Chương 7: KIỂM ĐỊNH GIẢ THIẾT THỐNG KÊ

1. Một số khái niệm:

- **Giả thiết kiểm định H_0** (*Null Hypothesis)*

Giả thiết H_0 là giả thiết về yếu tố cần kiểm định của tổng thể ở trạng thái bình thường, không chịu tác động của các hiện tượng liên quan. Yếu tố trong H_0 phải được xác định cụ thể, ví dụ:

+ H_0 : Tỉ lệ nảy mầm của 1 loại hạt giống là 70%.

+ H_0: Thời gian công nhân hoàn thành 1 sản phẩm là BNN có phân phối chuẩn với kỳ vọng là 20 phút và phương sai là 9 phút^2.

+ H_0: Mức độ yêu thích của khán giả với chương trình truyền hình *"Tìm kiếm tài năng "* không phụ thuộc vào lứa tuổi.

- **Giả thiết đối H_1** *(Alternative Hypothesis)* là một mệnh đề mâu thuẫn với H_0, H_1 thể hiện xu hướng cần kiểm định.

Vì ta sẽ dựa vào thông tin thực nghiệm của mẫu để kết luận xem có thừa nhận các giả thiết nêu trên hay không nên công việc này gọi là *kiểm định thống kê*.

- **Tiêu chuẩn kiểm định** là hàm thống kê $G = G(X_1, X_2, .., X_n, \theta_0)$, xây dựng trên mẫu ngẫu nhiên $W = (X_1, X_2, .., X_n)$ và tham số θ_0 liên quan đến H_0; Điều kiện đặt ra với thống kê G là nếu H_0 đúng thì quy luật phân phối xác suất của G phải hoàn toàn xác định.

- **Miền bác bỏ giả thiết W_α** là miền thỏa $P(G \in W_\alpha / H_0 \text{ đúng}) = \alpha$. α là một số khá bé, thường không quá 0,05 và gọi là ***mức ý nghĩa*** của kiểm định. Có vô số miền W_α như vậy.

- Quy tắc kiểm định: Từ mẫu thực nghiệm, ta tính được một giá trị cụ thể của tiêu chuẩn kiểm định là thống kê $g_{qs} = G(x_1, x_2, .., x_n, \theta_0)$. Theo nguyên lý xác suất bé, biến cố $G \in W_\alpha$ có xác suất nhỏ nên với 1 mẫu thực nghiệm, nó không thể xảy ra. Do đó:

+ Nếu $g_{qs} \in W_\alpha$ thì bác bỏ H_0, thừa nhận giả thiết H_1.

+ Nếu $g_{qs} \notin W_\alpha$: ta chưa đủ dữ liệu khẳng định H_0 sai. Ta nói "có thể chấp nhận H_0" hay "không bác bỏ H_0".

Kết luận của một bài toán kiểm định có thể mắc các sai lầm sau:

- **Sai lầm loại I**: Bác bỏ giả thiết H_0 trong khi H_0 đúng. Xác suất mắc phải sai lầm này nếu H_0 đúng chính bằng mức ý nghĩa α. Nguyên nhân mắc phải sai lầm loại I thường có thể do kích thước mẫu quá nhỏ, có thể do phương pháp lấy mẫu ...
- **Sai lầm loại II**: Thừa nhận H_0 trong khi H_0 sai, tức là mặc dù thực tế H_1 đúng nhưng giá trị thực nghiệm g_{qs} không thuộc W_α.

Tình huống / Quyết định	**H_0 đúng**	**H_0 sai**
Bác bỏ H_0	Sai lầm loại I. Xác suất = α	Quyết định đúng.
Không bác bỏ H_0	Quyết định đúng.	Sai lầm loại II. Xác suất =β

Ví dụ: Người bán hàng nói rằng tỉ lệ phế phẩm trong mỗi lô hàng không quá 5%. Người mua quyết đinh kiểm ngẫu nhiên 10 sản phẩm, nếu được cả 10 sản phẩm tốt thì mới mua lô hàng. Sai lầm loại I xảy ra khi người mua từ chối mua hàng trong khi thực sự lô hàng có không quá 5% phế phẩm; α là mức rủi ro cho bên bán. Sai lầm loại II xảy ra khi người mua nhận hàng nhưng tỉ lệ phế phẩm thực ra trên 5%; β chính là mức rủi ro cho bên mua.

Với một mẫu xác định, khi ta giảm α đi thì đồng thời sẽ làm tăng β và ngược lại. Chỉ có thể cùng giảm α, β nếu tăng kích thước mẫu. Người ta thường có xu hướng coi trọng xác suất mắc sai lầm loại I nên sẽ hạn chế trước giá trị α tùy thực tế, và sau đó phải tìm miền W_α sao cho xác suất mắc sai lầm loại II là nhỏ nhất. Miền W_α thỏa yêu cầu này được gọi là miền bác bỏ ***tốt nhất*** dựa trên các cơ sở toán học chặt chẽ.

Đặt các giả thiết H_0; H_1

Tìm tiêu chuẩn kđ G

Tìm miền bác bỏ W_α tốt nhất

Mẫu thực nghiệm

$\Rightarrow$ Tính g_{qs}

Kiểm tra $g_{qs} \overset{?}{\in} W_\alpha$

$g_{qs} \in W_\alpha$: Bác bỏ H_0, chấp nhận H_1

$g_{qs} \notin W_\alpha$: Chưa bác bỏ được H_o

(tạm chấp nhận H_0)

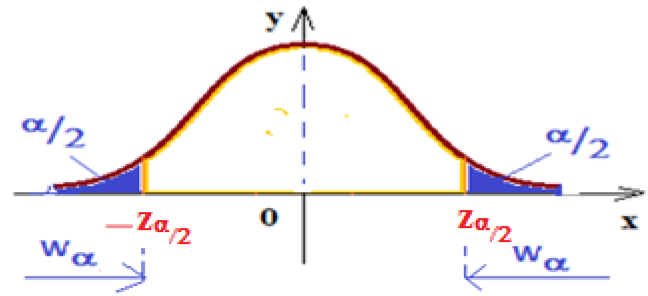
Quy trình thực hiện bài toán kiểm định.

Ví dụ minh họa cho các miền bác bỏ khi tiêu chuẩn kiểm định Z có phân phối chuẩn N(0,1).

