

Kiểm Định Giả Thuyết (Hypothesis Testing)

Khái Quát

- Định nghĩa kiểm định giả thuyết.
- Các thành phần của kiểm định giả thuyết.
- Phương pháp thực hiện kiểm định giả thuyết.

Vấn Đề Thúc Đẩy

- Cân bằng giới tính mới sinh của bé trai và bé gái như thế nào? Có lớn hơn một giá trị nào đó hay không?
- Thu thập dữ liệu về lương trung bình của người đi làm ở Việt Nam. Mức lương trung bình có trên 8 triệu 1 tháng hay không?
- Nhà máy đóng chai nước giải khát. Quá trình đóng chai luôn có sai số về lượng nước giải khát trong chai. Nhà máy muốn tính sai số có vượt mức 5% in trên bao bì hay không?

Hypothesis Testing – Kiểm Định Giả Thuyết



Ý tưởng chung:

- Anh Cường được đồn thổi (giả định) là rất giàu. Một hôm thấy anh chạy xe **Ferrari Purosangue** ngoài đường, bằng chứng quan sát này có ủng hộ tin đồn hay không? Cùng là anh Cường đó nhưng nếu anh chạy **Honda Dream 2** thì kết luận của bạn thế nào.
- Người ra tòa được mặc định là vô tội. Bằng chứng sẽ được đưa ra để chứng minh họ có tội.
- Chúng ta sẽ sử dụng lý thuyết thống kê để tiếp cận vấn đề trên.

Những Thành Phần Của Kiểm Định Giả Thuyết (Hypothesis Testing)

- Giả thuyết vô hiệu (H_0).
- Giả thuyết thay thế (H_1).
- Một giá trị thống kê (a test statistic).
- Khoảng bác bỏ (the rejection region).

Giả Thuyết Vô Hiệu (H_0) và Giả Thuyết Thay Thế (H_1)

- Chúng ta bắt đầu bằng hai loại giả thuyết:
 - Giả thuyết vô hiệu H_0 (the null hypothesis).
 - Giả thuyết thay thế H_1 (the alternative hypothesis).
- Giả thuyết vô hiệu H_0 được giả định là đúng.

Giả Thuyết Vô Hiệu H_0

- Kiểm định giả thuyết là quá trình đi tìm bằng chứng để có thể bác bỏ hoặc không thể bác bỏ H_0 .
- Trong nhiều trường hợp, giả thuyết vô hiệu H_0 thường là một phát biểu về rằng một tác nhân 'không có tác động' hoặc 'không mang lại thay đổi' đối với đối tượng nghiên cứu.
 - Ví dụ: chính sách vay vốn không làm thay đổi thu nhập trung bình của nhóm thụ hưởng chính sách.
- Giả thuyết vô hiệu H_0 chỉ mang dấu $\neq, \geq, \leq$. Ví dụ:
$$H_0: \mu = 0; \quad H_0: \mu \geq 166; \quad H_0: \mu \leq 5.$$

Giả Thuyết Thay Thế H_1

- Giả thuyết thay thế sẽ trái ngược hoàn toàn với H_1 :

$$H_0: \mu = 0 \quad \text{vs} \quad H_1: \mu \neq 0$$

$$H_0: \mu \geq 166 \quad \text{vs} \quad H_1: \mu < 166$$

$$H_0: \mu \leq 5.5 \quad \text{vs} \quad H_1: \mu > 5.5$$

- Giả thuyết thay thế thường dùng để nêu lên giả thuyết **cần chứng minh** trong nghiên cứu khoa học.
- Khi có dữ liệu giúp chúng ta để bác bỏ H_0 , thì chúng ta có thể nói là đó là bằng chứng để chứng minh H_1 .

Giá Trị Kiểm Định Thống Kê (1)

- Giá trị kiểm định thống kê (the test statistic) là giá trị tính từ mẫu quan sát được dùng để thực hiện kiểm định thống kê. Cụ thể nó là một hàm $g(X_1, X_2, \dots, X_n)$ cho mẫu $X_1, X_2, \dots, X_n$
- Giá trị kiểm định là một biến ngẫu nhiên, theo một phân phối xác suất nhất định. Với mỗi mẫu quan sát, chúng ta sẽ các hiện thực hóa khác nhau cho giá trị này.
- Chúng ta đi khảo sát xem giá trị quan sát cho giá trị kiểm định có khả năng “cao” hay “thấp” tạo ra giả định.
 - Ví dụ: khi chạy Dream 2, người chạy khả năng cao hay thấp là giàu không?

