	7.0	7.1	7.2	7.3	7.4	7.5
Số xe	5	7	6	6	3	3

Giả sử lượng xăng tiêu hao là đại lượng ngẫu nhiên có phân phối chuẩn.

a) Lời khẳng định: "Lượng xăng tiêu hao trung bình của dòng xe này là 7.2 lít (trên 100 km)" có được chấp nhận hay không? (Mức ý nghĩa $\alpha = 5\%$)

b) Những xe có lượng xăng tiêu hao từ 7.3 lít trở lên được xếp vào loại không đạt tiêu chuẩn về nhiên liệu. Giám đốc công ty khẳng định rằng tỷ lệ xe không đạt tiêu chuẩn về nhiên liệu tối đa bằng 10%. Với dữ liệu khảo sát đã cho, ta có đủ bằng chứng để bác bỏ ý kiến trên không? $\alpha = 1\%$.

c) Khảo sát lượng xăng tiêu hao Y (Đv: lít/100 km) trên một mẫu gồm 25 xe của dòng xe **W** cùng hãng, tính được $\bar{y} = 7.35$ (lít/100 km) và $s_Y = 0.2$. Có ý kiến cho rằng dòng xe **Z** tiêu hao ít nhiên liệu hơn dòng xe **W**. Với mức ý nghĩa $\alpha = 1\%$, hãy kiểm định ý kiến trên. Giả sử phương sai của X và Y bằng nhau.

✍ **LỜI GIẢI.**

a) Lời khẳng định: "Lượng xăng tiêu hao trung bình của dòng xe này là 7.2 lít (trên 100 km)" có được chấp nhận hay không? (Mức ý nghĩa $\alpha = 5\%$)
Theo đề bài $X \sim N(\mu; \sigma^2)$
σ^2: chưa biết.
Từ bảng số liệu đề bài, ta tính được: $n = 30$; $\bar{x} \approx 7.2133$; $s \approx 0.1570$.
• Giả Thuyết KĐ: $\begin{cases} H_0 : \mu = 7.2 \\ H_1 : \mu \neq 7.2 \end{cases}$: $\quad$ KĐ 2 phía, $\mu_0 = 7.2$.
• Mức ý nghĩa: $\alpha = 0.05$.
• Giá trị Thống kê kiểm định:

$$t_0 = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{7.2133 - 7.2}{\frac{0.1570}{\sqrt{30}}} \approx 0.4640.$$

• Miền bác bỏ: Bác bỏ H_0 nếu $|t_0| > t_{\alpha/2;n-1}$.
Ta có $\alpha = 0.05$
$\Rightarrow t_{\alpha/2;n-1} = t_{0.025;29} = 2.045.$
• So sánh và kết luận:
Ta có: $\quad |t_0| > t_{\alpha/2;n-1}$
$\quad |0.4640| > 2.045$
$\quad \Leftrightarrow 0.4640 > 2.045 \quad$ (sai)
$\Rightarrow$ Chưa đủ cơ sở để bác bỏ $H_0 : \mu = 7.2$.
Kết luận: Với mức ý nghĩa 5% thì lượng xăng tiêu hao trung bình của dòng xe này là 7.2 lít (trên 100 km).

b) Những xe có lượng xăng tiêu hao từ 7.3 lít trở lên được xếp vào loại không đạt tiêu chuẩn về nhiên liệu. Giám đốc công ty khẳng định rằng tỷ lệ xe không đạt tiêu chuẩn về nhiên liệu tối đa bằng 10%. Với dữ liệu khảo sát đã cho, ta có đủ bằng chứng để bác bỏ ý kiến trên không? $\alpha = 1\%$.

Gọi Y_1 là số xe không đạt tiêu chuẩn.

Ta có: $n = 30$, $y_1 = 6 + 3 + 3 = 12 \Rightarrow$ tỷ lệ mẫu $\hat{p} = \dfrac{y_1}{n} = \dfrac{12}{30} = 0.4$.