1. Miền bác bỏ 2 phía:

$W_{\alpha} = (-\infty, -Z_{\alpha/2}) \cup (Z_{\alpha/2}, +\infty)$

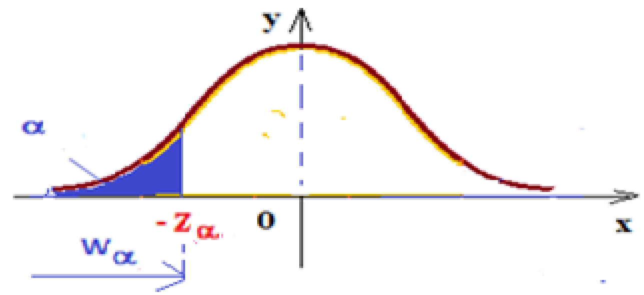
ở đây $F(z_{\alpha/2}) = 1 - \dfrac{\alpha}{2}$



2. Miền bác bỏ bên trái:

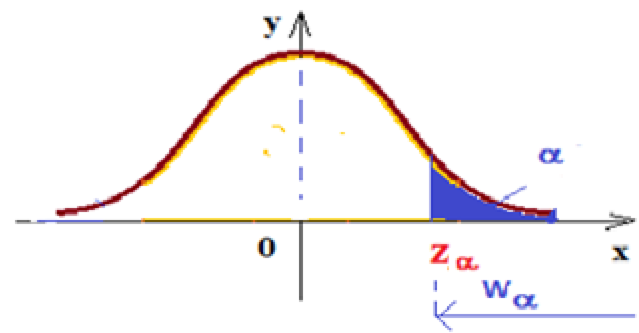
$W_{\alpha} = (-\infty, -Z_{\alpha})$

ở đây $F(z_{\alpha}) = 1 - \alpha$



3. Miền bác bỏ bên phải:

$W_{\alpha} = (Z_{\alpha}, +\infty)$



2. Bài toán kiểm định tham số:

2.1 Bài toán kiểm định tỉ lệ:

Bảng 3: Tóm tắt một số công thức của bài toán kiểm định tỉ lệ

	Giả thiết KĐ H_0	Giả thiết đối H_1	Tiêu chuẩn kiểm định	Miền bác bỏ H_0 với mức ý nghĩa α
BT 1 mẫu *$n \geq 30$*	$p = p_0$	$p \neq p_0$	$Z_{qs} = \dfrac{F - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$ $Z_{qs} \sim N(0; 1)$	$W_\alpha = (-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty)$
		$p < p_0$		$W_\alpha = (-\infty, -z_\alpha)$
		$p > p_0$		$W_\alpha = (z_\alpha, +\infty)$
BT 2 mẫu *$n_1 \geq 30$* *$n_2 \geq 30$*	$p_1 = p_2$	$p_1 \neq p_2$	$Z_{qs} = \dfrac{F_1 - F_2}{\sqrt{\overline{f}(1-\overline{f})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$; mẫu gộp: $\overline{f} = \dfrac{n_1 F_1 + n_2 F_2}{n_1 + n_2}$ $Z_{qs} \sim N(0; 1)$	$W_\alpha = (-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty)$
		$p_1 < p_2$		$W_\alpha = (-\infty, -z_\alpha)$
		$p_1 > p_2$		$W_\alpha = (z_\alpha, +\infty)$

Ở BT 2 mẫu, khi dùng mẫu cụ thể $f_1 = \dfrac{m_1}{n_1}; f_2 = \dfrac{m_2}{n_2} \Rightarrow \overline{f} = \dfrac{m_1 + m_2}{n_1 + n_2}$

<u>Ví dụ 11</u>**:** Theo số liệu công bố của một công ty dịch vụ tin học, tỷ lệ khách hàng hài lòng với dịch vụ của công ty là 85%. Một khảo sát độc lập cho thấy trong mẫu gồm 145 khách hàng của công ty có 120 khách hàng hài lòng. Với mức ý nghĩa 3%, có thể coi số liệu của công ty là đáng tin cậy không?

<u>Hướng dẫn:</u> Gọi p là tỉ lệ khách hàng hài lòng với dịch vụ của CT.

GtKđ H_0: $p = 85\%$

Giả thiết đối H_1: $p \neq 85\%$

+ Mức ý nghĩa $\alpha = 3\% \Rightarrow F(z_{\alpha/2}) = 1 - 0.03/2 = 0.985 \Rightarrow z_{\alpha/2} = 2.17$

Miền b/bỏ $W_\alpha = (-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty) = (-\infty; -2.17) \cup (2.17; +\infty)$

Kích thước mẫu: $n = 145$; Tỉ lệ mẫu: $f = 120/145 = 0.8276$

+ Tiêu chuẩn kđ:

$$Z_{qs} = \frac{f - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} = \frac{0.8276 - 0.85}{\sqrt{0.85(1-0.85)}}\sqrt{145} = -0.7559$$

Do $Z_{qs} \notin W_\alpha$ nên ta chưa đủ dữ kiện bác bỏ H_0.

Có thể tạm xem như số liệu của công ty là đáng tin.

Ví dụ 12: Theo tiêu chuẩn của công ty thì một lô hàng nguyên liệu được chấp nhận nếu không có quá 3% phế phẩm. Kiểm tra ngẫu nhiên 400 sản phẩm từ lô hàng này thì thấy 16 phế phẩm. Với mức ý nghĩa 5%, hãy xem xét lô hàng này có thể được chấp nhận không?

Hướng dẫn: + Gọi p là tỉ lệ phế phẩm thực sự của lô hàng.

GtKđ H_0: $p = 3\%$ (hay $p \leq 3\%$)

Giả thiết đối H_1: $p > 3\%$

+ Myn $\alpha = 5\% \Rightarrow F(z_\alpha) = 1- 0.05 = 0.95 \Rightarrow z_\alpha = 1.645$

Miền bác bỏ $W_\alpha = (z_\alpha ;+\infty) = (1.645 ;+\infty)$

Kích thước mẫu: n = 400; Tỉ lệ mẫu: f = 16/400 = 0.04.

+ TC kiểm định: $$Z_{qs} = \frac{f - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} = \frac{0.04-0.03}{\sqrt{0.03(1-0.03)}}\sqrt{400} = 1.172$$

Do $Z_{qs} \notin W_\alpha$ nên ta chưa bác bỏ H_0, tức là chưa thể kết luận tỉ lệ phế phẩm của lô hàng vượt ngưỡng cho phép.

