Kiểm định giả thuyết (Hypthesis testing)

Đặng Thanh Hải (Ph.D)
School of Engineering and Technology, VNUH
Email: hai.dang@vnu.edu.vn

Nguyên lý

Căn cứ trên các số liệu “**thống kê**” được (tính toán được) trên mẫu, chúng ta cần kết luận xem **một giả thiết thống kê** (về các tham số như giá trị trung bình, phương sai, …) có đúng trên toàn bộ quần thể không?.

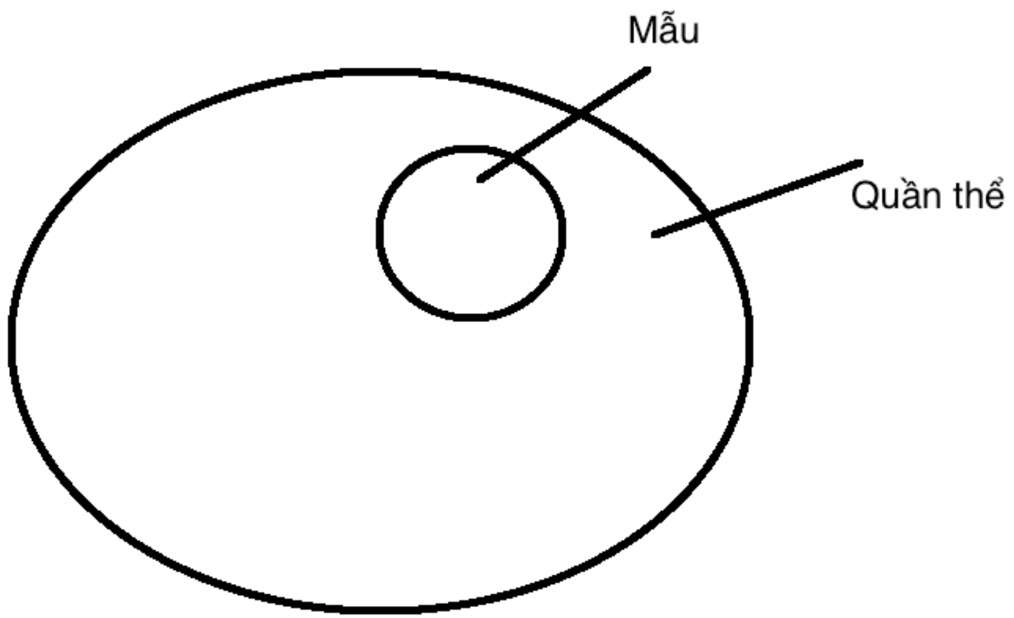
Ví dụ

1) Tập hợp chính có phân bố chuẩn với kỳ vọng bằng 3

2) Phương pháp điều trị A chữa khỏi 90% bệnh nhân

3) Tuổi thọ trung bình của hai loại bóng đèn A và B là như nhau

Vì sao cần kiểm thử giả thuyết



Nghiên cứu một thuộc tính của quần thể dựa vào 1 tập mẫu. Sử dụng dữ liệu thống kê được từ tập mẫu để kiểm định giả thuyết về các thuộc tính của quần thể.

Các loại giả thuyết

- Giả thuyết không (H0)
  - Là một phát biểu về tham số của tổng thể
  - Thường là một tuyên bố bị nghi ngờ
  - Vẫn cho là đúng cho đến khi nó được chứng minh là sai

- Giả thuyết thay thế (Đối thiết) (Ha)
  - Nhà nghiên cứu mong muốn ủng hộ và chứng minh là đúng
  - Là phát biểu ngược với H0
  - Được cho là đúng nếu H0 bị bác bỏ

Kiểm định giải thuyết nhằm mục đích bác bỏ hoặc không bác bỏ H0.

Kiểm định giả thiết về kỳ vọng (mẫu lớn)

- Giả thuyết "có thay đổi": H0: $\mu = \mu0$
  Ha: $\mu \# \mu0$

- Giả thuyết "thay đổi lớn hơn": H0: $\mu <= \mu0$
  Ha: $\mu > \mu0$
  Lưu ý: Chúng ta phải bác bỏ H0 để giả thuyết Ha đúng

- Giả thuyết "thay đổi nhỏ hơn": H0: $\mu >= \mu0$
  Ha: $\mu < \mu0$
  Lưu ý: Chúng ta phải bác bỏ H0 để giả thuyết Ha đúng

Trong đó $\mu0$ là giá trị cho trước.

Kiểm định giả thuyết

$$\text{Accept } H_0 \quad \text{if} \quad \mathbf{X}_n \in \tilde{R}^c$$
$$\text{Reject } H_0 \quad \text{if} \quad \mathbf{X}_n \in \tilde{R}.$$

Two kinds of errors can occur when executing this decision rule:

Type I error: Reject H_0 when H_0 is true.
Type II error: Accept H_0 when H_0 is false.

If the hypothesis is true, then we can evaluate the probability of a Type I error:

$$\alpha \triangleq P[\text{Type I error}] = \int_{\mathbf{x}_n \in \tilde{R}} f_{\mathbf{X}}(\mathbf{x}_n \mid H_0)\, d\mathbf{x}_n.$$

We call α the **significance level** of a test

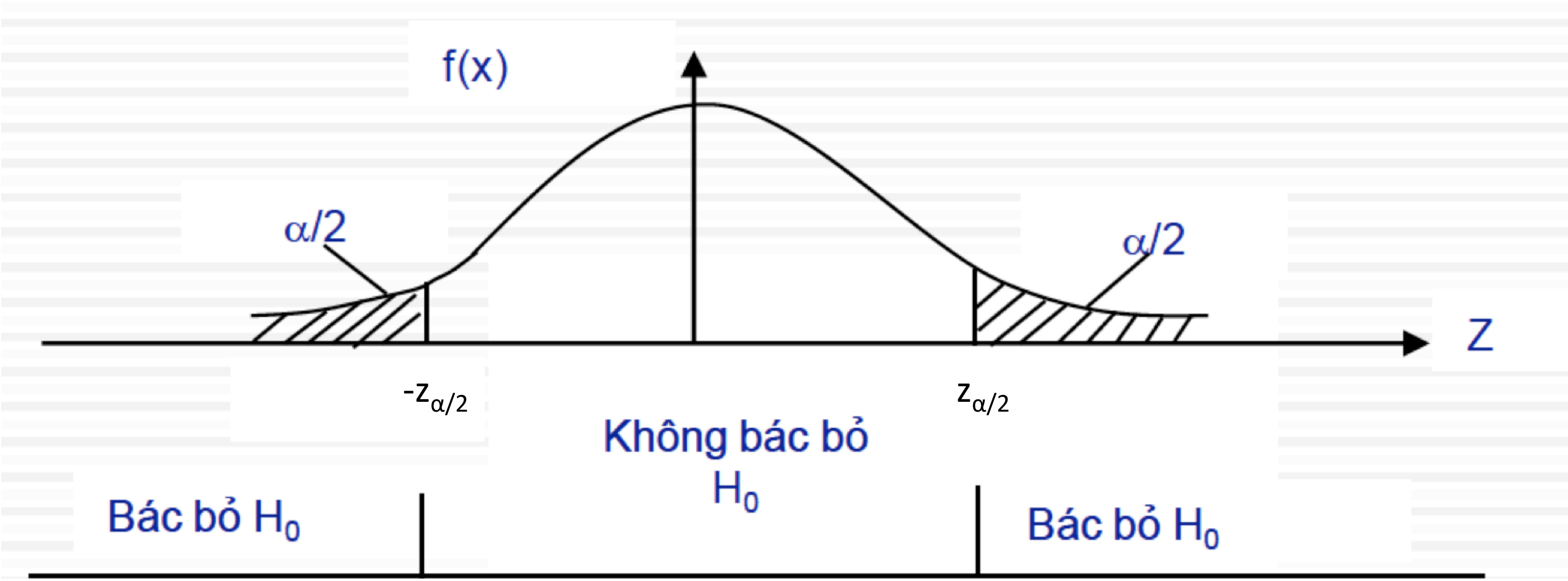
Kiểm định trên 1 mẫu

Kiểm định giả thuyết “có thay đổi”