Hai Giả Thuyết Kiểm Định Thống Kê

- Tổng quát người ta thường viết như sau:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

- μ_0 sẽ là giá trị chúng ta giả định là đúng (giá trị sinh ra mẫu quan sát được thu thập).

Giá Trị Kiểm Định Thống Kê (2)

- Chúng ta có bài toán xác định trung bình chiều cao của nam giới Việt Nam có phải là 166cm hay không?
- Giả thuyết H_0 và H_1 cho kiểm định giả thuyết hai phía (two-tailed/two sided tests):

$$H_0: \mu = 166$$

$$H_1: \mu \neq 166$$

Giá Trị Kiểm Định Thống Kê (4)

- Giá trị kiểm định chúng là sẽ là Z .

$$Z = \frac{\bar{X} - 166}{\sigma/\sqrt{n}} \sim N(0,1)$$

- Với mỗi mẫu khác nhau cùng $n = 36$, chúng ta sẽ có từng giá trị Z khác nhau.
 - Mẫu 1 cho $\bar{X}_1 = 166.5 \rightarrow Z = \frac{166.5-166}{6/6} = 0.5$
 - Mẫu 2 cho $\bar{X}_2 = 168.2 \rightarrow Z = \frac{168.2-166}{6/6} = 2.2$
 - Mẫu 3 cho $\bar{X}_3 = 167.8 \rightarrow Z = \frac{167.8-166}{6/6} = 1.8$

Giá Trị Kiểm Định Thống Kê (3)

- Giả thuyết H_0 cho rằng tất cả các quan sát $X_1, X_2, \dots, X_n$ đến từ một **phân phối y** **hệ thống** có trung bình là $\mu = 166$. Nếu mẫu được thu thập ngẫu nhiên, thì chúng ta có điều kiện I.I.D.
- Do đó: $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$ sẽ có cùng phân phối với $\mu = 166$ và $\sigma^2 = 36$ cũng được cho trước biết.
- Theo định lý giới hạn trung tâm, chúng ta biết $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$:

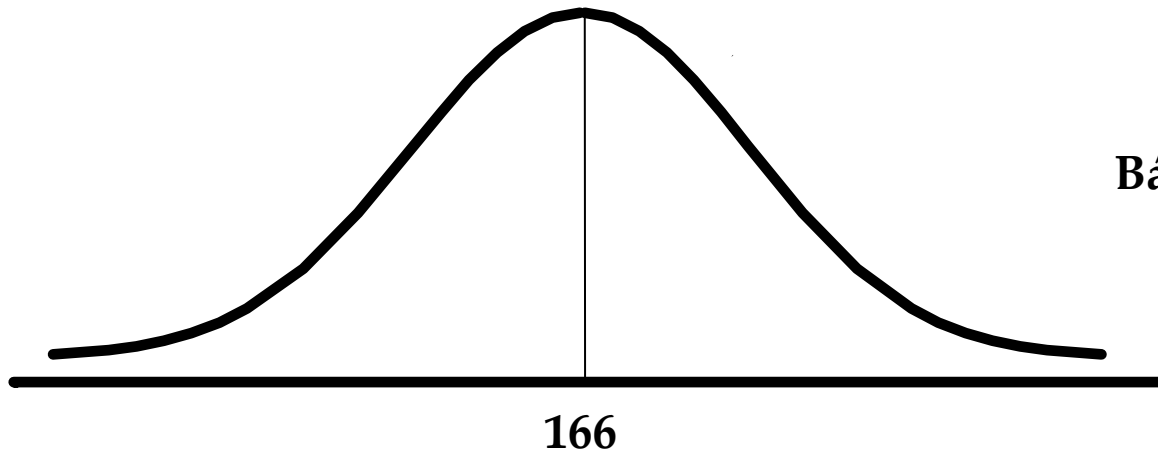
$$Z = \frac{\bar{X} - 166}{\sigma/\sqrt{n}} \sim N(0,1)$$

- Giá trị kiểm định chúng ta sẽ là giá trị Z .