Ví dụ 13: Tỉ lệ bệnh nhân bị bệnh T được chữa khỏi bệnh bằng thuốc A là 85%. Khi dùng thuốc B điều trị thì trong 1100 bệnh nhân bị bệnh T người ta thấy có 903 người khỏi bệnh. Có thể nói rằng thuốc B điều trị ít hiệu quả hơn thuốc A được không, kết luận với mức ý nghĩa 4%?

Hướng dẫn: + Gọi p là tỉ lệ BN khỏi bệnh khi dùng thuốc B.

GtKđ H_0: $p = 85\%$

Giả thiết đối H_1: $p < 85\%$

+ Myn $\alpha = 4\% \Rightarrow F(z_\alpha) = 1- 0,04 = 0,96 \Rightarrow z_\alpha = 1.75$

Miền bác bỏ $W_\alpha = (-\infty; -z_\alpha) = (-\infty; -1.75)$

n = 1100;

+ Tckđ :

$$Z_{qs} = \frac{f - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n} = \frac{\frac{903}{1100} - 0.85}{\sqrt{0.85(1-0.85)}}\sqrt{1100} = -2.7021$$

+ Do $Z_{qs} \in W_\alpha$ nên bác bỏ H_0, chấp nhận H_1. Xem như tỉ lệ BN khỏi bệnh khi dùng thuốc B là thấp hơn so với dùng thuốc A.

Ví dụ 14: Khảo sát ngẫu nhiên 80 sinh viên nam thấy có 56 bạn thường xuyên đi xe buýt; trong 60 SV nữ thì con số này là 48. Có thể coi như tỷ lệ SV nam đi xe buýt thường xuyên là thấp hơn so với SV nữ hay không? Hãy kết luận với mức ý nghĩa 5%?

Hướng dẫn: + Gọi p_1; p_2 lần lượt là tỉ lệ SV nam & nữ đi xe buýt tx.

GtKđ H_0: $p_1 = p_2$; Giả thiết đối H_1: $p_1 < p_2$

+ Myn $\alpha = 5\% \Rightarrow F(z_\alpha) = 1 - 0,05 = 0,95 \Rightarrow z_\alpha = 1.645$

Miền bác bỏ $W_\alpha = (-\infty; -1.645)$

$n_1 = 80$; $f_1 = 56/80$; $n_2 = 60$; $f_2 = 48/60$; $\overline{f} = (56+48)/(60+80)$

+ Tiêu chuẩn kđ:

$$Z_{qs} = \frac{f_1 - f_2}{\sqrt{\overline{f}(1-\overline{f})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{\frac{56}{80}-\frac{48}{60}}{\sqrt{\frac{104}{140}\left(1-\frac{104}{140}\right)\left(\frac{1}{80}+\frac{1}{60}\right)}} = -1.3397$$

+ Do $Z_{qs} \notin W_\alpha$ nên chưa bác bỏ được H_0. Xem như tỉ lệ sinh viên nam thường xuyên đi xe buýt không thấp hơn so với SV nữ.

<u>Ví dụ 15:</u> Tỉ lệ phế phẩm của 1 nhà máy là 10%. Sau khi cải tiến quy trình sản xuất, người ta kiểm tra thử 250 sản phẩm thì thấy có 17 phế phẩm.

<u>Bài toán a)</u> Với mức ý nghĩa 5%, hãy cho biết có thể coi như việc cải tiến quy trình sản xuất đã làm thay đổi tỷ lệ phế phẩm của nhà máy không?

<u>Bài toán b)</u> Với mức ý nghĩa 5%, có thể cho rằng việc cải tiến quy trình sản xuất đã có hiệu quả hay không?

<u>Ví dụ 16:</u>

Người ta bảo quản cùng 1 loại hạt giống theo 2 phương pháp khác nhau trong thời gian như nhau. Gieo thử ngẫu nhiên 500 hạt giống đã được bảo quản theo phương pháp I thì thấy có 450 hạt nảy mầm; gieo thử 700 hạt giống đã được bảo quản theo phương pháp II thì có 600 hạt nảy mầm. Với myn 2%, có thể xem như phương pháp I có hiệu quả hơn phương pháp II hay không?

2.2 Bài toán kiểm định trung bình:

2.2.1 Bài toán 1 mẫu:

<table>
<tr><td></td><td>GT KĐ H_0</td><td>GT đối H_1</td><td>Tiêu chuẩn kiểm định</td><td colspan="2">Miền bác bỏ H_0 với mức ý nghĩa α</td></tr>
<tr><td rowspan="4">BT 1 mẫu</td><td></td><td></td><td rowspan="4">- Nếu đã biết σ^2:
$Z_{qs} = \frac{\overline{X} - a_0}{\sigma}\sqrt{n}$
- Nếu chưa biết σ^2:
$Z_{qs} = \frac{\overline{X} - a_0}{s}\sqrt{n}$</td><td>- Tổng thể phân phối chuẩn, đã biết σ^2.
- Hoặc tổng thể tùy ý, kích thước mẫu $n \geq 30$</td><td>- Tổng thể phân phối chuẩn; chưa biết σ^2 và $n < 30$.
(Nếu $n > 30$, có thể dùng W_α như trường hợp bên.</td></tr>
<tr><td rowspan="3">$a = a_0$</td><td>$a \neq a_0$</td><td>$W_\alpha = (-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty)$</td><td>$W_\alpha = (-\infty, -t_{\frac{\alpha}{2}}(n-1)) \cup (t_{\frac{\alpha}{2}}(n-1), +\infty)$</td></tr>
<tr><td>$a < a_0$</td><td>$W_\alpha = (-\infty, -z_\alpha)$</td><td>$W_\alpha = (-\infty, -t_\alpha^{(n-1)})$</td></tr>
<tr><td>$a > a_0$</td><td>$W_\alpha = (z_\alpha, +\infty)$</td><td>$W_\alpha = (t_\alpha^{(n-1)}, +\infty)$</td></tr>
</table>

Ví dụ 17: Một công ty sản xuất phomát nghi ngờ một nhà cung cấp sữa cho công ty đã pha thêm nước vào sữa để làm tăng lượng sữa cung cấp. Nếu sữa có pha nhiều nước quá mức bình thường thì nhiệt độ đông của nó sẽ thấp hơn so với sữa tự nhiên. Biết rằng điểm đông của sữa tự nhiên tuân theo quy luật phân phối chuẩn với trung bình khoảng -0.545^{0}C, độ lệch chuẩn 0.008^{0}C. Người ta kiểm định chất lượng sữa trong các container hàng mới nhập bằng cách lấy ra 25 mẫu ngẫu nhiên thì thấy nhiệt độ đông trung bình của sữa trong mẫu là -0.55^{0}C. Hãy kết luận về chất lượng sữa mà công ty mua với mức ý nghĩa 1%.