• **GTKĐ:** $\begin{cases} H_0 : p \leq 0.1 \\ H_1 : p > 0.1 \end{cases}$: KĐ 1 phía, $p_0 = 10\% = 0.1$.

• **Mức ý nghĩa:** $\alpha = 1\% = 0.01$.

• **Giá trị Thống kê kiểm định:**

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.4 - 0.1}{\sqrt{\dfrac{0.1(1-0.1)}{30}}} \approx 5.4772.$$

• **Miền bác bỏ:** Bác bỏ H_0 nếu $z_0 > z_{1-\alpha}$.

Ta có $\alpha = 0.01$

$\Rightarrow z_{1-\alpha} = z_{0.99} \approx 2.33$.

• **So sánh và kết luận:**

Ta có: $z_0 > z_{1-\alpha}$

$\Leftrightarrow 5.4772 > 2.33$ (đúng)

$\Rightarrow$ bác bỏ $H_0 : p \leq 0.1$.

Kết luận: Với mức ý nghĩa 1%, ta có đủ bằng chứng để bác bỏ ý kiến trên của vị giám đốc.

c) Khảo sát lượng xăng tiêu hao Y (Đv: lít/100 km) trên một mẫu gồm 25 xe của dòng xe **W** cùng hãng, tính được $\bar{y} = 7.35$ (lít/100 km) và $s_Y = 0.2$. Có ý kiến cho rằng dòng xe **Z** tiêu hao ít nhiên liệu hơn dòng xe **W**. Với mức ý nghĩa $\alpha = 1\%$, hãy kiểm định ý kiến trên. Giả sử phương sai của X và Y bằng nhau.

Theo đề bài $X \sim N(\mu_X; \sigma_X^2)$; $Y \sim N(\mu_Y; \sigma_Y^2)$

$\sigma_X = \sigma_Y$ chưa biết.

Mẫu 1: $n_X = 30$; $\bar{x} \approx 7.2133$ và $s_X \approx 0.1570$.

Mẫu 2: $n_Y = 35$; $\bar{y} = 7.35$ và $s_Y = 0.2$.

• **Giả thuyết kiểm định** $\begin{cases} H_0 : \mu_X \geq \mu_Y \\ H_1 : \mu_X < \mu_Y \end{cases}$; $\Rightarrow \mu_X - \mu_Y = 0$; $\alpha = 0.01$.

• **Giá trị thống kê kiểm định**

$$s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2} = \frac{(30-1) \cdot 0.1570^2 + (25-1) \cdot 0.2^2}{30 + 25 - 2} \approx 0.0316$$

$$t_0 = \frac{\bar{x} - \bar{y} - (\mu_X - \mu_Y)}{\sqrt{\dfrac{s_p^2}{n_X} + \dfrac{s_p^2}{n_Y}}} = \frac{7.2133 - 7.35 - 0}{\sqrt{\dfrac{0.0316}{30} + \dfrac{0.0316}{25}}} = -2.8397.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $t_0 < -t_{\alpha; n_X + n_Y - 2}$.

Ta có $\alpha = 0.01$

$\Rightarrow -t_{\alpha; n_X+n_Y-2} = -t_{0.01;53} \approx -z_{0.99} \approx -2.33$.

• **So sánh và kết luận:**

Ta có: $t_0 < -t_{\alpha; n_X+n_Y-2}$

$-2.8397 < -2.33$ (đúng)

$\Rightarrow$ Bác bỏ $H_0 : \mu_X \geq \mu_Y$.

Kết luận: Với mức ý nghĩa 1% thì dòng xe **Z** tiêu hao ít nhiên liệu hơn dòng xe **W**.