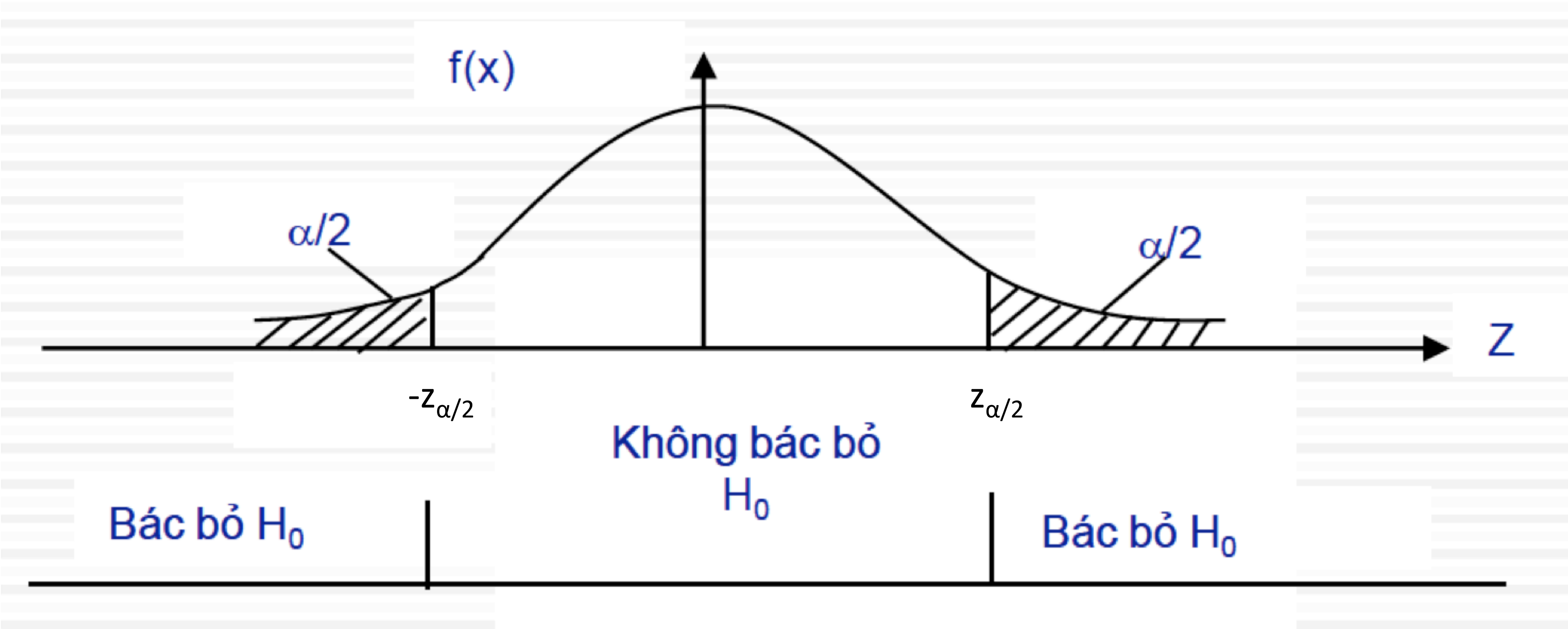
H0: μ = μ0
Ha: μ # μ0

Đây thường là kiểm thử giả thuyết một số yếu tố thay đổi và làm thay đổi một thuộc tính nào đó của quần thể.

Kiểm định 2 phía với α là *mức ý nghĩa*. Quy tắc bác bỏ H0 như sau



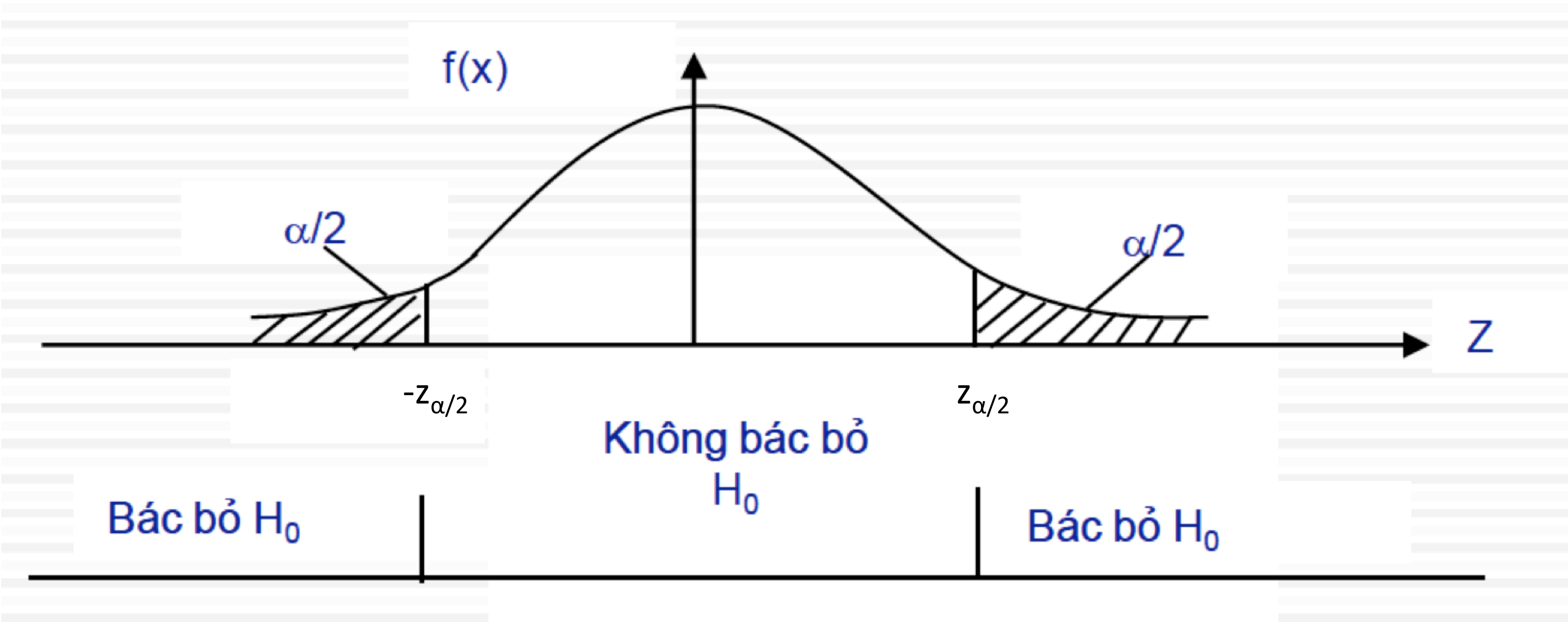
Kiểm định giả thuyết “có thay đổi”



Giá trị kiểm định:

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

Kiểm định giả thuyết “có thay đổi”



So sánh giá trị kiểm định z với giá trị $-z_{\alpha/2}$ và $z_{\alpha/2}$.
Nếu $z < -z_{\alpha/2}$ hoặc $z > z_{\alpha/2}$ thì bác bỏ giả thuyết H0 ($\mu = \mu 0$).
Nếu không thì không bác bỏ giả thuyết H0.

Kiểm định giả thuyết “có thay đổi”

Ví dụ: Một tập hợp chính có phân bố chuẩn với kỳ vọng μ (chưa biết) và độ lệch chuẩn σ=5.2, người ta lấy một mẫu kích thước n=100 và tính x = 27.56. Với mức ý nghĩa α=0.05 hãy kiểm định giả thuyết

H_0: μ=26

Với đối thiết H_a: $\mu \neq 26$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{27.56 - 26}{5.2 / \sqrt{100}} = 3$$

α=0.05 nên $z_{\alpha/2}$ = 1.96

Giá trị kiểm định rơi vào miền bác bỏ H_0. Như vậy bác bỏ H_0: μ=26

Kiểm định giả thuyết “có thay đổi”

- Một người nông dân sử dụng 1 loại phân bón mới cho 1 vườn táo và thu được 2756kg trên một 100 cây. Biết rằng mức trung binh khi chưa sử dụng loại phân bón mới này là 26kg/1 cây với độ lệch chuẩn là 5,2kg. Hãy kiểm định giả thuyết sản lượng của cây táo có thay đổi bởi loại phân bón này với mức ý nghĩa là 0.05.