Hướng dẫn:

+ Gọi a là nhiệt độ đông trung bình của lượng sữa mới nhập.

GtKđ H_0: $a = -0.545^0C$; Giả thiết đối H_1: $a < -0.545^0C$

$n = 25 < 30$; $\sigma = 0.008$ (đã biết); $a_0 = -0.545^0C$. $\overline{x} = -0.55^0C$

+ Myn $\alpha = 1\% \Rightarrow \Phi(z_{2\alpha}) = 1 - 0,01 = 0,99 \Rightarrow z_\alpha = 2.33$

Miền bác bỏ $W_\alpha = (-\infty; -2.33)$

+ T.chuẩn kđ: $$Z_{qs} = \frac{\overline{x} - a_0}{\sigma}\sqrt{n} = \frac{(-0.55) - (-0.545)}{0.008}\sqrt{25} = -3.125$$

+ Do $Z_{qs} \in W_\alpha$ nên bác bỏ H_0, chấp nhận H_1. Ta kết luận lượng sữa công ty mới mua đã bị pha nước.

<u>Ví dụ 18</u> Người ta đã thực hiện một cải tiến kỹ thuật trong bộ hòa khí của xe ôtô với hy vọng sẽ tiết kiệm được xăng hơn. Cho xe chạy thử 12 lần thì họ có số km chạy được cho 1 lít xăng:

20.6	20.5	20.8	20.8	20.7	20.6
21	20.6	20.5	20.4	20.3	20.7

Nếu trước khi cải tiến, 1 lít xăng trung bình chạy được 20.4 km thì với số liệu này người ta đã có thể kết luận việc cải tiến mang lại hiệu quả đáng kể hay không, với mức ý nghĩa 5% ?

Giả thiết quãng đường xe chạy được khi tiêu thụ 1 lít xăng là biến ngẫu nhiên có phân phối chuẩn.

Hướng dẫn:

+ Gọi a là quãng đường trung bình ôtô chạy được với 1 lít xăng sau khi cải tiến kỹ thuật.

n = 12 <30 ; $\overline{x}$ = 20.625 a_0= 20.4 s = 0.1913

Giả thiết kiểm định H_0: a = 20.4

Giả thiết đối H_1: a > 20.4

+ Myn α = 5% $\Rightarrow$ Tra bảng Student 1 phía: $t_\alpha(n-1) = t_{0.05}(11) = 1.796$

Miền bác bỏ W_α = $(1.796; +\infty)$

+ Tiêu chuẩn kiểm định:

$$Z_{qs} = \frac{\overline{x} - a_0}{s}\sqrt{n} = \frac{20.625 - 20.4}{0.1913}\sqrt{12} = 4.0743$$

+ Do $Z_{qs} \in W_\alpha$ nên bác bỏ H_0, chấp nhận H_1.

Việc cải tiến kỹ thuật đã có hiệu quả.

Ví dụ 19 Ở một phân xưởng, người ta định mức thời gian gia công 1 chi tiết cho mỗi công nhân là 12 phút. Sau khi thay đổi nguyên liệu, người ta khảo sát ngẫu nhiên quá trình gia công của 1 số chi tiết và thu được số liệu dưới đây. Với myn 5%, hãy quyết định xem có cần thay đổi định mức gia công hay không?

Thời gian gia công 1 chi tiết (phút)	10-10.5	10.5-11	11-11.5	11.5-12	12-12.5	12.5-13	13-13.5
Số chi tiết t/ư	4	12	26	37	43	28	10

Hướng dẫn:

+ Gọi a là thời gian gia công TB 1 chi tiết ở thời điểm hiện tại.

+ GTKĐ H_0: a = 12 phút. GTĐ H_1: a ≠ 12 phút

+ Miền bác bỏ $W_\alpha = (-\infty; -1.96) \cup (1.96; +\infty)$

+ Tckđ: $$Z_{qs} = \frac{\overline{x} - a_0}{s}\sqrt{n} = \frac{11.9594 - 12}{0.7170}\sqrt{160} = -0.7163$$

+ Do $Z_{qs} \notin W_\alpha$ nên chưa bác bỏ được H_0. Vì vậy chưa cần thay đổi định mức.

2.2.2 Bài toán 2 mẫu: (kđ so sánh trung bình 2 tổng thể)

	Dạng bài	Gt Ho	Gt H1	Miền bác bỏ W_α	Tiêu chuẩn kiểm định	
1	- X1,X2 có phân phối Chuẩn. - Biết phương sai tổng thể $\sigma_1^2; \sigma_2^2$. - 2 mẫu độc lập. **z- test**	$a_1 = a_2$	$a_1 \neq a_2$	$(-\infty; -Z_\alpha) \cup (Z_\alpha; +\infty)$	$Z_{qs} = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0;1)$	
			$a_1 < a_2$	$(-\infty; -Z_{2\alpha})$		
			$a_1 > a_2$	$(Z_{2\alpha}; +\infty)$		
2	- X1,X2 có phân phối Chuẩn. - Chưa biết $\sigma_1^2; \sigma_2^2$ giả thiết $\sigma_1^2 \neq \sigma_2^2$. - 2 mẫu độc lập. **t- test**	$a_1 = a_2$	$a_1 \neq a_2$	$(-\infty; \ -t_{\frac{\alpha}{2}}(v)) \cup (t_{\frac{\alpha}{2}}(v); \ +\infty)$	$T_{qs} = \dfrac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \sim T(v)$ $v = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$	Trường hợp X1, X2 phân phối bất kỳ & $n_1, n_2 > 30$, xem như Tqs ~ N(0, 1) và xđ miền bác bỏ như sau: * 2 phía: $(-\infty; -Z_\alpha) \cup (Z_\alpha; +\infty)$ * Bên trái: $(-\infty; -Z_{2\alpha})$ * Bên phải: $(Z_{2\alpha}; +\infty)$
			$a_1 < a_2$	$(-\infty; \ -t_\alpha(v))$		
			$a_1 > a_2$	$(t_\alpha(v); \ +\infty)$; $v \in N^+$.		