Xác Định Mức Ý Nghĩa Và Khoảng Bác Bỏ (1)

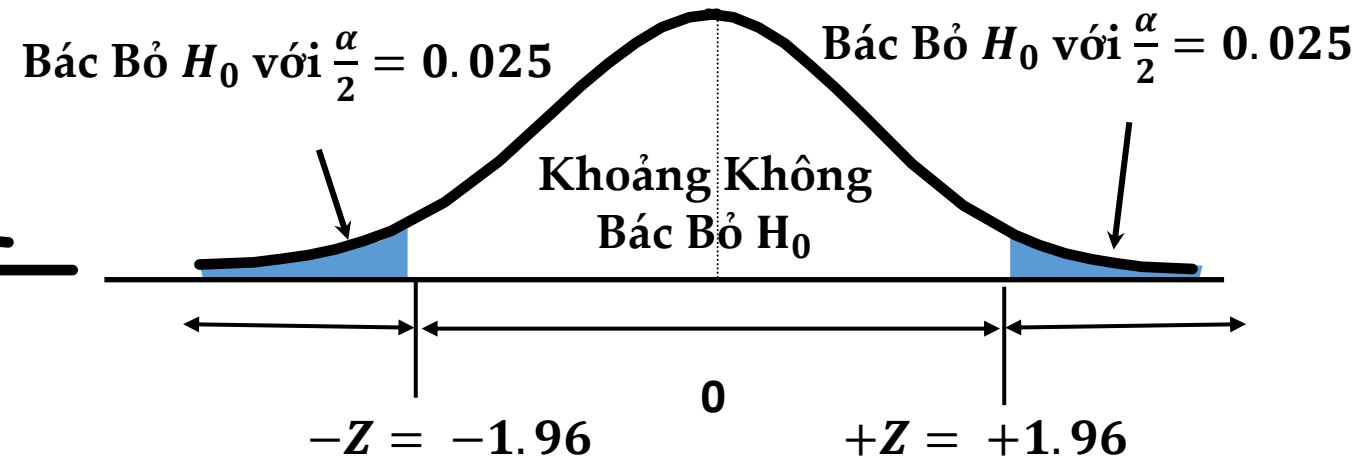
- Các mức ý nghĩa (significance level) phổ biến là 10%, 5% và 1%.
- Mức ý nghĩa sẽ tương ứng với khoảng bác bỏ (rejection region) trong phân phối xác suất:

Phân phối mẫu của $\bar{X}$



$\mu = 166$ khi H_0 xảy ra

Phối chuẩn $Z \sim N(0, 1)$



Xác Định Mức Ý Nghĩa Và Khoảng Bác Bỏ (2)

- Khi mẫu đưa ra các giá trị thống kê kiểm định khác nhau kiểm định:
 - Mẫu 1 cho $\bar{X}_1 = 166.5 \rightarrow Z = \frac{166.5-166}{6/6} = 0.5$
 - Mẫu 2 cho $\bar{X}_2 = 168.2 \rightarrow Z = \frac{168.2-166}{6/6} = 2.2$
 - Mẫu 3 cho $\bar{X}_3 = 167.8 \rightarrow Z = \frac{167.8-166}{6/6} = 1.8$
- Các kết luận sẽ xảy ra như thế nào với từng bằng chứng ở trên với mức ý nghĩa 5%? Chúng ta có thể bác bỏ H_0 hay không?
- Thuật ngữ tiếng Anh: “we cannot/ fail to reject the null hypothesis” or “we can reject the null hypothesis”.

Xác Định Mức Ý Nghĩa Và Khoảng Bỏ (3)

- Nếu chúng ta thay đổi giả thuyết thành:

$$H_0: \mu = 167$$

$$H_1: \mu \neq 167$$

- Kết luận chúng ta sẽ thay đổi thế nào với:
 - Mẫu 1 cho $\bar{X}_1 = 166.5$?
 - Mẫu 2 cho $\bar{X}_2 = 168.2$?
 - Mẫu 3 cho $\bar{X}_3 = 167.8$?