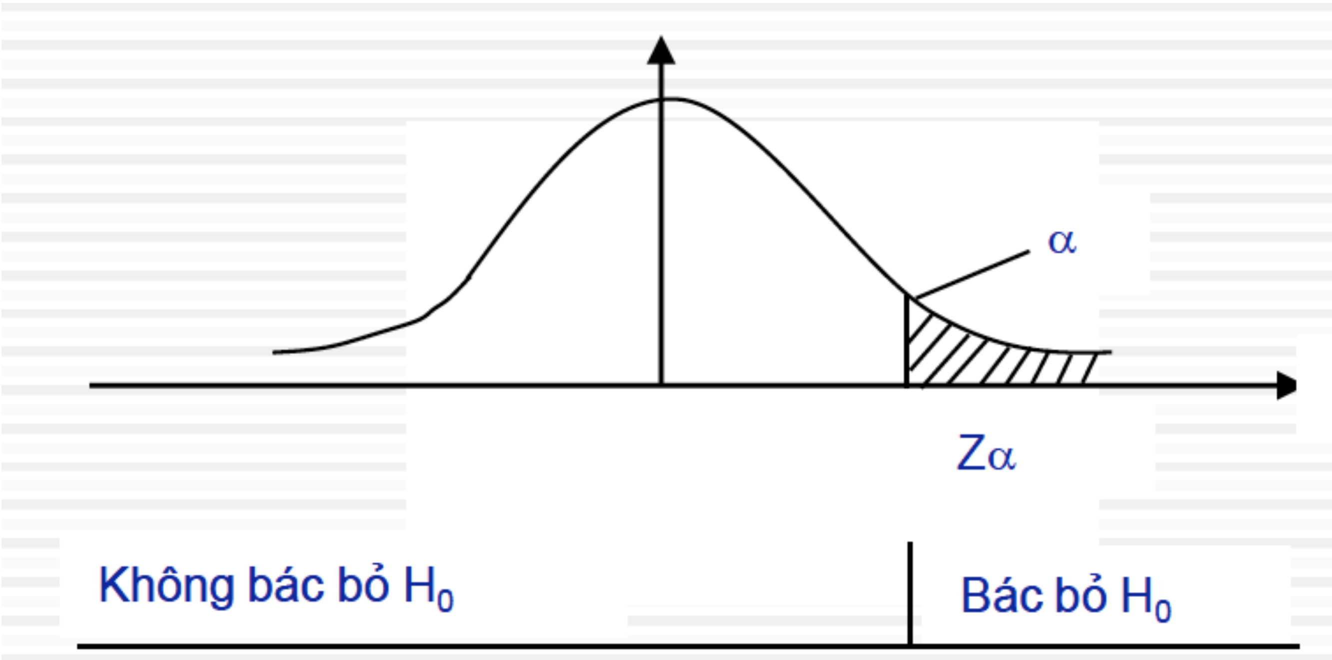
- Sau khi thay đổi giám đốc mới, nhà máy sản xuất thép ghi nhận sản lượng trong 100 ngày, có trung bình và độ lệch chuẩn của mẫu lần lượt là 880 tấn và 50 tấn. Hãy kiểm định giả thuyết rằng sản lượng bình quân hàng ngày của nhà máy hiện nay khác với với mức sản lượng trung bình 892 tấn/ngày đã được ghi nhận cách đây 1 năm với mức ý nghĩa là 0.05.

Kiểm định giả thuyết “thay đổi lớn hơn”

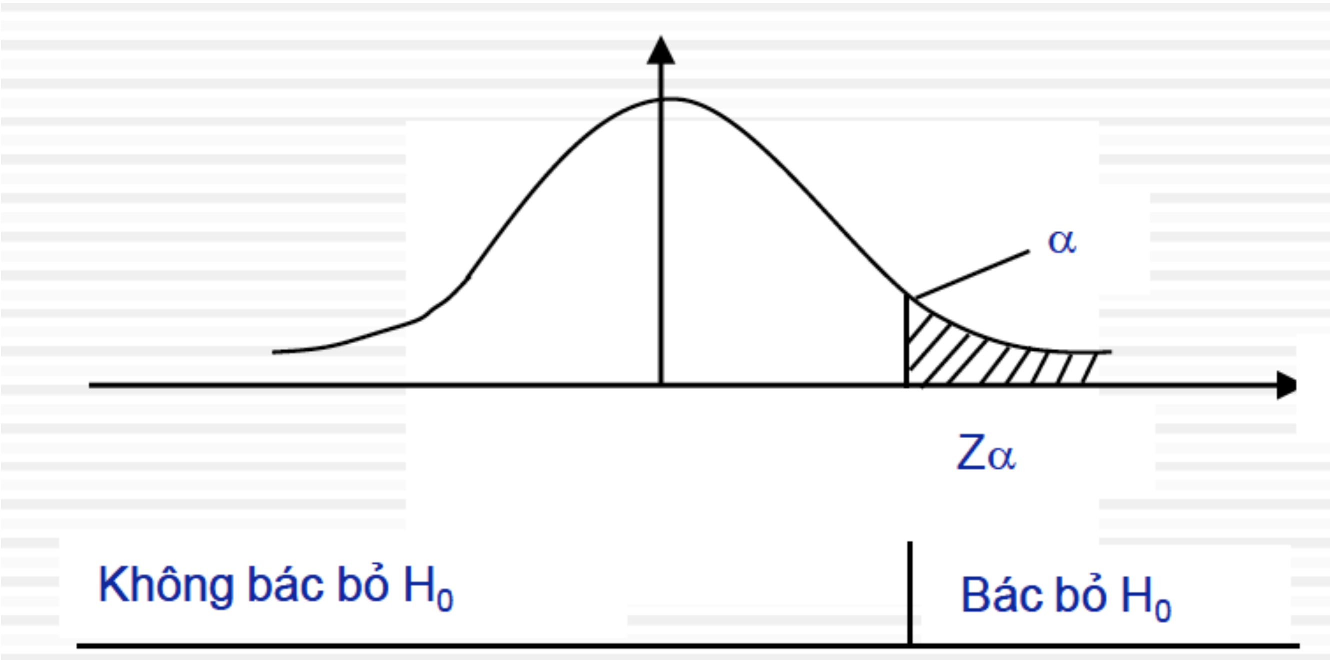
H0: μ <= μ0
Ha: μ > μ0

Đây thường là kiểm thử giả thuyết một số yếu tố thay đổi dẫn đến thay đổi tăng thêm một thuộc tính nào đó của quần thể**. Lưu ý: Chúng ta phải bác bỏ H0 để giả thuyết Ha đúng**

Kiểm định 1 phía với α là *mức ý nghĩa*. Quy tắc bác bỏ H0 như sau



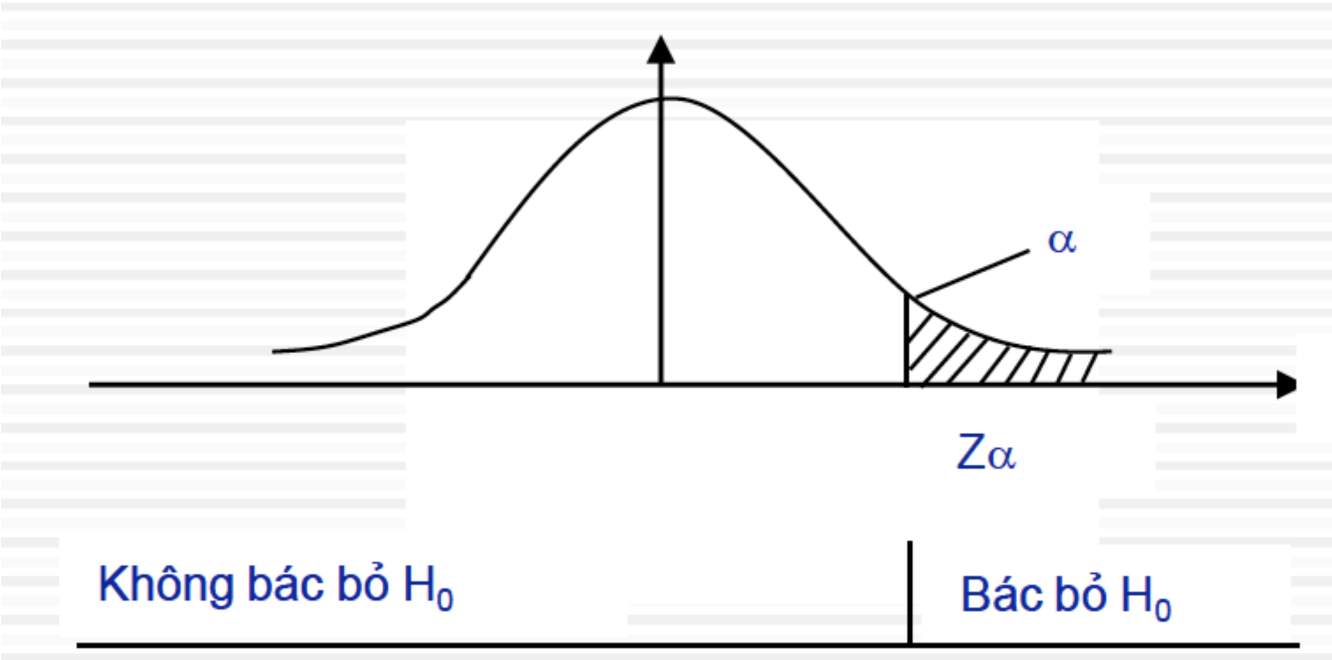
Kiểm định giả thuyết "thay đổi lớn hơn"



Giá trị kiểm định:

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

Kiểm định giả thuyết “thay đổi lớn hơn”



So sánh giá trị kiểm định z với giá trị z_α.
Nếu $z > z_\alpha$ thì bác bỏ giả thuyết H0 ($\mu <= \mu 0$) hay **Ha đúng**.
Nếu không thì không bác bỏ giả thuyết H0.