□

BÀI 3.11 (Câu 4 - Đề 2 CKI 21-22). Trong một nhà máy sản xuất linh kiện điện tử, các kỹ sư cân nhắc lựa chọn giữa hai loại nhựa để tiến hành sản xuất linh kiện. Một yếu tố quan trọng được cân nhắc là sức chịu lực phá hủy (đơn vị: psi) của từng loại nhựa. Đối với loại nhựa 1, các kỹ sư kiểm tra trên một mẫu cỡ $m = 13$, tính được sức chịu lực phá hủy trung bình là $\bar{x} = 163.25$ và độ lệch chuẩn mẫu $s_x = 1.1$. Thử nghiệm trên một mẫu cỡ $n = 14$ đối với loại nhựa 2, tính được sức chịu lực phá hủy trung bình là $\bar{y} = 163.5$ và độ lệch chuẩn mẫu $s_y = 1.95$. Với mức ý nghĩa $\alpha = 3\%$, ta có thể khẳng định rằng sức chịu lực phá hủy của hai loại nhựa này là như nhau hay không? Giả sử rằng sức chịu lực phá hủy của hai loại nhựa tuân theo phân phối chuẩn và có phương sai khác nhau.

✍ **LỜI GIẢI.**

Gọi X_1 : sức chịu lực phá hủy của nhựa loại 1.

X_2 : sức chịu lực phá hủy của nhựa loại 2.

Mẫu 1: $n_1 = 13$; $\bar{x}_1 = 163.25$; $s_1 = 1.1$.

Mẫu 2: $n_2 = 14$; $\bar{x}_2 = 163.5$; $s_2 = 1.95$.

$\sigma_1^2 \neq \sigma_2^2$: chưa biết.

- **Giả Thuyết KĐ:** $\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$: KĐ 2 phía, $\Delta_0 = \mu_1 - \mu_2 = 0$.
- **Mức ý nghĩa:** $\alpha = 0.03$.
- **Giá trị Thống kê kiểm định:**

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{163.25 - 163.5 - 0}{\sqrt{\dfrac{1.1^2}{13} + \dfrac{1.95^2}{14}}} \approx -0.4140.$$

- **Miền bác bỏ:** Bác bỏ H_0 nếu $|t_0| > t_{\alpha/2;v}$ trong đó

$$v = \left\lfloor \frac{\left[(s_1^2/n_1) + (s_2^2/n_2)\right]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \right\rfloor \approx \lfloor 20.7915 \rfloor = 20.$$

$\alpha = 0.03 \Rightarrow t_{\alpha/2;v} = t_{0.015;20} = 2.34$.

- **So sánh và kết luận:**

Ta có: $|t_0| > t_{\alpha/2;v}$

$\Leftrightarrow |-0.4140| > 2.34$

$\Leftrightarrow 0.4140 > 2.34$ (sai)

$\Rightarrow$ chưa đủ cơ sở để bác bỏ $H_0 : \mu_1 = \mu_2$.

Kết luận: Với mức ý nghĩa 3%, ta có thể khẳng định rằng sức chịu lực phá hủy của hai loại nhựa này là như nhau.

□

2. KIỂM ĐỊNH GIẢ THUYẾT SO SÁNH HAI TỶ LỆ

Gọi Y_1, Y_2 lần lượt là số phần tử có tính chất A trong mẫu 1 (có n_1 phần tử), mẫu 2 (có n_2 phần tử) tương ứng.

Mẫu 1: $n_1 = ...$; $y_1 = ...$ $\Rightarrow$ tỷ lệ mẫu $\hat{P}_1 = \dfrac{y_1}{n_1}$.

Mẫu 2: $n_2 = ...$; $y_2 = ...$ $\Rightarrow$ tỷ lệ mẫu $\hat{P}_2 = \dfrac{y_2}{n_2}$.

$\Rightarrow \hat{P} = \dfrac{y_1 + y_2}{n_1 + n_2}$

☑ **B1.** Phát biểu giả thuyết kiểm định $\Rightarrow \Delta_0 = p_1 - p_2$.

☑ **B2.** Xác định mức ý nghĩa α.

☑ **B3.** Tính giá trị thống kê kiểm định (TKKĐ)

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2 - \Delta_0}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

B4. Xác định miền bác bỏ hoặc tính $p-$giá trị.