Giả Thích Mức Ý Nghĩa

- **Mức ý nghĩa (significance level)** tương ứng với **sai lầm loại I (type I error)**.
- Trong kiểm định, ta không thể giảm cả 2 loại sai lầm cùng 1 lúc. Ta ấn định trước 1 loại sai lầm và tìm giá trị kiểm định có sai lầm loại kia nhỏ nhất có thể.
- Sai lầm loại I nghiêm trọng hơn, nên được chọn ở mức nhỏ nhất tương đối.

	H_0 đúng	H_1 đúng
Bác bỏ H_0	Sai lầm loại I (Type I error) $P(\text{Bác bỏ} H_0 \text{ đúng}) = \alpha$	Kết quả chính xác $P(\text{Bác bỏ} H_0 \text{ sai}) = 1 - \alpha$
Không bác bỏ H_0	Kết quả chính xác $P(\text{Không bác bỏ} H_0 \text{ đúng}) = 1 - \beta$	Sai lầm loại II (Type II error) $P(\text{Không bác bỏ} H_0 \text{ sai}) = \beta$

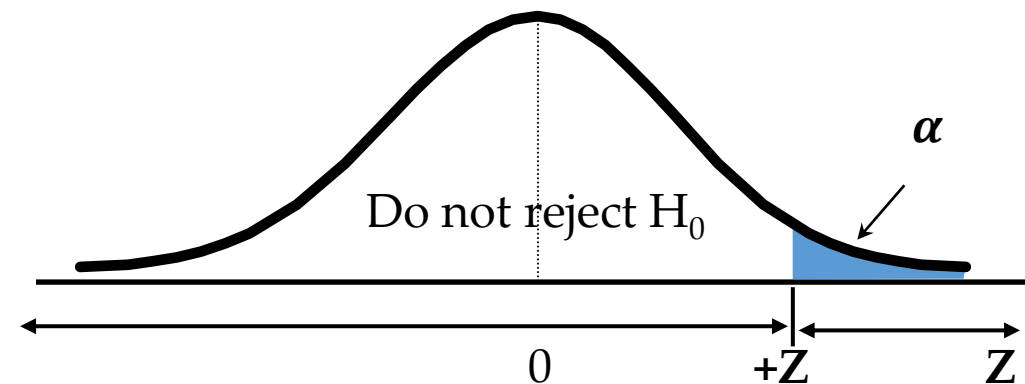
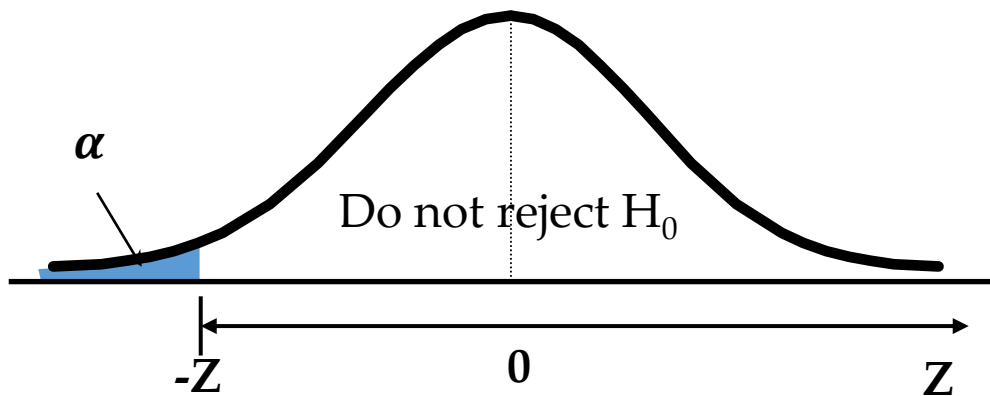
Kiểm Định Một Phía

- Kiểm định giả thuyết thống kê một phía:

$$H_0: \mu \geq 166 \text{ vs } H_1: \mu < 166$$

$$H_0: \mu \leq 166 \text{ vs } H_1: \mu > 166$$

- Khoảng bác bỏ sẽ khác tùy vào giả thuyết dấu $\leq$ hay $\geq$ của H_0



Bảng Giá Trị Tới Hạn Cho Phân Phối Chuẩn

Mức ý nghĩa (α)	Hai Phía	Một Phía, Đuôi-Trên	Một Phía, Đuôi Dưới
	Giá Trị Tới Hạn (Critical Values)		
10%	1.645	1.280	-1.280
5%	1.960	1.645	-1.645
1%	2.580	2.330	-2.330

- Chú ý: các giá trị này lấy ra từ bảng giá trị hàm phân phối chuẩn hoặc sử dụng hàm Excel `NORM.INV()`.