Kiểm định giả thuyết “thay đổi lớn hơn”

Ví dụ: Một tập hợp chính có phân bố chuẩn với kỳ vọng μ (chưa biết) và độ lệch chuẩn $\sigma=40$, người ta lấy một mẫu kích thước n=64 và tính x = 136.5. Với mức ý nghĩa $\alpha=0.01$ hãy kiểm định giả thuyết

H_0: $\mu\leq 130$

Với đối thiết H_a: $\mu>130$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{136.5 - 130}{40 / \sqrt{64}} = 1.3$$

$\alpha=0.01$ nên $z_\alpha = 2.33$

z=1.3<2.33. Như vậy chưa có cơ sở bác bỏ H_0

Kiểm định giả thuyết “thay đổi lớn hơn”

- Một người nông dân sử dụng 1 loại phân bón mới cho 1 vườn táo và thu được 2956kg trên một 100 cây. Biết rằng mức trung binh khi chưa sử dụng loại phân bón mới này là 26kg/1 cây với độ lệch chuẩn là 5,2kg. Hãy kiểm định giả thuyết sản lượng của cây táo tăng lên bởi loại phân bón này với mức ý nghĩa là 0.05.

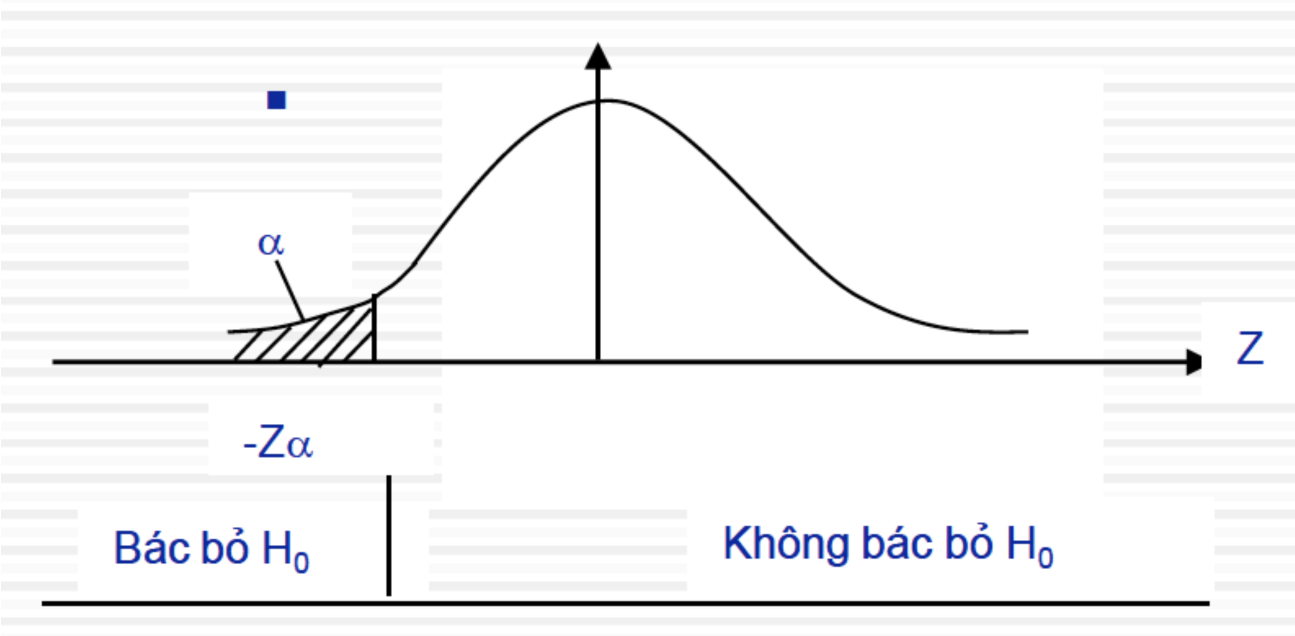
- Sau khi thay đổi giám đốc mới, nhà máy sản xuất thép ghi nhận sản lượng trong 100 ngày, có trung bình và độ lệch chuẩn của mẫu lần lượt là 980 tấn và 50 tấn. Hãy kiểm định giả thuyết rằng sản lượng bình quân hàng ngày của nhà máy hiện tăng hơn so với mức sản lượng trung bình 892 tấn/ngày đã được ghi nhận cách đây 1 năm với mức ý nghĩa là 0.05.

Kiểm định giả thuyết “thay đổi nhỏ hơn”

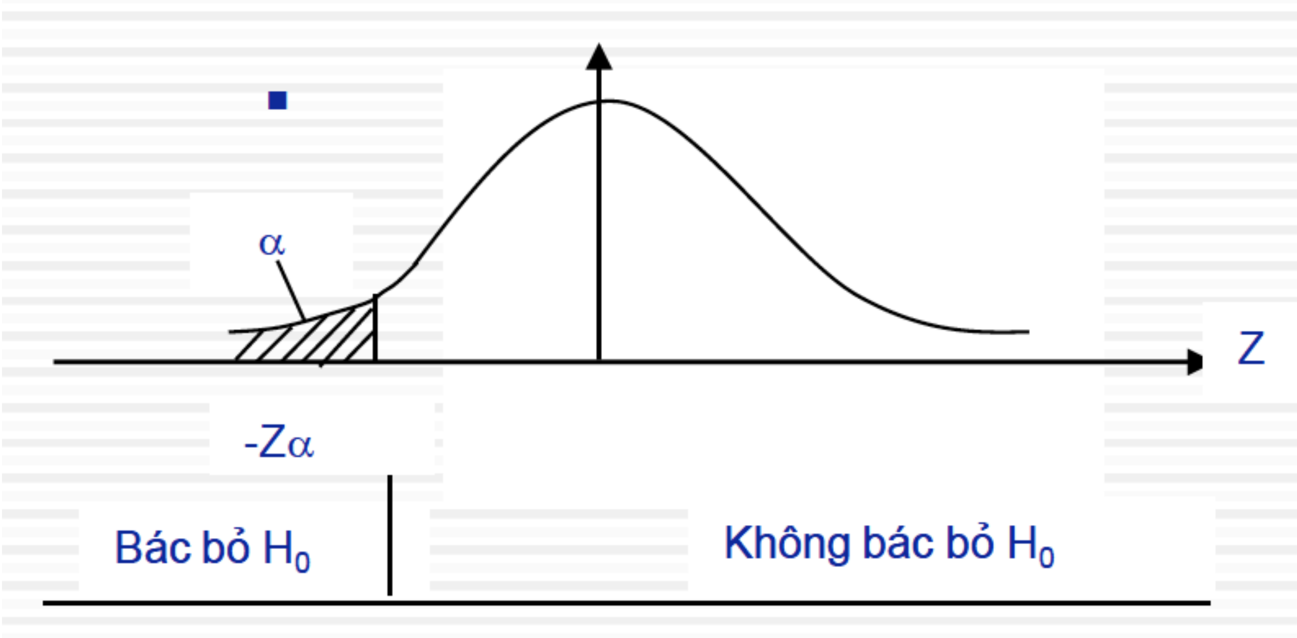
H0: μ >= μ0
Ha: μ < μ0

Đây thường là kiểm thử giả thuyết một số yếu tố thay đổi dẫn đến thay đổi nhỏ đi một thuộc tính nào đó của quần thể. **Lưu ý: Chúng ta phải bác bỏ H0 để giả thuyết Ha đúng**