Trong trường hợp tổng quát thì Ho: $a_1 = a_2 + d_0$.

Khi đó TCKĐ thay đổi tương ứng. VD trong dạng 1: $Z_{qs} = \left(\overline{X}_1 - \overline{X}_2 - d_0\right) \Big/ \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

	Dạng bài	Gt Ho	Gt H1	Miền bác bỏ W_α	Tiêu chuẩn kiểm định
3	- X1, X2 có phân phối Chuẩn. - Chưa biết $\sigma_1^2;\sigma_2^2$ giả thiết $\sigma_1^2=\sigma_2^2$. - 2 mẫu độc lập. **t- test**	$a_1=a_2$	$a_1 \neq a_2$	$(-\infty;\ -t_{\frac{\alpha}{2}}(n_1+n_2-2)) \cup (t_{\frac{\alpha}{2}}(n_1+n_2-2);\ +\infty)$	$T_{qs}=\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{\dfrac{S_p^2}{n_1}+\dfrac{S_p^2}{n_2}}} \sim T(n_1+n_2-2)$ ở đây phương sai gộp: $S_p^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$
			$a_1 < a_2$	$(-\infty;\ -t_\alpha(n_1+n_2-2))$	
			$a_1 > a_2$	$(t_\alpha(n_1+n_2-2);\ +\infty)$	
4	- X1, X2 có phân phối Chuẩn. - Chưa biết $\sigma_1^2;\sigma_2^2$. - **2 mẫu phụ thuộc tương ứng theo cặp**, có cùng kích thước n - Đặt D=X1-X2 **t- test**	$a_1=a_2$ hay $a_D=0$	$a_D \neq 0$	$(-\infty;\ -t_{\frac{\alpha}{2}}(n-1)) \cup (t_{\frac{\alpha}{2}}(n-1);\ +\infty)$	$T_{qs}=\dfrac{\overline{D}}{S_D}\sqrt{n}$ $T_{qs} \sim T(n-1)$
			$a_D < 0$	$(-\infty;\ -t_\alpha(n-1))$	
			$a_D > 0$	$(t_\alpha(n-1);\ +\infty)$	

Khi n > 30 và bảng tra Student không đủ số liệu, ta dùng bảng tra của phân phối chuẩn để thay thế.

Ví dụ 20: Người ta trồng cùng 1 giống lúa trên 2 thửa rộng như nhau và bón 2 loại phân khác nhau, đến ngày thu hoạch họ lấy mẫu trên 2 thửa ruộng và có kết quả khảo sát như sau:

	Số bông được khảo sát	Số hạt trung bình trên 1 bông	Độ lệch mẫu hiệu chỉnh
Thửa ruộng 1	1000	70	10
Thửa ruộng 2	500	72	20

Với mức ý nghĩa 5%, hãy kết luận xem sự khác nhau giữa 2 trung bình mẫu là ngẫu nhiên hay bản chất?

Hướng dẫn: + Gọi a_1; a_2 là số hạt lúa TB trên 1 bông ở mỗi thửa.

+ GTKĐ H_0: $a_1 = a_2$. GTĐ H_1: $a_1 \neq a_2$

+ Miền bác bỏ $W_\alpha = (-\infty; -1.96) \cup (1.96; +\infty)$

+ Tckđ: $$Z_{qs} = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{70-72}{\sqrt{\frac{10^2}{1000} + \frac{20^2}{500}}} = 2.1082$$

+ Do $Z_{qs} \in W_\alpha \Rightarrow$ bác bỏ H_0. Số hạt TB trên 1 bông ở 2 thửa ruộng là khác nhau, hay sự khác nhau giữa 2 TB mẫu là do bản chất.

Ví dụ 21:

Khảo sát thu nhập (đơn vị: triệu đồng) trong 3 tháng đầu năm của các công nhân trong 2 nhà máy có điều kiện làm việc như nhau, người ta có được kết quả:

Nhà máy 1	**18.5**	**19**	**19.3**	**20**	**20.2**	**21**	**21.5**	**19**	**19.7**	**20**
Nhà máy 2	**17.3**	**18**	**19**	**20**	**20.6**	**20.9**	**18.2**	**19.6**	**20.8**	

Với mức ý nghĩa 5%, có thể cho rằng thu nhập trung bình của công nhân 2 nhà máy đó trong 3 tháng đầu năm là như nhau hay không, biết thu nhập của công nhân ở 2 nhà máy có phân phối chuẩn và có phương sai bằng nhau.

Hướng dẫn:

Đây là bài toán t-test với giả thiết 2 phương sai tổng thể như nhau.

Gọi a_1; a_2 là thu nhập trung bình 3 tháng đầu năm của công nhân 2 nhà máy.

Giả thiết kiểm định H_o: $a_1 = a_2$; H_1: $a_1 \neq a_2$

$n_1 = 10$ $\overline{x_1} = 19.82$ $s_1^2 = 0.8662$

$n_2 = 9$ $\overline{x_2} = 19.3777$ $s_2^2 = 1.7519$

+ Miền bác bỏ $W_\alpha = (-\infty; -t_{\alpha/2}(n_1+n_2-2)) \cup (t_{\alpha/2}(n_1+n_2-2); +\infty)$

$= (-\infty; -2.1098) \cup (2.1098; +\infty)$

+ Do giả thiết phương sai 2 tổng thể chưa biết và $\sigma_1^2 = \sigma_2^2$, nên ta cần tính thêm phương sai gộp:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = 1.2830$$

suy ra tiêu chuẩn kiểm định:

$$T_{qs} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} = 0.8497 \notin W\alpha$$

+ Có thể chấp nhận H_0. Thu nhập của CN 2 nhà máy là như nhau.

<u>Ví dụ 22 :</u>

Tại một xí nghiệp, người ta xây dựng 2 phương án gia công cùng một loại chi tiết. Để đánh giá xem chi phí trung bình về nguyên liệu theo 2 phương án ấy có khác nhau hay không, người ta tiến hành sản xuất thử và thu được kết quả sau:

Phương án 1:	2.5	3.2	3.5	3.8	3.5	
Phương án 2:	2.0	2.7	2.5	2.9	2.3	2.6

Hãy cho kết luận về vấn đề trên biết rằng chi phí nguyên liệu theo cả 2 phương án gia công đều là biến ngẫu nhiên phân phối chuẩn với $\sigma_1^2 = \sigma_1^2 = 0.16$. Xét mức ý nghĩa 2%.

