

### Giá trị xác suất của kiểm định: giá trị P (P-value):

Chúng ta nhìn lại bài toán kiểm định 2 phía:  $H_0: a = a_0$ ;  $H_1: a \neq a_0$  với trường hợp  $n \geq 30$  và chưa biết phương sai tổng thể. Giả sử mức ý nghĩa đang được xem xét là  $\alpha_1 = 0,05$  thì  $z_{\alpha_1} = 1,96$  và miền bác bỏ tương ứng là  $W_{\alpha_1} = (-\infty; -1,96) \cup (1,96; +\infty)$ .

Nếu từ một mẫu cụ thể ta tính được  $z_{qsA} = 2,0 \in W_{\alpha_1}$  thì giả thiết  $H_0$  tương ứng bị bác bỏ. Giả thiết từ một mẫu cụ thể khác ta tính được  $z_{qsB} = 10$  chẳng hạn thì giả thiết  $H_0$  cũng bị bác bỏ. Ta nhận thấy việc bác bỏ  $H_0$  trong trường hợp mẫu sau có vẻ “thuyết phục” hơn.

Mặt khác, nếu thay đổi mức ý nghĩa đang được xem xét thành  $\alpha_2 = 0,02$  thì  $z_{\alpha_2} = 2,33$  và miền bác bỏ tương ứng là  $W_{\alpha_2} = (-\infty; -2,33) \cup (2,33; +\infty)$ . Lúc này ta vẫn bác bỏ  $H_0$  nếu  $z_{qsB} = 10$  nhưng lại phải chấp nhận  $H_0$  nếu dùng  $z_{qsA} = 2,0$ .

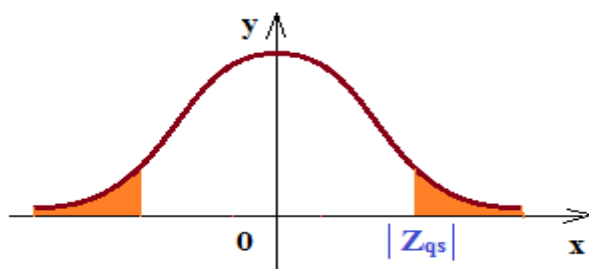
Qua đó ta thấy việc bác bỏ  $H_0$  với  $z_{qsB} = 10$  khá thuyết phục nhưng việc bác bỏ  $H_0$  với  $z_{qsA} = 2,0$  lại ít thuyết phục hơn.

Quá trình kiểm định như trên được gọi là kiểm định theo cách tiếp cận cổ điển. Bây giờ ta tìm hiểu một cách tiếp cận khác bài toán kiểm định. Thay vì kiểm định giả thiết với một mức ý nghĩa  $\alpha$  định trước thì người ta cho rằng sau khi định rõ các giả thiết kiểm định  $H_0$  và giả thiết đối  $H_1$ , ta thu thập các số liệu mẫu và xác định **mức độ khẳng định việc bác bỏ giả thiết  $H_0$** . Mức độ khẳng định này thường được gọi là giá trị xác suất P hay **P-value**.

Ta nói rằng mức ý nghĩa nhỏ nhất tại đó giả thiết  $H_0$  bị bác bỏ được gọi là giá trị P kết hợp với mẫu quan sát được. Người ta còn gọi giá trị P là *mức ý nghĩa quan sát*, nó cho biết xác suất mắc sai lầm loại I tối đa khi bác bỏ giả thiết  $H_0$  với một mẫu quan sát cụ thể.

Xét bài toán kiểm định trung bình tổng thể trong trường hợp mẫu lớn ( $n \geq 30$ ) và chưa biết phương sai tổng thể. Tiêu chuẩn kiểm định là thống kê  $Z = \frac{\bar{X} - a_0}{s} \sqrt{n} \sim N(0,1)$  nếu chấp nhận giả thiết  $H_0$ : “ $a = a_0$ ” đúng.

a) Giả thiết kiểm định  $H_0$ : “ $a = a_0$ ”.  $H_1$ : “ $a \neq a_0$ ”.