Kiểm định 1 phía với α là *mức ý nghĩa*. Quy tắc bác bỏ H0 như sau



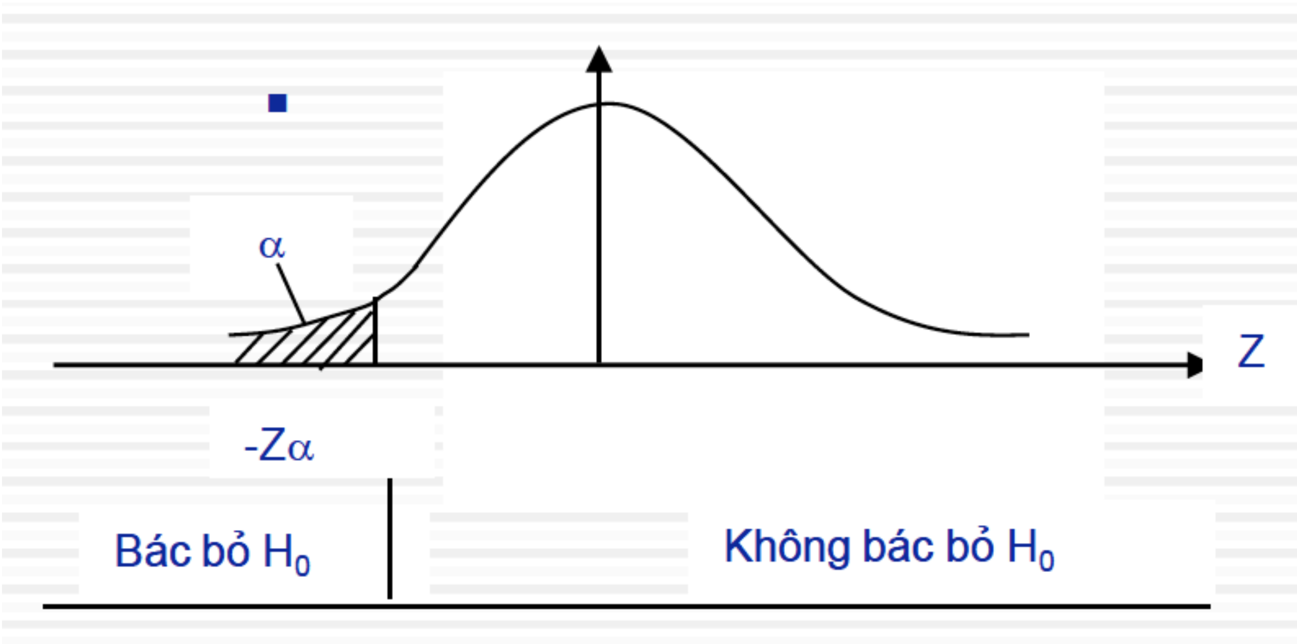
Kiểm định giả thuyết "thay đổi nhỏ hơn"



Giá trị kiểm định:

$$z = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$$

Kiểm định giả thuyết "thay đổi nhỏ hơn"



So sánh giá trị kiểm định z với giá trị $-z_\alpha$.
Nếu $z < -z_\alpha$ thì bác bỏ giả thuyết H0 ($\mu >= \mu0$), tức là Ha đúng.
Nếu không thì không bác bỏ giả thuyết H0.

Kiểm định giả thuyết “thay đổi nhỏ hơn”

Ví dụ 2: Một tập hợp chính có phân bố chuẩn với kỳ vọng μ (chưa biết) và độ lệch chuẩn $\sigma=0.4$, người ta lấy một mẫu kích thước n=100 và tính x = 31.9. Với mức ý nghĩa $\alpha=0.01$ hãy kiểm định giả thuyết

H_0: $\mu\geq 32$

Với đối thiết H_a: $\mu<32$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{31.9 - 32}{0.4 / \sqrt{100}} = -2.5$$

$\alpha=0.01$ nên $z_\alpha = 2.33$

z=-2.5<-2.33. Như vậy bác bỏ H_0 và kết luận

Kiểm định giả thuyết “thay đổi nhỏ hơn”

- Một người nông dân sử dụng 1 loại phân bón mới cho 1 vườn táo và thu được 2356kg trên một 100 cây. Biết rằng mức trung binh khi chưa sử dụng loại phân bón mới này là 26kg/1 cây với độ lệch chuẩn là 5,2kg. Hãy kiểm định giả thuyết sản lượng của cây táo bị giảm đi do loại phân bón này với mức ý nghĩa là 0.05.

- Sau khi thay đổi giám đốc mới, nhà máy sản xuất thép ghi nhận sản lượng trong 100 ngày, có trung bình và độ lệch chuẩn của mẫu lần lượt là 820 tấn và 50 tấn. Hãy kiểm định giả thuyết rằng sản lượng bình quân hàng ngày của nhà máy hiện tại giảm đi so với mức sản lượng trung bình 892 tấn/ngày đã được ghi nhận cách đây 1 năm với mức ý nghĩa là 0.05.

Phương sai chưa biết

- Nếu tập mẫu có kích thước lớn (n>=30), phương sai của quần thể có thể được ước lượng bằng phương sai của tập mẫu.

$$\hat{\sigma}_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X}_n)^2.$$

Sử dụng thống kê $T = \dfrac{\overline{X}_n - \mu}{\hat{\sigma}_n/\sqrt{n}}.$ T tuân theo **phân bố t của Student với n-1 bậc tự do**

Tương tự như slides trước

Phương sai chưa biết: Ví dụ

Thí dụ 6. Một nhóm nghiên cứu công bố rằng trung bình một người vào siêu thị A tiêu hết 140 ngàn đồng. Chọn một mẫu ngẫu nhiên gồm 50 người mua hàng ta tính được số tiền trung bình họ tiêu là 154 nghìn với độ lệch tiêu chuẩn là 62 nghìn. Với mức ý nghĩa 0,02 hãy kiểm định xem công bố của nhóm nghiên cứu có đúng hay không ?

Thí dụ 7. Một công ty có một hệ thống máy tính có thể xử lí 1200 hóa đơn trong một giờ. Công ty mới nhập một hệ thống máy tính mới. Hệ thống này khi chạy kiểm tra trong 40 giờ cho thấy số hóa đơn được xử lí trung bình trong 1 giờ là 1260 với độ lệch tiêu chuẩn là 215. Với mức ý nghĩa 5% hãy nhận định xem hệ thống mới có tốt hơn hệ thống cũ hay không ?

Phương sai chưa biết: Ví dụ

Một người Mỹ trung bình đọc 10 cuốn sách/năm. Khảo sát 136 người, trung bình là 12 cuốn, độ lệch tiêu chuẩn mẫu 9 cuốn.

Nhận định một người Mỹ đọc nhiều hơn 10 cuốn trong 1 năm với alpha 5%

Nhà sản xuất tuyên bố trung bình 1 chiếc bánh có 88 calo, chọn ngẫu nhiên 36 chiếc, trung bình là 90 calo với độ lệch tiêu chuẩn 4 calo. Kiểm định: 1 chiếc chứa nhiều hơn 88 calo với mức ý nghĩa 5%.

Phương sai chưa biết: Ví dụ

Mỗi nhân viên bán trung bình được 780 ngàn, trong đợt khuyến mại 80 nhân viên bán trung bình được 920 với độ lệch tiêu chuẩn 620

Với alpha 0.1, kiểm chứng nhiều hơn ngày thường

Kiểm định giả thiết về tỷ lệ

Nghiên cứu một quần thể mà mỗi cá thể có thể có hoặc không có một thuộc tính A nào đó.

- P là tỉ lệ cá thể có thuộc tính A trong quần thể
- f = k/n là tỉ lệ (tần suất) cá thể có thuộc tính A trong mẫu nghiên cứu

Câu hỏi: Ta muốn kiểm định giải thiết liên quan đến p dựa vào tần suất f.