Đối thuyết	**Miền bác bỏ H_0**	**$p-$giá trị**
$H_1 : p_1 - p_2 \neq \Delta_0$	$\lvert z_0\rvert > z_{1-\alpha/2}$	$p-$giá trị $= 2\left[1 - \Phi(\lvert z_0\rvert)\right]$
$H_1 : p_1 - p_2 < \Delta_0$	$z_0 < -z_{1-\alpha}$	$p-$giá trị $= \Phi(z_0)$
$H_1 : p_1 - p_2 > \Delta_0$	$z_0 > z_{1-\alpha}$	$p-$giá trị $= 1 - \Phi(z_0)$

Ngược lại, chưa đủ cơ sở bác bỏ H_0.

☑ **B5.** So sánh và kết luận.

BÀI 3.12 (Câu 4 - Đề 2 HKI 22-23). Trong một nghiên cứu để ước tính tỷ lệ cư dân trong một thành phố nào đó và các vùng ngoại ô của nó ủng hộ việc xây dựng nhà máy năng lượng hạt nhân, người ta thấy rằng 63 trong 100 cư dân thành thị ủng hộ việc xây dựng, trong khi chỉ 59 trong 125 cư dân ngoại ô là ủng hộ. Hỏi có sự khác biệt giữa tỷ lệ cư dân thành thị và ngoại ô ủng hộ việc xây dựng nhà máy hạt nhân hay không với mức ý nghĩa 2%? Sử dụng $p-$giá trị.

✍ **LỜI GIẢI.**

Mẫu 1: $y_1 = 63$; $n_1 = 100 \Rightarrow \hat{p}_1 = \dfrac{63}{100} = 0.63$.

Mẫu 2: $y_2 = 59$; $n_2 = 125 \Rightarrow \hat{p}_2 = \dfrac{59}{125} = 0.427$.

$$\Rightarrow \hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{63 + 59}{100 + 125} = \frac{122}{225}.$$

☑ Giả thuyết kiểm định: $\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}$; $\quad \Rightarrow \Delta_0 = p_1 - p_2 = 0.$

☑ Giá trị thống kê kiểm định

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2 - \Delta_0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.63 - 0.427 - 0}{\sqrt{\frac{122}{225}\left(1 - \frac{122}{225}\right)\left(\frac{1}{100} + \frac{1}{125}\right)}} \approx 3.03699.$$

☑ $p-$giá trị $= 2\left(1 - \Phi(\lvert 3.03699\rvert)\right) = 2\left(1 - 0.99881\right) = 0.00238$.
Ta có $p-$giá trị $= 0.00238 < 0.02 = \alpha$
$\Rightarrow$ bác bỏ $H_0 : p_1 = p_2$.
Kết luận: Với mức ý nghĩa 2% thì có sự khác biệt có ý nghĩa giữa tỷ lệ cư dân thành thị và ngoại ô trong ủng hộ việc xây dựng nhà máy hạt nhân.

□

BÀI 3.13 (Câu 3 - Đề 1 CKII 21-22). Hai loại giải pháp khác nhau để đánh bóng thấu kính nội nhãn (được dùng trong mắt người sau phẫu thuật đục thủy tinh thể) đang được đánh giá để sử dụng. Trong 300 thấu kính đã được đánh bóng bằng giải pháp 1 thì có 253 thấu kính không có khuyết tật do đánh bóng. Trong 300 thấu kính khác được đánh bóng bằng giải pháp 2 thì có 196 thấu kính không có khuyết tật do đánh bóng. Có lý do nào để tin rằng hai giải pháp đánh bóng là khác nhau không? Sử dụng $\alpha = 0.05$. $p-$giá trị cho kiểm định này là bao nhiêu?

✍ **LỜI GIẢI.**

Gọi Y_1 là số thấu kính được đánh bóng bằng cách sử dụng giải pháp đánh bóng thứ nhất

Y_2 là số thấu kính được đánh bóng bằng cách sử dụng giải pháp đánh bóng thứ hai.