Cách Hiểu Kết Quả Kiểm Định

- Cho ta biết mẫu đưa đến một giá trị “tổng hợp”, giá trị “tổng hợp” có khả năng thấp hay cao được tạo ra bởi phân phối dưới điều kiện giả thuyết vô hiệu đúng.
- Đây là kết quả mang tính thống kê.
- Cách sử dụng kết quả mang tính thống kê để bác bỏ hay chứng minh các lý thuyết, giả thuyết khoa học. Trong thực tế việc sử dụng phương pháp kiểm định thống kê là vấn đề rất gây tranh cãi.

Tổng Kết Ngắn

Các bước của kiểm định giả thuyết:

- ❑ Xác lập giả thuyết vô hiệu (H_0) và giả thuyết thay thế (H_1).
- ❑ Tính toán giá trị kiểm định thống kê (the test statistic).
- ❑ Chọn mức ý nghĩa thống kê (statistical significance) và khoảng bác bỏ (rejection region).
- ❑ So sánh và đưa ra kết luận.

Kiểm Định Giả Thuyết Khi Không Biết Phương Sai Tổng Thể σ^2

- Trường hợp σ^2 biết trước không thực tế, chúng ta quay lại trường hợp σ^2 là ẩn số.
- Khi đó, giá trị kiểm định sẽ sử dụng ước lượng của phương sai và theo phân phối t :

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n - 1)$$

- Các bước còn lại được thực hiện tương tự trường hợp σ^2 biết trước. Lưu ý: S^2 là ước lượng không chệch của σ^2 .
- Đây thường được là **t -test** hay **kiểm định t** .

Kiểm Định t (1)

- Cho biết:

$$H_0: \mu = 166$$

$$H_1: \mu \neq 166$$

- Thu thập một mẫu gồm 25 nam; giá trị trung bình và phương sai mẫu lần lượt là $\bar{X} = 168.4$ và $S^2 = 36$. Hãy kiểm định thống kê với mức ý nghĩa 5% và 1%.

Kiểm Định t (2)

- Tính giá trị kiểm định:

$$t = \frac{168.4 - 166}{6/\sqrt{25}} = 2$$

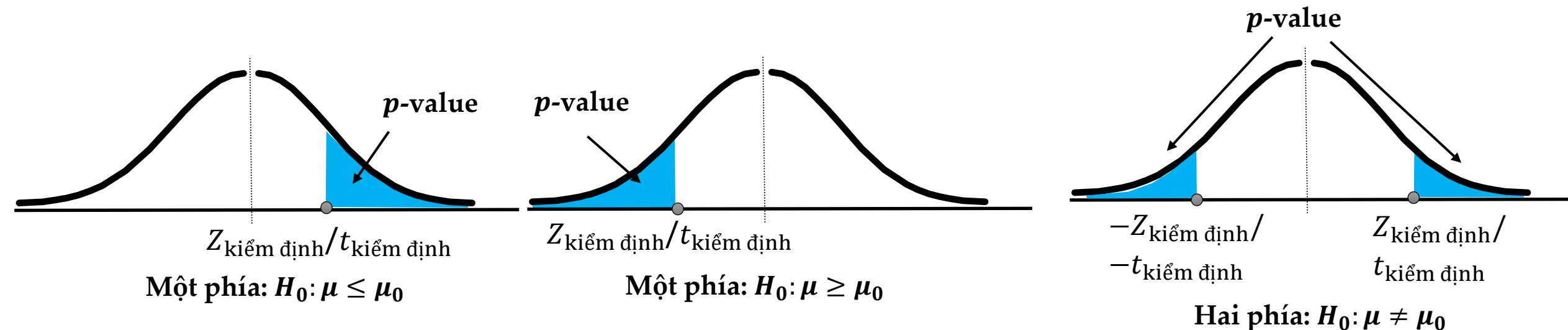
- Sử dụng bảng giá trị hoặc hàm Excel `T.INV()` (cho 24 bậc tự do), chúng ta có giá trị tới hạn ở mức 5% và 1% lần lượt là ± 2.064 và ± 2.797 .
- Kết luận?