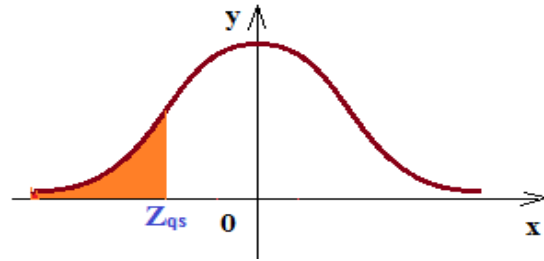


$$\begin{aligned} P\text{-value} &= 2 \times P(Z > |Z_{qs}|) \\ &= 2 \times [0,5 - \Phi(|Z_{qs}|)] = 1 - \Phi(|Z_{qs}|) \end{aligned}$$

VD: Nếu  $z_{qs} = 2,01 \Rightarrow P\text{-value} = 1 - \Phi(2,01) = 4,44 \%$

b) Giả thiết kiểm định  $H_0 : "a = a_0"$ .  $H_1: "a < a_0"$ .

$$P\text{-value} = P(Z < Z_{qs}) = 0,5 - \Phi(Z_{qs})$$



\* Nếu  $z_{qs} = -2,01 \Rightarrow P\text{-value} = 0,5 - \Phi(-2,01) = 0,5 - 0,4778 = 2,22\%$

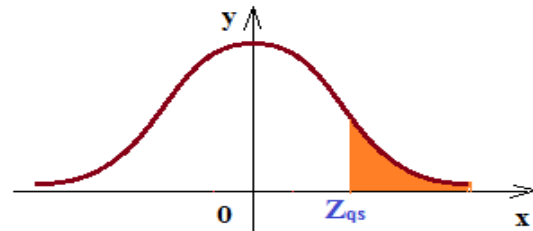
\* Nếu  $z_{qs} = 2,01 \Rightarrow P\text{-value} = 0,5 + \Phi(2,01) = 97,78\%$

c) Giả thiết kiểm định  $H_0 : "a = a_0"$ .  $H_1: "a > a_0"$ .

$$P\text{-value} = P(Z > Z_{qs}) = 0,5 - \Phi(Z_{qs})$$

Nếu  $z_{qs} = 2,01 \Rightarrow$

$$\Rightarrow P\text{-value} = 0,5 - \Phi(2,01) = 2,22\%$$



Trong VD ở mục a)  $P\text{-value} = 4,44\%$  tức là giả thiết  $H_0$  sẽ bị bác bỏ khi mức ý nghĩa  $\alpha$  được yêu cầu trong bài toán lớn hơn  $4,44\%$ . Nếu mức ý nghĩa  $\alpha$  trong bài nhỏ hơn  $4,44\%$  thì ta phải công nhận giả thiết  $H_0$ .

Các phần mềm thống kê hiện nay đều đưa ra  $P\text{-value}$  cho mỗi bài toán kiểm định để độc giả tự đánh giá kết quả. Ý tưởng chủ đạo là  $P\text{-value}$  càng nhỏ thì càng bác bỏ giả thiết mạnh,  $P\text{-value}$  càng lớn thì càng chấp nhận giả thiết mạnh.

Thông thường người ta tiến hành kiểm định theo nguyên tắc:

- Nếu  $0,1 \leq P\text{-value}$ : ta thường thừa nhận  $H_0$ .
- Nếu  $0,05 \leq P\text{-value} < 0,1$ : cân nhắc cẩn thận trước khi bác bỏ  $H_0$ .
- Nếu  $0,01 \leq P\text{-value} < 0,05$ : nghiêng về hướng bác bỏ  $H_0$  nhiều hơn.
- Nếu  $0,001 \leq P\text{-value} < 0,01$ : có thể ít băn khoăn khi bác bỏ  $H_0$ .
- Nếu  $P\text{-value} < 0,001$ : có thể hoàn toàn yên tâm khi bác bỏ  $H_0$ .

Trong các VD ở mục a) và c) ta tính được giá trị P tương đối nhỏ nên kết luận nghiêng về xu hướng bác bỏ  $H_0$ . Còn VD ở mục b) khi  $z_{qs} = -2,01$  thì giá trị P quá lớn nên ta luôn chấp nhận  $H_0$ .

Những trường hợp tiêu chuẩn kiểm định có các phân phối khác như phân phối Student, phân phối Khi bình phương, phân phối Fisher..., chúng ta có thể tìm giá trị P tương ứng với giá trị quan sát được lấy từ mẫu. Tuy nhiên do hạn chế của phần tra bảng nên chúng ta không trình bày VD ở đây, để tính toán chúng, sinh viên có thể sử dụng các hàm tương ứng trong Excel.