Mẫu 1: $n_1 = 300$; $y_1 = 253 \Rightarrow \hat{p}_1 = \dfrac{y_1}{n_1} = \dfrac{253}{300}$.

Mẫu 2: $n_2 = 300$; $y_2 = 196 \Rightarrow \hat{p}_2 = \dfrac{y_2}{n_2} = \dfrac{196}{300} = \dfrac{49}{75}$.

$$\Rightarrow \hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{253 + 196}{300 + 300} = \frac{449}{600}$$

• **Giả thuyết kiểm định:** $\begin{cases} H_0 : p_1 = p_2 \\ H_1 : p_1 \neq p_2 \end{cases}; \qquad \Rightarrow \Delta_0 = p_1 - p_2 = 0;\ \alpha = 0.05.$

• **Giá trị thống kê kiểm định**

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2 - \Delta_0}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{\dfrac{253}{300} - \dfrac{49}{75} - 0}{\sqrt{\dfrac{449}{600}\left(1 - \dfrac{449}{600}\right)\left(\dfrac{1}{300} + \dfrac{1}{300}\right)}} = 5.3621$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $|z_0| > z_{1-\alpha/2}$.

Ta có $\alpha = 0.05$

$\Rightarrow z_{1-\alpha/2} = z_{0.975} = 1.96$.

• **So sánh và kết luận:**

Ta có: $\quad |z_0| > z_{1-\alpha/2}$

$|5.3621| > z_{1-\alpha/2}$

$5.3621 > 1.96 \quad$ (đúng)

$\Rightarrow$ bác bỏ H_0.

Kết luận: Với mức ý nghĩa 5% thì hai giải pháp đánh bóng có tỷ lệ kính khuyết tật khác nhau.

⁎ p - giá trị:

$z_0 = 5.3621 \Rightarrow p$-giá trị $= 2\left(1 - \Phi(|5.3621|)\right) \approx 2\left(1 - 1\right) = 0$.

□

BÀI 3.14 (Câu 2 - Đề 1 CKI 21-22). Trong một nhà máy sản xuất vỏ bao bì, các kỹ sư áp dụng một phương pháp sản xuất mới với mục đích làm giảm tỷ lệ bao bì hỏng hoặc kém chất lượng. Để kiểm tra hiệu quả của phương pháp sản xuất mới này trước khi đi vào áp dụng chính thức, các kỹ sư so sánh giữa 2 phân xưởng sản xuất: phân xưởng I vẫn áp dụng phương pháp sản xuất cũ và phân xưởng II được thử nghiệm với phương pháp sản xuất mới. Đối với 123 bao bì được sản xuất ở phân xưởng I, có 14 bao bì kém chất lượng hoặc hỏng. Ở phân xưởng II, các kỹ sư thấy có 10 bao bì hỏng/kém chất lượng trong 194 bao bì được sản xuất. Với mức ý nghĩa $\alpha = 0.01$, ta có đủ bằng chứng để kết luận rằng phương pháp sản xuất mới có hiệu quả hơn phương pháp sản xuất cũ không?

✍ **LỜI GIẢI.**

Gọi Y_1 là số bao bì kém chất lượng hoặc bị hỏng ở phân xưởng I;

Y_2 là số bao bì kém chất lượng hoặc bị hỏng ở phân xưởng II.

Mẫu 1: $n_1 = 123$; $y_1 = 14 \Rightarrow \hat{p}_1 = \dfrac{y_1}{n_1} = \dfrac{14}{123}$.

Mẫu 2: $n_2 = 194$; $y_2 = 10 \Rightarrow \hat{p}_2 = \dfrac{y_2}{n_2} = \dfrac{10}{194} = \dfrac{5}{97}$.

$$\Rightarrow \hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{14 + 10}{123 + 194} = \frac{24}{317}$$

• **Giả thuyết kiểm định:** $\begin{cases} H_0 : p_1 \le p_2 \\ H_1 : p_1 > p_2 \end{cases}$; $\quad \Rightarrow \Delta_0 = p_1 - p_2 = 0$; $\alpha = 0.01$.