p -value Trong Kiểm Định Giả Thuyết

- p -value cho kiểm định t từ mẫu dữ liệu quan sát được là giá trị α_0 nhỏ nhất mà chúng ta có thể bác bỏ giả thuyết vô hiệu H_0 tại mức ý nghĩa α_0 .
- Chúng ta có thể tính toán p -value sử dụng bảng tính hoặc hàm số Excel `T.DIST()`.
- p -value “nhạy cảm” với độ lớn của mẫu.

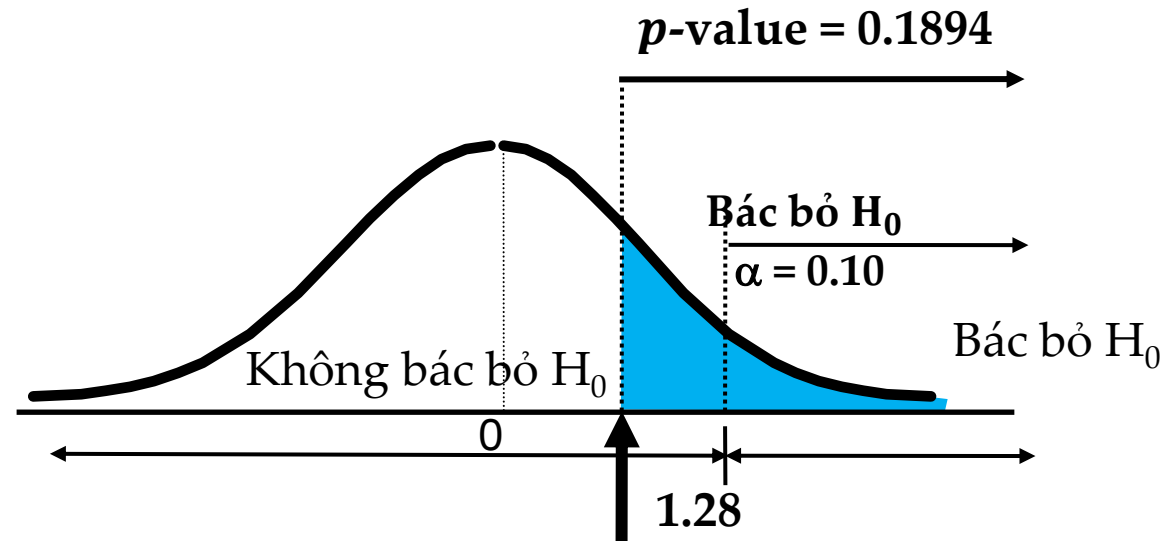
Cách tính p -value

- Bước 1: tính giá trị kiểm định $Z_{\text{kiểm định}}$ hoặc $t_{\text{kiểm định}}$.
- Bước 2: vẽ hình cho kiểm định 1 phía hoặc 2 phía: p -value là diện tích hình màu xanh.
- Bước 3: tra bảng hoặc sử dụng hàm `NORM.DIST()` hoặc `T.DIST()`



Ví Dụ:

- Kiểm định giả thuyết với mà giá trị $Z = 1.28$, xác định p -value của kiểm định.
- Chúng ta có bác bỏ H_0 ở mức 10% không?
- Kết luận thay đổi thế nào khi $Z = 1.8$.



Thực Hành:

- Cho biết:

$$H_0: \mu = 166$$

$$H_1: \mu \neq 166$$

1. Nếu biết $\sigma = 8$ và $n = 16$, hãy tính p -value trong các trường hợp sau:
 - a) Mẫu 1: $\bar{X}_1 = 166.5$
 - b) Mẫu 2: $\bar{X}_2 = 165$
 - c) Mẫu 3: $\bar{X}_3 = 170$
2. p -value thay đổi thế nào biết $\sigma = 6$ và $n = 25$.

Tổng Kết

- Kiểm định giả thuyết thống kê là gì?
- Các bước kiểm định gồm có gì?
- Khi p -value thấp thì bác bỏ hay không bác bỏ H_0 ?