Ví dụ 23:

Khảo sát chi phí (triệu đồng) để hoàn thành 1 đồ án tốt nghiệp của từng sinh viên năm 4, ta thu được số liệu mẫu:

Khoa Cơ khí: **4,6 4,7 4,3 4,5 4,4 4,9**
4,9 4,6 4,4 4,7 4,8 4,5

Khoa Máy tính:

Chi phí cho 1 đề tài	3,5 - 4	4 - 4,5	4,5 - 5	5 - 5,5	5,5 - 6
Số SV tương ứng	4	14	16	12	4

a) Theo dự báo của VPK Cơ khí thì chi phí trung bình làm ĐATN của 1 sv là 4,5tr. Với myn 2,5%, hãy cho biết số liệu của VPK có thể xem như thấp hơ n số liệu khảo sát thực tế hay ko?

b) Hãy so sánh chi phí trung bình một SV làm ĐATN giữa 2 khoa, kết luận với myn 5%. Giả thiết các mức chi phí đều tuân theo pp chuẩn .

Ví dụ 24 :

Để so sánh độ bền của 2 loại sơn phản quang trong giao thông, người ta kẻ 12 lần sơn mỗi loại trên một đoạn đường có nhiều xe lưu thông, thứ tự sơn được chọn 1 cách ngẫu nhiên.

Sau 1 thời gian, người ta dùng máy đo cường độ phản chiếu của các lần sơn (chỉ số đọc càng cao thì cường độ phản chiếu càng lớn) và ghi lại được các số liệu sau đây:

Sơn A:	12.5	11.7	9.9	9.6	10.3	9.6
	9.4	11.3	8.7	11.5	10.6	9.7
Sơn B:	9.4	11.6	9.7	10.4	6.9	7.3
	8.4	7.2	7.0	8.2	12.7	9.2

Người ta cho rằng loại sơn A bền hơn loại sơn B. Hãy kiểm định thông tin này với mức ý nghĩa 1%.

Ví dụ 25 :

Để so sánh tốc độ xử lý của 2 phần mềm thống kê A và B, người ta chọn 1 bộ dữ liệu và đặt ra 10 yêu cầu cần xử lý trên bộ dữ liệu này theo thứ tự.

Dưới đây là số liệu thu được về thời gian xử lý từng lệnh của mỗi phần mềm trên bộ dữ liệu này:

Hãy cho biết có thể coi phần mềm B xử lý nhanh hơn phần mềm A hay không, kết luận với myn 5%?

Giả thiết thời gian xử lý của 2 phần mềm đều tuân theo phân phối chuẩn.

Lệnh	Thời gian xử lý (giây)	
	Phần mềm A	Phần mềm B
1	9.98	9.88
2	9.88	9.86
3	9.84	9.75
4	9.99	9.8
5	9.94	9.87
6	9.84	9.84
7	9.86	9.87
8	10.12	9.86
9	9.9	9.83
10	9.91	9.86

3. PHÂN TÍCH PHƯƠNG SAI (ANOVA)

Phân tích phương sai là một mô hình để xem xét sự biến động của một biến ngẫu nhiên định lượng X chịu tác động trực tiếp của một hay nhiều yếu tố nguyên nhân (định tính).

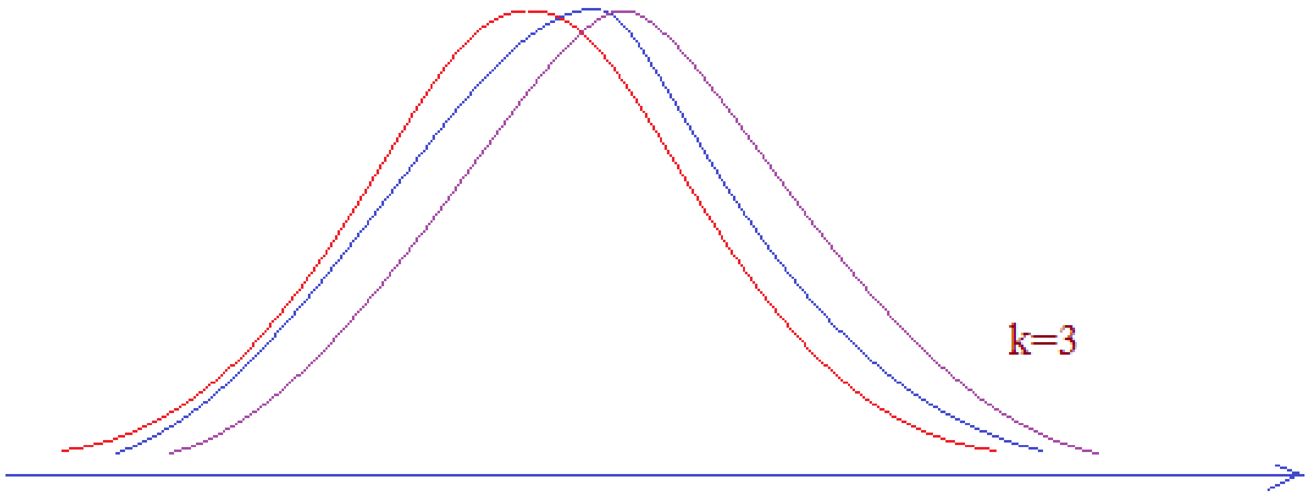
- *Dạng 1: Phân tích phương sai 1 yếu tố*
- *Dạng 2: Phân tích phương sai 2 yếu tố không lặp (BTL)*
- *Dạng 3: Phân tích phương sai 2 yếu tố có lặp (BTL).*

Trong mô hình phân tích phương sai 1 yếu tố, chúng ta kiểm định so sánh trung bình của biến ngẫu nhiên X ở những nhóm khác nhau dựa vào các mẫu quan sát lấy từ những nhóm này. (Các nhóm được phân biệt theo các mức của yếu tố đang xem xét. Bài toán còn được gọi là bài toán so sánh trung bình các tổng thể).

Giả thiết của bài toán: *(điều kiện bài toán)*

- Các tổng thể có phân phối chuẩn

  (với trung bình tương ứng là a_1; a_2; ...; a_k chưa biết).
- Các tổng thể có phương sai bằng nhau.
- Các mẫu quan sát được lấy độc lập.