• **Giá trị thống kê kiểm định**

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2 - \Delta_0}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{\dfrac{14}{123} - \dfrac{5}{97} - 0}{\sqrt{\dfrac{24}{317}\left(1 - \dfrac{24}{317}\right)\left(\dfrac{1}{123} + \dfrac{1}{194}\right)}} = 2.0425.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $z_0 > z_{1-\alpha}$.
Ta có $\alpha = 0.01$
$\Rightarrow z_{1-\alpha} = z_{0.99} = 2.33$.
• **So sánh và kết luận:**
Ta có: $\quad z_0 > z_{1-\alpha}$

$2.0425 > 2.33 \quad$ (sai)

$\Rightarrow$ chưa đủ cơ sở để bác bỏ $H_0 : p_1 \le p_2$.
Kết luận: Với mức ý nghĩa 1% thì ta chưa đủ bằng chứng để kết luận rằng phương pháp sản xuất mới có hiệu quả hơn phương pháp sản xuất cũ. □

BÀI 3.15 (Câu 4 - Đề 2 CKI 21-22). Các tác giả của bài báo "Adjuvant Radiotherapy an Chemotherapy in Node-Positive Premenopausal Women with Breast Cancer" (New Engl. J. of Med., 1997: 956-962) báo cáo các kết quả nghiên cứu của họ về một thí nghiệm được thiết kế để so sánh hiệu quả của 2 phương pháp điều trị bệnh ung thư: chỉ điều trị bằng hóa trị (phương pháp 1) và điều trị kết hợp giữa hóa trị và xạ trị (phương pháp 2). Trong số 183 bệnh nhân được điều trị bằng phương pháp 1, có 90 bệnh nhân sống sót ít nhất 15 năm, trong khi 108 trong số 173 bệnh nhân mà được điều trị bằng phương pháp 2 sống sót ít nhất 15 năm. Có ý kiến cho rằng phương pháp điều trị kết hợp giữa hóa trị và xạ trị có hiệu quả hơn là chỉ điều trị bằng hóa trị. Dựa trên kết quả thí nghiệm, ta có đủ bằng chứng ủng hộ ý kiến trên không với mức ý nghĩa $\alpha = 2\%$?

✍ **LỜI GIẢI.**

Gọi Y_1 là số bệnh nhân sống sót ít nhất 15 năm khi điều trị bằng phương pháp 1;

Y_2 là số bệnh nhân sống sót ít nhất 15 năm khi điều trị bằng phương pháp 2.

Mẫu 1: $n_1 = 183$; $y_1 = 90 \Rightarrow \hat{p}_1 = \dfrac{y_1}{n_1} = \dfrac{90}{183} = \dfrac{30}{61}$.

Mẫu 2: $n_2 = 173$; $y_2 = 108 \Rightarrow \hat{p}_2 = \dfrac{y_2}{n_2} = \dfrac{108}{173}$.

$$\Rightarrow \hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{90 + 108}{183 + 173} = \frac{99}{178}$$

• **Giả thuyết kiểm định:** $\begin{cases} H_0 : p_1 \ge p_2 \\ H_1 : p_1 < p_2 \end{cases}$; $\quad \Rightarrow \Delta_0 = p_1 - p_2 = 0$; $\alpha = 2\% = 0.02$.

• **Giá trị thống kê kiểm định**

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2 - \Delta_0}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} = \frac{\dfrac{30}{61} - \dfrac{108}{173} - 0}{\sqrt{\dfrac{99}{178}\left(1 - \dfrac{99}{178}\right)\left(\dfrac{1}{183} + \dfrac{1}{173}\right)}} = -2.5145.$$