Nhắc lại: Tần suất f là một ĐLNN có phân bố xấp xỉ phân bố chuẩn với kì vọng Ef = p và phương sai Df = p(1-p)/n với điều kiện np>5 và n(1-p)>5.

$$Z=(f-p_0)/\sqrt{p_0(1-p_0)/n}$$

Kiểm định giả thiết về tỷ lệ

Ví dụ: Một đảng chính trị trong một cuộc bầu cử tổng thống ở Mỹ tuyên bố rằng 45% cử trị sẽ bỏ phiếu cho ứng viên A của họ.
Chọn ngẫu nhiên 200 cử tri để thăm dò ý kiến cho thấy 80 người trong số đó tuyên bố bỏ phiếu cho ông A
Với mức $\alpha=5\%$, hãy kiểm định xem dự đoán của đảng trên đúng không?
H0: p=0.45
Ha: p≠0.45
f=80/200=0.4
np_0=200(0.45)=90 >=5
$n(1-p_0)$=200(0.55)=110 >=5
$Z=(f-p_0)\sqrt{n}/\sqrt{p_0(1-p_0)}=(0.4-045)\sqrt{200}/\sqrt{(0.45)(0.55)}=-1,43$
$\alpha=0.05$, nên $z_{\alpha/2}=1.96$
→ Không có cơ sở bác bỏ H_0

Kiểm định giả thiết về tỷ lệ: Ví dụ

18% gia đình ở thành phố A có máy tính ở nhà. Để kiểm tra chọn ngẫu nhiên 80 thấy 22 gia đình có. Với mức alpha 0.02, kiểm định gia đình có máy tính cao hơn tỉ lệ chung?

Kiểm định giả thiết về tỷ lệ: Ví dụ

Thí dụ 13. Một báo cáo nói rằng 18% gia đình ở thành phố A có máy tính cá nhân ở nhà. Để kiểm tra, người ta chọn ngẫu nhiên 80 gia đình trong thành phố có trẻ em đang đi học và thấy rằng có 22 gia đình có máy tính. Với mức ý nghĩa $\alpha = 0{,}02$ hãy kiểm định xem liệu trong các gia đình có trẻ em đang đi học, tỉ lệ gia đình có máy tính có cao hơn tỉ lệ chung hay không.

Hãy kiểm thử với mức ý nghĩa

- 5%
- 1%

Kiểm định giả thiết về tỷ lệ: Ví dụ

Độ bền của một loại dây thép sản xuất theo công nghệ cũ là 150. Sau khi cải tiến kỹ thuật người ta lấy mẫu gồm 100 sợi dây thép để thử độ bền thì thấy độ bền trung bình là 185 và $s = 25$. Với mức ý nghĩa $\alpha = 0,05$, hỏi công nghệ mới có tốt hơn công nghệ cũ hay không?

Độ dày của một chi tiết máy do một máy sản xuất là một đại lượng ngẫu nhiên phân phối theo qui luật chuẩn với độ dày trung bìng $1,25mm$. Nghi ngờ máy hoạt động không bình thường người ta kiểm tra 10 chi tiết máy thì thấy độ dài trung bình là $1,325$ với độ lệch tiêu chuẩn $0,075mm$. Với mức ý nghĩa $\alpha = 0,01$, hãy kết luận về điều nghi ngờ nói trên?

Thí dụ 14. Một công ty A sản xuất bánh kẹo tuyên bố rằng $\frac{2}{3}$ số trẻ em thích ăn bánh của công ty. Trong một mẫu gồm 100 trẻ em được hỏi, có 55 em tỏ ra thích bánh của công ty A. Với mức ý nghĩa 5%, số liệu nói trên có chứng tỏ là tuyên bố của công ty là hơi quá đáng hay không ?

Hãy kiểm thử với mức ý nghĩa

- 2%
- 1%

Kiểm định phương sai

$$H_0\text{: } X \text{ is Gaussian with } \sigma_X^2 = \sigma_0^2 \text{ and } m \text{ unknown}$$

$$H_1\text{: } X \text{ is Gaussian with } \sigma_X^2 \neq \sigma_0^2 \text{ and } m \text{ unknown.}$$

Sử dụng thống kê $\chi^2 = \frac{(n-1)\hat{\sigma}_n^2}{\sigma_X^2} = \frac{1}{\sigma_X^2} \sum \sum_{j=1}^{n} (X_j - \overline{X}_n)^2$, tuân theo **phân bố Chi-bình phương với n-1 bậc tự do**

Miền bác bỏ Ho: $\tilde{R}^c = \left\{ \mathbf{x} : a \leq \frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} \leq b \right\}$

Khoảng tin cậy:

$$1 - \alpha = P\left[a \leq \frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} \leq b \right] = P\left[\chi^2_{1-\alpha/2, n-1} < \frac{(n-1)\hat{\sigma}_n^2}{\sigma_0^2} < \chi^2_{\alpha/2, n-1} \right]$$

Kiểm định trên 2 mẫu

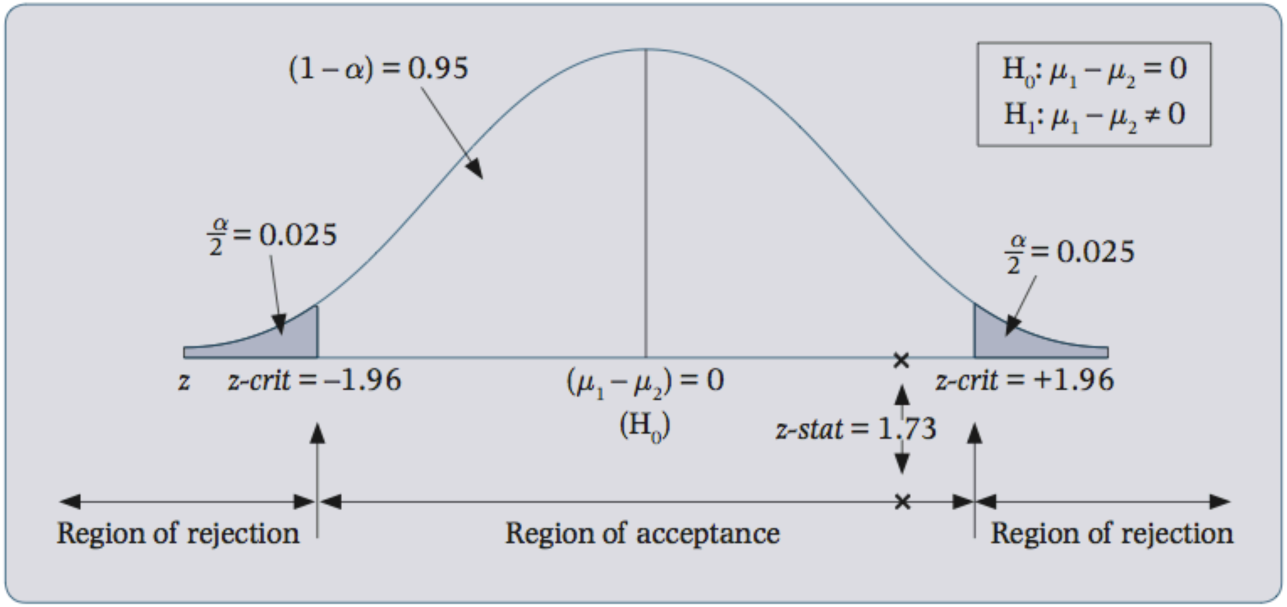
So sánh 2 kỳ vọng từ 2 mẫu độc lập

- Trường hợp 2 phương sai σ_1^2, σ_2^2 đã biết

Sử dụng thống kê

$$z\text{-}stat = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

, tuân theo **phân bố Chuẩn tắc N(0,1)**



So sánh 2 kỳ vọng từ 2 mẫu độc lập

- Trường hợp 2 phương sai chưa biết và giả sử 2 phương sai bằng nhau

Sử dụng thống kê

$$t\text{-}stat = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\text{where } s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

, tuân theo **phân bố t của Student với (n1 + n2 – 2) bậc tự do**

So sánh 2 kỳ vọng từ 2 mẫu độc lập

- Trường hợp 2 phương sai chưa biết và giả sử 2 phương sai khác nhau

Sử dụng thống kê

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Với $D_0 = \mu_1 - \mu_2$

, tuân theo **phân bố t của Student với *df* bậc tự do**

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

So sánh 2 kỳ vọng từ 2 mẫu phụ thuộc nhau

- Các mẫu [phụ thuộc] khớp đôi nhau:
  - Dữ liệu trong hai mẫu được đo trên cùng 1 đối tượng tại 2 thời điểm khác nhau hay dưới 2 điều kiện khác nhau
  - Ví dụ:

- An employee's job performance is recorded both *before* and *after* a training programme to observe whether any difference in performance has occurred due to the training programme.
- A patient's blood pressure is recorded *before* a drug is administered and then again *after* the treatment to see whether the treatment has had any effect on blood pressure.
- Internet usage of subscribers *last year* is compared to the same subscribers' internet usage *this year* to determine whether there is any change in usage levels.