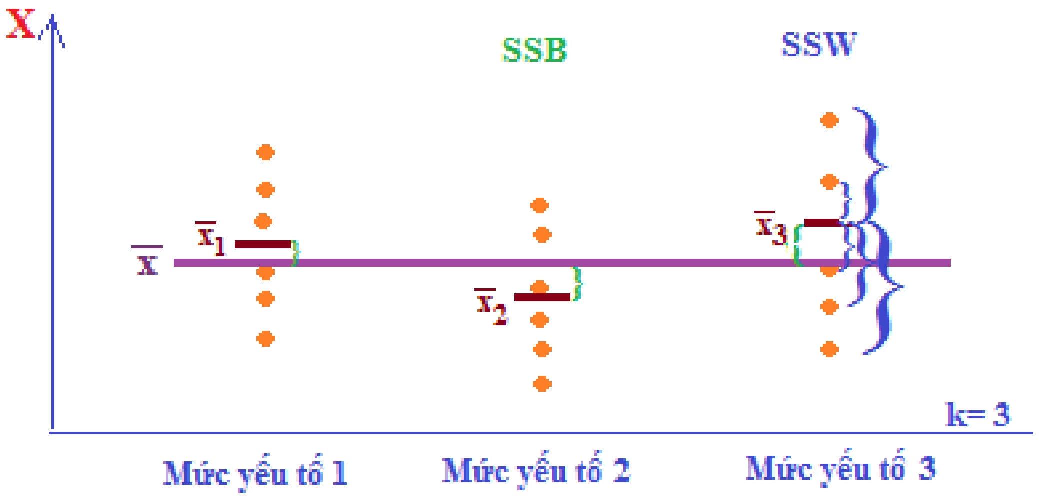


Giả thiết kiểm định H_0 **: $a_1 = a_2 = \ldots = a_k$.**

Giả thiết đối H_1 **: $\exists\, a_i \neq a_j$; với $i \neq j$**

<u>Các ký hiệu cho bảng tính phía sau:</u>

- **SSB** : Sum of squares between group (hay SSG)
- **SSW** : Sum of squares within group.
- **SST**: Total sum of squares.



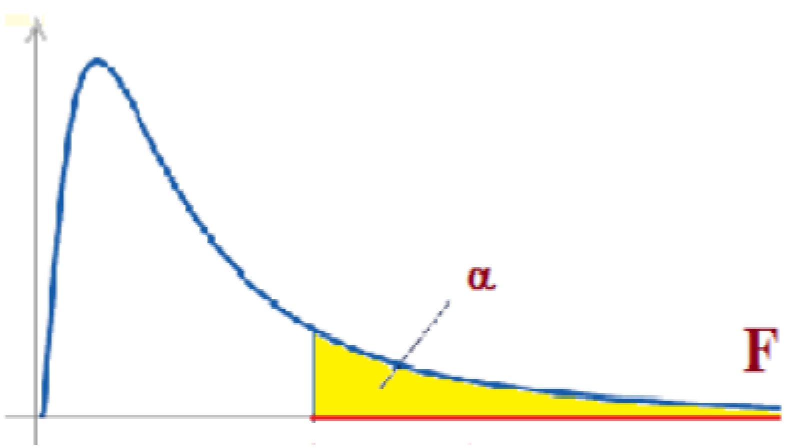
Các mức yếu tố / Các quan sát trong nhóm	1	2	....	k	Mẫu gộp
1 2 ..	(x_{i1})	(x_{i2})		(x_{ik})	
(1) Kích thước mẫu	n_1	n_2	...	n_k	$N = n_1 + n_2 + \ldots + n_k$
(2) Trung bình mẫu của từng nhóm	$\overline{x}_1$	$\overline{x}_2$	...	$\overline{x}_k$	$\overline{x} = \left(\sum_{j=1}^{k} n_j \times \overline{x}_j \right) / N$
(3) Tổng bình phương chênh lệch giữa các nhóm	$n_1 \times \left(\overline{x}_1 - \overline{x} \right)^2$	$n_2 \times \left(\overline{x}_2 - \overline{x} \right)^2$	...	$n_k \times \left(\overline{x}_k - \overline{x} \right)^2$	$SSB = \sum_{j=1}^{k} n_j \times \left(\overline{x}_j - \overline{x} \right)^2$
(4) Tổng bình phương chênh lệch trong nội bộ nhóm	$\sum_{i=1}^{n_1} \left(x_{i1} - \overline{x}_1 \right)^2$	*Bấm* $n \times \widehat{s}^{2}$	...	$\sum_{i=1}^{n_k} \left(x_{ik} - \overline{x}_k \right)^2$	$SSW = \sum_{j=1}^{k} \sum_{i=1}^{n_i} \left(x_{ij} - \overline{x}_j \right)^2$
(5) $\sum_{i=1}^{n_i} x_{ij}^2$	$\sum_{i=1}^{n_1} x_{i1}^2$	*Bấm . ..SUM -* $\sum x^2$	...	$\sum_{i=1}^{n_k} x_{ik}^2$	$T = \sum_{i=1}^{n_i} \sum_{j=1}^{k} x_{ij}^2$
(6) Tổng bình phương chênh lệch toàn bộ	$SST = \sum_{j=1}^{k} \sum_{i=1}^{n_i} \left(x_{ij} - \overline{x} \right)^2$ *Bấm* $T - N \times \left(\overline{x} \right)^2$ hoặc $SST = SSB + SSW$				

Tiêu chuẩn kiểm định:

Source of Variation	Tổng Bình phương chênh lệch	Bậc tự do	Phương sai (Trung bình BPCL)	Tiêu chuẩn kiểm định F
Between Groups	$SSB = \sum_{j=1}^{k} n_j \times \left(\overline{x}_j - \overline{x}\right)^2$	**k -1**	$MSB = \frac{SSB}{k-1}$	$F = \frac{MSB}{MSW}$
Within Groups	$SSW = \sum_{j=1}^{k}\sum_{i=1}^{n_i}\left(x_{ij} - \overline{x}_j\right)^2$	**N - k**	$MSW = \frac{SSW}{n-k}$	
Total	$SST = \sum_{j=1}^{k}\sum_{i=1}^{n_i}\left(x_{ij} - \overline{x}\right)^2$	N –1		

Miền bác bỏ:

$$\mathbf{W_\alpha = (f_\alpha (k\text{-}1;\ N - k);\ +\infty)}$$



Nhận xét:

- **SSB** (hay SSG) **:** Phần biến thiên của giá trị X do các mức độ của yếu tố đang xem xét tạo ra.
- **SSW:** Phần biến thiên của giá trị X do các yếu tố nào đó không được xem xét tạo ra.
- **SST:** Tổng các biến thiên của X do tất cả các yếu tố tạo ra.