• **Miền bác bỏ:** Bác bỏ H_0 khi $z_0 < -z_{1-\alpha}$.
Ta có $\alpha = 0.02$
$\Rightarrow -z_{1-\alpha} = -z_{0.98} = -2.05$.
• **So sánh và kết luận:**
Ta có: $\quad z_0 < -z_{1-\alpha}$

$$-2.5145 < -2.05 \quad \text{(đúng)}$$

$\Rightarrow$ bác bỏ $H_0 : p_1 \geq p_2$.
Kết luận: Với mức ý nghĩa 2% thì ta có đủ bằng chứng ủng hộ ý kiến cho rằng phương pháp điều trị kết hợp giữa hóa trị và xạ trị có hiệu quả hơn là chỉ điều trị bằng hóa trị. □

BÀI 3.16 (Câu 3 - Đề 2 CKII 21-22 (CNTT hệ CLC)). Các tác giả của bài báo "Adjuvant Radiotherapy an Chemotherapy in Node-Positive Premenopausal Women with Breast Cancer" (New Engl. J. of Med., 1997: 956-962) báo cáo các kết quả nghiên cứu của họ về một thí nghiệm được thiết kế để so sánh hiệu quả của 2 phương pháp điều trị bệnh ung thư: chỉ điều trị bằng hóa trị (phương pháp I) và điều trị kết hợp giữa hóa trị và xạ trị (phương pháp II). Trong số 181 bệnh nhân được điều trị bằng phương pháp I, có 89 bệnh nhân sống sót ít nhất 15 năm, trong khi 107 trong số 186 bệnh nhân mà được điều trị bằng phương pháp II sống sót ít nhất 15 năm.

a) Tìm khoảng tin cậy 95% cho tỷ lệ bệnh nhân sống sót ít nhất 15 năm khi điều trị bằng phương pháp I.

b) Có ý kiến cho rằng phương pháp điều trị kết hợp giữa hóa trị và xạ trị có hiệu quả hơn là chỉ điều trị bằng hóa trị. Dựa trên kết quả thí nghiệm, ta có đủ bằng chứng ủng hộ ý kiến trên không với mức ý nghĩa $\alpha = 1\%$?

BÀI 3.17 (Câu 2 - Đề 1 CKII 20-21). Vào mùa đông của đại dịch cúm, bố mẹ của 2000 bé đã được các nhà nghiên cứu khảo sát tại một công ty dược phẩm nổi tiếng để xác định liệu thuốc mới của công ty có hiệu quả sau hai ngày hay không. Trong số 120 bé bị cúm và được cho dùng thuốc, 29 bé khỏi bệnh trong hai ngày. Trong 280 bé bị cúm những không được cho dùng thuốc mới, có 56 bé phục hồi trong hai ngày. Hỏi có dấu hiệu có ý nghĩa nào ủng hộ lời tuyên bố của công ty về hiệu quả của thuốc hay không với mức ý nghĩa 5%? (*Yêu cầu dùng cả 2 phương pháp: miền bác bỏ và p−giá trị*)

BÀI 3.18 (Câu 1 - Đề 1 CKI 18-19). Gọi X (giờ) là thời gian tự học hàng ngày của sinh viên, khảo sát 120 sinh viên trường Đại học KHXHNV. Kết quả cho bởi bảng sau:

Thời gian tự học (giờ)	1	2	3	4	5	6	7	8
Số sinh viên	13	18	14	23	15	16	17	4

Giả sử thời gian tự học của sinh viên có phân phối chuẩn.

a) Ước lượng thời gian tự học trung bình của sinh viên trường KHXHNV với độ tin cậy 98%. (1.5đ)

b) Với độ tin cậy 95%, ước lượng tỉ lệ những sinh viên có thời gian tự học trên 5 giờ mỗi ngày. (1.5đ)

c) Khảo sát thời gian tự học của 90 sinh viên trường Đại học Kinh tế

Thời gian tự học (giờ)	1	2	3	4	5	6
Số sinh viên	7	8	17	24	20	14

Có ý kiến cho rằng thời gian tự học của sinh viên trường KHXHNV lớn hơn sinh viên trường Kinh tế. Với mức ý nghĩa 5%, hãy kiểm định ý kiến trên. (2đ)

d) So sánh tỷ lệ những sinh viên có thời gian tự học trên 5 giờ mỗi ngày giữa hai trường KHXHNV và Kinh tế. ($\alpha = 1\%$) (2đ)