So sánh 2 kỳ vọng từ 2 mẫu phụ thuộc nhau

Sử dụng thống kê

$$t\text{-}stat = \frac{(\bar{x}_d - \mu_d)}{\frac{s_d}{\sqrt{n}}}$$

, tuân theo **phân bố t của Student với *df=n-1* bậc tự do**

Trong đó:

n = the number of pairs of data (or number of respondents)

$$s_d = \sqrt{\frac{\Sigma(x_d - \bar{x}_d)^2}{n-1}}$$

$$\bar{x}_d = \frac{\Sigma x_d}{n}$$

$$\mu_d = \mu_1 - \mu_2$$

$$x_d = x_1 - x_2 \qquad \text{(i.e. differences between pairs)}$$

So sánh 2 tỷ lệ từ 2 mẫu độc lập nhau

- **Điều kiện**: Approximately has a normal distribution if each of the sample sizes n_1 and n_2 is large. Here n_1 and n_2 are large enough if n_1p_1, $n_1(1 - p_1)$, n_2p_2, and $n_2(1 - p_2)$ are all at least 5.

Sử dụng thống kê

$$z\text{-}stat = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\hat{\pi}(1 - \hat{\pi})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{\pi} = \frac{x_1 + x_2}{n_1 + n_2}$ $\quad p_1 = \frac{x_1}{n_1}$ $\quad p_2 = \frac{x_2}{n_2}$

$\hat{\pi}$ = pooled sample proportion

, tuân theo **phân bố Chuẩn tắc N(0,1)**

So sánh 2 phương sai của 2 mẫu độc lập nhau

- Kiểm định một phía $H_0: \sigma_1^2 = \sigma_2^2$ versus $H_a: \sigma_1^2 > \sigma_2^2$,

Sử dụng thống kê

$$F = \frac{s_1^2}{s_2^2}$$

, tuân theo **phân bố F với *df_1 = n_1 -1* bậc tự do tử số và *df_2 = n_2 – 1* bậc tự do mẫu số**

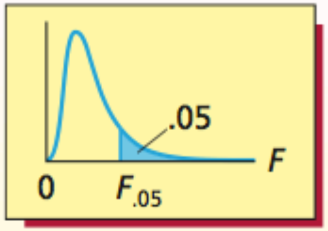
- Kiểm định một phía $H_0: \sigma_1^2 = \sigma_2^2$ versus $H_a: \sigma_1^2 < \sigma_2^2$,

Sử dụng thống kê

$$F = \frac{s_2^2}{s_1^2}$$

, tuân theo **phân bố F với *df_1 = n_2 -1* bậc tự do tử số và *df_2 = n_1 – 1* bậc tự do mẫu số**

TABLE 11.2 A Portion of an F Table: Values of $F_{.05}$



df_2 \ df_1	Numerator Degrees of Freedom (df_1) 1	2	3	4	5	6	7	8	9	10	12	15	20	24
Denominator Degrees of Freedom (df_2) 1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84
7	5.59	4.71	4.25	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29

Source: M. Merrington and C. M. Thompson, "Tables of Percentage Points of the Inverted Beta (F) Distribution," *Biometrika,* Vol. 33 (1943), pp. 73–88. Reproduced by permission of Oxford University Press and *Biometrika* trustees.

So sánh 2 phương sai của 2 mẫu độc lập nhau

- Kiểm định 2 phía $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_a\colon \sigma_1^2 \neq \sigma_2^2$,

Sử dụng thống kê

$$F = \frac{\text{the larger of } s_1^2 \text{ and } s_2^2}{\text{the smaller of } s_1^2 \text{ and } s_2^2}$$

F tuân theo **phân bố F với *df_1* bậc tự do tử số và *df_2* bậc tự do mẫu số**

Trong đó:

df_1 = {the size of the sample having the larger variance} − 1

df_2 = {the size of the sample having the smaller variance} − 1

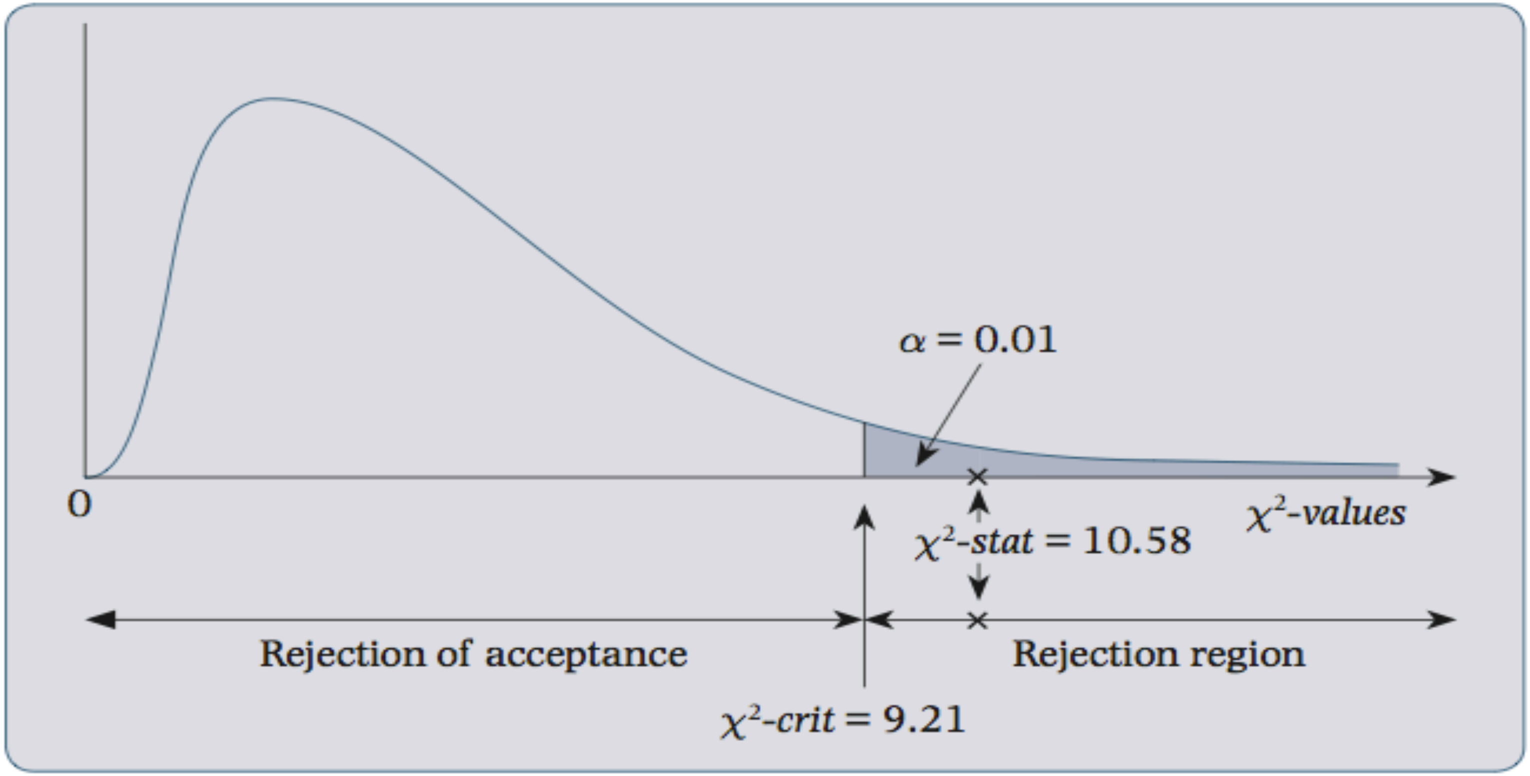
Kiểm định giả thuyết Chi-bình-phương

Kiểm định phù hợp

- Xét một thí nghiệm (ngẫu nhiên) đa thức trong đó mỗi phần tử trong n phần tử được chọn ngẫu nhiên được phân vào 1 trong k nhóm.
  - f_i = số phần tử được phân vào nhóm *i* (chính là tần suất quan sát được thứ *i*)
  - $E_i = np_i$ = số phần tử kỳ vọng sẽ được phân vào nhóm *i* nếu pi là xác suất để một phần tử ngẫu nhiên được phân vào nhóm *i* (tần suất kỳ vọng thứ *i*)
- Kiểm định:
  - Ho: các giá trị xác suất là $p_1, p_2, ..., p_k$ – trong đó p_i là xác suất để một phần tử ngẫu nhiên được phân vào nhóm *i*.
  - Ha: ít nhất một giá trị xác suất trên không đúng như được phát biểu trong Ho.