Hệ số xác định R^2: $$R^2 = \frac{SSB}{SST} \times 100\%$$

Hệ số xác định R^2 của mô hình Phân tích phương sai được sử dụng để đo mức độ ảnh hưởng của yếu tố được xem xét trong mô hình đối với sự biến động của các giá trị của biến ngẫu nhiên X quanh giá trị trung bình của nó. R^2 càng lớn thì mô hình càng gọi là thích hợp.

Ví dụ 26:

Khi theo dõi tác động của các điều kiện ngoại cảnh đến sự sinh trưởng của 1 loại cây non, người ta gieo trồng cùng 1 loại hạt giống trong 3 điều kiện ngoại cảnh A, B, C khác nhau và thu được số liệu mẫu sau:

Điều kiện ngoại cảnh	Chiều cao của cây (cm)					
A	48	51	57	62	59	55
B	46	42	45	50	47	51
C	44	55	53	56	54	

Hãy dùng phương pháp Anova để so sánh chiều cao trung bình của các cây con trong 3 điều kiện ngoại cảnh trên và kết luận với mức ý nghĩa 5%. (Lưu ý bổ sung thêm các giả thiết cần có để thực hiện được yêu cầu bài toán)

Hướng dẫn:

* Gọi a_1; a_2; a_3 lần lượt là chiều cao trung bình của các cây con được trồng trong các điều kiện ngoại cảnh A; B; C.

Giả thiết kiểm định H_0: $a_1 = a_2 = a_3$

Giả thiết đối H_1: $\exists\, a_i \neq a_j$ với $i \neq j$

Các giả thiết cần có: *Xem điều kiện bài toán .*

* Miền bác bỏ $W_\alpha = (f_{0.05}(2; 14); +\infty) = (3.7389; +\infty)$

Tra bảng Fisher $\alpha = 0.05$; bậc tử: $n_1 = 2$;

bậc mẫu: $n_2 = 14$

* Tính tiêu chuẩn kiểm định:

Xem 2 bảng phía sau.

	A	B	C	Mẫu gộp
	48 62 **51 59** **57 55**	**46 50** **42 47** **45 51**	**44 56** **55 54** **53**	
(1) Kích thước mẫu	$n_1 = 6$	$n_2 = 6$	$n_3 = 5$	$N = 17$
(2) Trung bình mẫu	$\overline{x}_1 = 55.3333$	$\overline{x}_2 = 46.8333$	$\overline{x}_3 = 52.4$	$\overline{x} = 51.4706$
(3) (SSG) **SSB**				**222.8686**
(4) **SSW**				**281.3667**
(5) Σx_i^2				
(6) **SST**	**504.2353**			

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	222.8686	2	111.4343	5.544653	0.016845	3.738892
Within Groups	281.3667	14	20.09762			
Total	504.2353	16				

$f_\alpha (k-1; N-k)$

* Kết luận:

Do tiêu chuẩn kiểm định $F_{qs} = 5.5447 \in W_\alpha$ nên bác bỏ H_0; chấp nhận $H_1 \Rightarrow$ Chiều cao trung bình của cây non sinh trưởng ở các điều kiện A; B; C là không bằng nhau.

Cách nói khác: Chiều cao (trung bình) của cây phụ thuộc vào điều kiện ngoại cảnh.

Ví dụ 27:

Khi đo mức độ bụi trong không khí tại 3 khu vực trong thành phố, người ta được số liệu sau (đơn vị mg/m^3):

Số thứ tự quan sát	Khu vực 1	Khu vực 2	Khu vực 3
1	0,54	0,48	0,56
2	0,60	0,49	0,62
3	0,72	0,55	0,60
4	0,67	0,62	
5	0,83	0,57	

Với mức ý nghĩa 5%, có thể coi như mức độ bụi trung bình ở các khu vực trên là như nhau không? Lưu ý bổ sung thêm các giả thiết cần có để thực hiện được yêu cầu bài toán.

Tìm hệ số R^2 và nêu ý nghĩa.

	Khu vực 1	Khu vực 2	Khu vực 3	Mẫu gộp
	0.54, 0.67 0.6, 0.83 0.72	0.48, 0.62 0.49, 0.57 0.55	0.56 0.62 0.6	
(1) Kích thước mẫu	$n_1 = 5$	$n_2 = 5$	$n_3 = 3$	$N = 13$
(2) Trung bình mẫu	$\overline{x_1} =$	$\overline{x_2} =$	$\overline{x_3} =$	$\overline{x} =$
(3) (SSG) **SSB**				**0.0427**
(4) **SSW**				**0.0652**
(5) $\Sigma x_i^{\ 2}$				
(6) **SST**				

Bài tập tham khảo:

Từ giáo trình Xác suất –

- thống kê & Phân tích số liệu; tài liệu (3).

Năm 1992, trong một nghiên cứu về ảnh hưởng của trục cuốn ép lên cường độ chịu nén của các loại thùng carton tiêu chuẩn RSC được sản xuất, Burgess đã tiến hành đo cường độ chịu nén của bốn loại thùng carton khác nhau. Dưới đây là dữ liệu thu được

Hộp	Lực nén					
1	655.5	788.3	734.3	721.4	679.1	699.4
2	789.2	772.5	786.9	686.1	732.1	774.8
3	737.1	639.0	727.1	671.7	717.2	727.1
4	535.1	628.7	542.4	559.0	586.9	520.0

Hãy so sánh cường độ chịu nén của bốn loại thùng carton với mức ý nghĩa $\alpha = 0.01$.

Một công ty dược phẩm so sánh ba công thức thuốc giảm đau cho chứng đau nửa đầu. Trong thí nghiệm, 27 người tình nguyện được chia ngẫu nhiên thành ba nhóm, mỗi nhóm tương ứng với một công thức thuốc. Những người này sẽ uống thuốc khi bị đau nửa đầu và và phản hồi mức đau đầu sau khi uống thuốc (10 là mức đau nhất). Dưới đây là dữ liệu thu được

Drug A	4	5	4	3	2	4	3	4	4
Drug B	6	8	4	5	4	6	5	8	6
Drug C	6	7	6	6	7	5	6	5	5

Hãy so sánh tác dụng của ba công thức thuốc trên với mức ý nghĩa 0.05.