BÀI 3.19 (Câu 2 - Đề 3 CKI 17-18). Một khảo sát về chiều cao X (cm) của một giống cây trồng người ta quan sát một mẫu và có kết quả như sau:

Chiều cao (cm)	100	110	120	130	140	150	160
Số cây	10	10	15	30	10	10	15

Giả sử chiều cao X có phân phối chuẩn.

a) Ước lượng chiều cao trung bình của giống cây trồng trên với độ tin cậy 95%.

b) Những cây trồng có chiều cao từ 135 cm trở lên được gọi là những cây "cao". Hãy ước lượng tỉ lệ những cây cao với độ tin cậy 95%.

c) Người ta áp dụng phương pháp mới trong việc trồng và chăm sóc cây. Sau một thời gian, khảo sát 100 cây trồng theo phương pháp mới được bảng số liệu sau:

Chiều cao (cm)	100	110	120	130	140	150	160
Số cây	6	10	20	34	12	7	11

Với mức ý nghĩa 5%, hãy kiểm định xem phương pháp mới có làm tăng chiều cao trung bình của cây hay không?

d) Có ý kiến cho rằng phương pháp mới làm tăng tỉ lệ cây "cao". Với mức ý nghĩa 5%, hãy kiểm tra ý kiến này.

BÀI 3.20 (Câu 3 - Đề 4 CKI 17-18). Để xem xét tình hình học tập môn xác suất thống kê (XSTK), một giảng viên (GV) đã tiến hành lấy mẫu (2 lần). Biết rằng điểm số sinh viên tuân theo phân phối chuẩn. Điểm sinh viên nam và sinh viên nữ độc lập nhau.

a) Ở lần lấy mẫu thứ nhất, GV này thu thập được thông tin của 20 sinh viên với điểm trung bình 5.5 và độ lệch chuẩn 0.75. Sử dụng thông tin này để ước lượng điểm trung bình của sinh viên với độ tin cậy 95%.

GV này tiếp tục thực hiện lấy mẫu lần thứ hai và thu thập được bảng thông tin sau:

Điểm	$[0,1)$	$[1,2)$	$[2,3)$	$[3,4)$	$[4,5)$	$[5,6)$	$[6,7)$	$[7,8)$	$[8,9)$	$[9,10]$
Nam	4	6	4	6	2	6	5	2	6	4
Nữ	5	10	8	8	5	4	3	2	3	2

(Bảng dữ liệu chứa thông tin: có 4 sinh viên nam và 5 sinh viên nữ có điểm số nằm trong $[0,1)$, có 6 sinh viên nam và 10 sinh viên nữ có điểm số nằm trong $[1,2)$,.... Hãy sử dụng thông tin ở lần lấy mẫu thứ hai để giải các câu hỏi (b,c,d)

b) Một sinh viên được gọi là có điểm cao nếu điểm lớn hơn hoặc bằng 7.0. Ước lượng khoảng tin cậy cho tỉ lệ học sinh có điểm cao với độ tin cậy 99%.

c) Dựa theo số liệu điểm môn XSTK trong học kì trước, điểm trung bình của sinh viên **nam** là 5.0. Hãy cho biết giá trị trên có phù hợp với dữ liệu quan sát (điểm sinh viên nam đã thu thập được) hay không với mức ý nghĩa 2%?

d) Có nhiều người cho rằng các bạn nữ chăm học hơn các bạn nam nên điểm trung bình của các bạn nữ cao hơn điểm trung bình của các bạn nam. Tuy nhiên, GV này cho rằng điểm trung bình của các bạn nữ **không cao hơn** điểm trung bình của các bạn nam. Hãy sử dụng dữ liệu quan sát để kiểm tra nhận định về hai điểm trung bình của GV với mức ý nghĩa 2%.