Kiểm định phù hợp

Sử dụng thống kê $$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - E_i)^2}{E_i}$$ tuân theo **phân bố Chi-bình-phương với *k-1* bậc tự do**



Kiểm định phù hợp với phân bố chuẩn

1 We will test the following null and alternative hypotheses:

H_0: the population has a normal distribution

H_a: the population does not have a normal distribution

2 Select a random sample of size n and compute the sample mean $\bar{x}$ and sample standard deviation s.

3 Define k intervals for the test. Two reasonable ways to do this are to use the classes of a histogram of the data or to use intervals based on the Empirical Rule.

4 Record the observed frequency (f_i) for each interval.

5 Calculate the expected frequency (E_i) for each interval under the normality assumption. Do this by computing the probability that a normal variable having mean $\bar{x}$ and standard deviation s is within the interval and by multiplying this probability by n. Make sure that each expected frequency is large enough. If necessary, combine intervals to make the expected frequencies large enough.

6 Calculate the chi-square statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - E_i)^2}{E_i}$$

and define the p-value for the test to be the area under the curve of the chi-square distribution having $k - 3$ degrees of freedom to the right of χ^2.

7 Reject H_0 in favor of H_a at level of significance α if either of the following equivalent conditions holds:

a $\chi^2 > \chi^2_\alpha$ **b** p-value $< \alpha$

Here, the χ^2_α point is based on $k - 3$ degrees of freedom.

Phân tích phương sai
(analysis of variance ANOVA):
so sánh kỳ vọng trên nhiều mẫu

One-Way Analysis of Variance

- To estimate and compare the effects of the different treatments on the response variable.
  - **Response variable** is assumed to be affected by one **factor**
- Các giả sử của One-Way ANOVA
  - **Constant variance**: the *p* populations of values of the ***response variable*** associated with the ***treatments*** have equal variances.
  - **Normality**: the *p* populations of values of the response variable associated with the treatments all have normal distributions.
  - **Independence**: the samples of experimental units associated with the treatments are randomly selected, independent samples.

One-Way Analysis of Variance

- Kiểm định sự khác nhau giữa các kỳ vọng

let n_i denote the size of the sample that has been randomly selected for treatment i,
let x_{ij} denote the j^{th} value of the response variable that is observed when using treatment i.

H_0: $\mu_1 = \mu_2 = \cdots = \mu_p$ (all treatment means are equal) versus H_a: At least two of $\mu_1, \mu_2, \ldots, \mu_p$ differ (at least two treatment means differ)

Sử dụng thống kê

$$F = \frac{MST}{MSE} = \frac{SST/(p-1)}{SSE/(n-p)}$$

tuân theo **phân bố F với** $df_1 = p-1$ **bậc tự do tử số và** $df_2 = n-p$ **bậc tự do mẫu số**

Trong đó:

$$MST = SST/(p-1)$$

$$MSE = SSE/(n-p)$$

$$SST = \sum_{i=1}^{p} n_i(\bar{x}_i - \bar{x})^2$$ **treatment sum of squares**

$$SSE = \sum_{j=1}^{n_1}(x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \bar{x}_2)^2 + \cdots + \sum_{j=1}^{n_p}(x_{pj} - \bar{x}_p)^2$$ **error sum of squares**

$$SSTO = SST + SSE$$ **total sum of squares**

Two-Way Analysis of Variance

- Studying the effects of **two factors** on a **response variable**
  - **factor 1,** has ***a* levels** (levels 1, 2, . . . , *a*)
  - **factor 2,** has ***b* levels** (levels 1, 2, . . . , *b*).
  - a **treatment** is considered to be a **combination of a level of factor 1 and a level of factor 2.**
  - assign ***m*** randomly selected **experimental units** to each **treatment**.

$x_{ij,k}$ = the kth value of the response variable observed when using level i of factor 1 and level j of factor 2

$\bar{x}_{ij}$ = the mean of the m values observed when using the ith level of factor 1 and the jth level of factor 2

$\bar{x}_{i\bullet}$ = the mean of the bm values observed when using the ith level of factor 1

$\bar{x}_{\bullet j}$ = the mean of the am values observed when using the jth level of factor 2

$\bar{x}$ = the mean of the abm values that we have observed in the experiment

Two-Way ANOVA Procedure

- **total sum of squares (*SSTO*)** is partitioned into four components
  - *SSTO = SS(1) + SS(2) + SS(int) + SSE*
    - ***SS(1)***: *factor 1 sum of squares,* ***SS(2)***: *factor 2 sum of squares,* ***SS(int)***: *interaction sum of squares,* ***SSE***: *error sum of squares*
- **Step 1:** Calculate *SSTO*, which measures the total amount of variability:

$$SSTO = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{m} (x_{ij,k} - \bar{x})^2$$

- **Step 2:** Calculate *SS*(1), which measures the amount of variability due to the different levels of factor 1:

$$SS(1) = bm \sum_{i=1}^{a} (\bar{x}_{i\bullet} - \bar{x})^2$$

Two-Way ANOVA Procedure

- **Step 3:** Calculate *SS*(2), which measures the amount of variability due to the different levels of factor 2:

$$SS(2) = am\sum_{j=1}^{b}(\bar{x}_{\bullet j} - \bar{x})^2$$

- **Step 4:** Calculate *SS*(int), which measures the amount of variability due to the interaction between factors 1 and 2:

$$SS(\text{int}) = m\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}_{ij} - \bar{x}_{i\bullet} - \bar{x}_{\bullet j} + \bar{x})^2$$

- **Step 5:** Calculate *SSE*, which measures the amount of variability due to the error:

$$SSE = SSTO - SS(1) - SS(2) - SS(\text{int})$$

Two-Way ANOVA Test

- Kiểm định sự độc lập
  - **H0**: no interaction exists between factors 1 and 2
  - **Ha**: interaction does exist

Sử dụng thống kê $F(\text{int}) = \dfrac{MS(\text{int})}{MSE}$ tuân theo **phân bố F với *df_1 = (a - 1)(b - 1)* bậc tự do tử số và *df_2 = (ab(m-1)* bậc tự do mẫu số**

Trong đó:

$$MS(\text{int}) = \frac{SS(\text{int})}{(a-1)(b-1)}$$

$$MSE = \frac{SSE}{ab(m-1)}$$

Two-Way ANOVA Test

- Kiểm định mức ý nghĩa của **factor 1**
  - **H0**: no differences exist between the effects of the different levels of factor 1
  - **Ha**: at least two levels of factor 1 have different effects

Sử dụng thống kê $$F(1) = \frac{MS(1)}{MSE}$$ tuân theo **phân bố F với $df_1 = (a - 1)$ bậc tự do tử số và $df_2 = ab(m-1)$ bậc tự do mẫu số**

Trong đó:

$$MS(1) = \frac{SS(1)}{a - 1}$$

$$MSE = \frac{SSE}{ab(m - 1)}$$

Two-Way ANOVA Test

- Kiểm định mức ý nghĩa của **factor 2**
  - **H0**: no differences exist between the effects of the different levels of factor 2
  - **Ha**: at least two levels of factor 2 have different effects

Sử dụng thống kê $$F(2) = \frac{MS(2)}{MSE}$$ tuân theo **phân bố F với** $df_1 = (a - 1)$ **bậc tự do tử số và** $df_2 = ab(m-1)$ **bậc tự do mẫu số**

Trong đó:

$$MS(2) = \frac{SS(2)}{b - 1}$$

$$MSE = \frac{SSE}{ab(m - 1)}$$

Kiểm định giả thuyết sử dụng P_value

P_value

- Không sử dụng khoảng bác bỏ, khoảng chấp nhận
- Kiểm định bằng cách đặt câu hỏi: "*Giả sử giả thuyết Ho đúng, thì xác suất để* ***biến ngẫu nhiên thống kê*** *lớn hơn hoặc bằng* ***giá trị thống kê*** *được từ mẫu là bao nhiêu**?***"
- **Kiểm định**:
  - TÍnh P_value = P[Biến ngẫu nhiên thống kê >= giá trị thống kê trên mẫu|Ho]
  - Nếu P_value <= mức ý nghĩa α thì bác bỏ Ho, ngược lại chấp nhận Ho.